



**UNIVERSIDADE FEDERAL DA PARAÍBA**  
**CENTRO DE CIÊNCIAS HUMANAS, LETRAS E ARTES**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA**  
**MESTRADO EM LINGUÍSTICA**

**MAYLTON DA SILVA FERNANDES**

**ORGANIZAÇÃO LEXICAL EM BILÍNGUES: UMA INVESTIGAÇÃO SOBRE OS  
EFEITOS DA PROFICIÊNCIA EM L2 NA DIRECIONALIDADE DO *PRIMING*  
ENCOBERTO DE PALAVRAS COGNATAS E NÃO COGNATAS**

**JOÃO PESSOA**

**2022**

MAYLTON DA SILVA FERNANDES

ORGANIZAÇÃO LEXICAL EM BILÍNGUES: UMA INVESTIGAÇÃO SOBRE OS  
EFEITOS DA PROFICIÊNCIA EM L2 NA DIRECIONALIDADE DO *PRIMING*  
ENCOBERTO DE PALAVRAS COGNATAS E NÃO COGNATAS

Dissertação apresentada ao Programa de Pós-graduação em Linguística do Centro de Ciências Humanas, Letras e Artes da UFPB - Universidade Federal da Paraíba, como um requisito parcial para a obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Márcio Martins Leitão.  
Coorientador: Prof. Dr. Gustavo Lopez Estivalet.

JOÃO PESSOA

2022

**Catálogo na publicação**  
**Seção de Catalogação e Classificação**

F363o Fernandes, Maylton da Silva.

Organização lexical em bilíngues: uma investigação sobre os efeitos da proficiência em L2 na direcionalidade do priming encoberto de palavras cognatas e não cognatas / Maylton da Silva Fernandes. - João Pessoa, 2023.

102 f. : il.

Orientação: Márcio Martins Leitão.

Coorientação: Gustavo Lopez Estivalet.

Dissertação (Mestrado) - UFPB/CCHLA.

1. Bilinguismo. 2. Processamento lexical. 3. Priming encoberto. I. Leitão, Márcio Martins. II. Estivalet, Gustavo Lopez. III. Título.

UFPB/BC

CDU 81'246.2(043)

ATA DE DEFESA DE DISSERTAÇÃO DE  
MAYLTON DA SILVA FERNANDES

Aos vinte e oito dias do mês de fevereiro de dois mil e vinte e dois (28/02/2022), às catorze hora, realizou-se, via Plataforma Zoom, a sessão pública de defesa de Dissertação intitulada “**ORGANIZAÇÃO LEXICAL EM BILÍNGUES: UMA INVESTIGAÇÃO SOBRE OS EFEITOS DA PROFICIÊNCIA EM L2 NA DIRECIONALIDADE DO PRIMING ENCOBERTO DE PALAVRAS COGNATAS E NÃO COGNATAS**”, apresentada pelo(a) mestrando(a) **MAYLTON DA SILVA FERNANDES**, Licenciado(a) em **Letras** pelo(a) **Universidade Estadual da Paraíba - UEPB**, que concluiu os créditos para obtenção do título de MESTRE(A) EM LINGUÍSTICA, área de concentração **Teoria e Análise Linguística**, segundo encaminhamento do(a) Prof(a). Dr(a). Jan Edson Rodrigues Leite, Coordenador(a) do Programa de Pós-Graduação em Linguística da UFPB e segundo registros constantes nos arquivos da Secretaria da Coordenação do Programa. O(A) Prof(a). Dr(a). Márcio Martins Leitão (PROLING - UFPB), na qualidade de orientador(a), presidiu a Banca Examinadora da qual fizeram parte os(a)s Professores(as) Doutores(as) Gustavo Lopez Estivalet (Coorientador/PROLING-UFPB), Rosana Costa de Oliveira (Examinadora/PROLING-UFPB) e Elena Ortiz Preuss (Examinadora/UFG). Dando início aos trabalhos, o(a) senhor(a) Presidente Prof(a). Dr(a). Márcio Martins Leitão convidou os membros da Banca Examinadora para compor a mesa. Em seguida, foi concedida a palavra ao(à) Mestrando(a) para apresentar uma síntese de sua Dissertação, após o que foi arguido(a) pelos membros da banca Examinadora. Encerrando os trabalhos de arguição os examinadores deram o parecer final sobre a Dissertação, ao qual foi atribuído o conceito **APROVADO**. Proclamados os resultados pelo(a) professor(a) Dr(a). Márcio Martins Leitão, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar a presente ata foi lavrada e assinada por todos os membros da Banca Examinadora. João Pessoa, 28 de fevereiro de 2022.

Observações

**A banca recomenda a publicação tendo em vista a qualidade do trabalho.**



Prof(a). Dr(a). Márcio Martins Leitão  
(Presidente da Banca Examinadora) (Coorientador)



Prof(a). Dr(a). Gustavo Lopez Estivalet



Prof(a). Dr(a). Rosana Costa de Oliveira  
(Examinadora)



Prof(a). Dr(a). Elena Ortiz Preuss  
(Examinadora)

## RESUMO

Há um interesse na Psicolinguística em compreender como o léxico ou léxicos são representados e acessados na memória bilíngue. Duñabeitia, Perea e Carreiras (2010) afirma em seu estudo que bilíngues simultâneos, cujas duas ou mais línguas foram adquiridas ao mesmo tempo ainda durante a infância, processam tanto a L1 quanto a L2, ou ainda, as duas L1, da mesma forma independente da direcionalidade do estímulo, o que não seria possível para bilíngues tardios cuja L2 fora adquirida após a infância/adolescência. Também, parece existir uma assimetria no acesso lexical bilíngue, no que diz respeito à direção do estímulo linguístico recebido. Conforme De Groot e Nas (1991), existe facilitação no acesso de palavras da L2 quando uma outra relacionada na L1 é apresentada primeiro, porém o mesmo ocorre de forma muito menor mesmo com falantes com alto nível de proficiência, ou pode até não ocorrer, quando uma palavra da L2 é exibida antes de uma palavra da L1. Este trabalho teve como objetivo investigar o processamento de palavras cognatas e não-cognatas considerando a direcionalidade do input linguístico e o nível de proficiência por meio do paradigma de *priming* encoberto, assim como a organização lexical em bilíngues tardios, falantes de Português Brasileiro como língua materna e falantes de Inglês como língua estrangeira, nos níveis intermediário e avançado. Para tal, selecionamos 33 participantes, 20 com nível intermediário de proficiência, 13 com nível avançado, por meio de um teste de nivelamento (NATION, 1990) com tempo máximo de realização de 10 minutos, conforme indicação de Souza e Soares-Silva (2015), assim como também aplicamos um questionário sociolinguístico. Em seguida, aplicamos uma tarefa experimental de *priming* encoberto, tanto na direção L2-L1, quanto na direção L1-L2 envolvendo palavras distribuídas entre cognatas e não-cognatas, além de pseudopalavras usadas como distratoras pseudorandomizadas através do programa *Mix* (VAN CASTEREN & DAVIS, 2006). Conforme análise estatística realizada a partir do tempo de reação dos participantes, foi observado um efeito principal para o nível de proficiência, em que os bilíngues com alto nível de proficiência apresentaram uma média no tempo de reação menor que os do grupo intermediário. Também foi observado efeito significativo em relação às condições de palavras experimentais e pseudopalavras, sendo o tempo de reação para as palavras cognatas e não cognatas menor que o tempo para pseudopalavras. Além disso, observamos efeitos significativos em ambas as direções de *priming*, bem como foi observado efeitos de *priming* em bilíngues tardios intermediários. De modo geral, os resultados encontrados vão de acordo com a assimetria de *priming* mencionada por De Groot e Nas (1991), assim como corroboram com a afirmação de Duñabeitia, Perea e Carreiras (2010), que diz que bilíngues não balanceados e não simultâneos processem a L2 de forma diferente da L1. Mesmo assim, os dados obtidos trazem nova luz a discussão sobre o processamento lexical em bilíngues tardios a luz de novos modelos de processamento linguístico.

**Palavras-chave:** Bilinguismo, Processamento Lexical, Priming Encoberto.

## ABSTRACT

There is an interest in Psycholinguistics in understanding how the lexicon or lexicons are represented and accessed in bilingual memory. Duñabeitia, Perea and Carreiras (2010) states in their study that simultaneous bilinguals, whose two or more languages were acquired at the same time during childhood, process both L1 and L2, or even both L1, in the same way regardless the direction of the stimulus, which would not be possible for late bilinguals whose L2 was acquired after childhood/adolescence. Also, there appears to be an asymmetry in bilingual lexical access regarding the direction of the linguistic stimulus. According to De Groot and Nas (1991), it is easier to access words from L2 when another one in L1 is presented first, but the same occurs much less even with speakers with a high level of proficiency, or may not even show, when an L2 word is displayed before an L1 word. This work aimed to investigate the facilitating effects of cognate and non-cognate words considering the direction of the linguistic input using a masked priming paradigm, as well as the lexical organization in late bilinguals, speakers of Brazilian Portuguese as their mother tongue and speakers of English as a foreign language, at intermediate and advanced levels. To this end, we selected 33 participants, 20 of them had an intermediate level of proficiency while the other 13 participants had an advanced level, according a placement test (NATION, 1990). This test was realized with a maximum time of 10 minutes, as indicated by Souza and Soares-Silva (2015), and we also applied a sociolinguistic questionnaire to check the participants' linguistic background. Next, we applied an experimental task of masked priming, both in the L2-L1 direction and in the L1-L2 direction, involving words distributed between cognates and non-cognates, in addition to pseudowords used as pseudorandom distractors through the Mix program (VAN CASTEREN & DAVIS, 2006). According to the statistical analysis carried out based on the reaction time of the participants, a main effect was observed for the proficiency level, in which bilinguals with a higher level of proficiency had an average reaction time lower than those in the intermediate group. A significant effect was also observed in relation to the conditions of experimental words and pseudowords, with the reaction time for cognate and non-cognate words being shorter than the time for pseudowords. Furthermore, we observed significant effects in both priming directions, as well as observed priming effects in intermediate late bilinguals. In general, the results found are in line with the priming asymmetry mentioned by De Groot and Nas (1991), as well as corroborate the statement by Duñabeitia, Perea and Carreiras (2010), which says that unbalanced and non-simultaneous bilinguals process the L2 differently from the L1. Even so, the data obtained bring new light to the discussion about lexical processing in late bilinguals in the light of new models of language processing.

**Keywords:** Bilingualism, Lexical Processing, Masked Priming.

Este trabalho é todo dedicado à minha mãe,  
Claudionoura, que sempre me apoiou e esteve  
ao meu lado.

## AGRADECIMENTOS

A Deus, pela graça diariamente distribuída e que me motiva a ser alguém melhor e a sempre buscar evoluir.

Ao meu orientador, professor Márcio Leitão, por me abrir as portas do LAPROL, pela parceria estabelecida durante esses anos e por todos os conselhos, ensinamentos e confiança em mim depositada.

Ao meu coorientador, professor Gustavo Estivalet, também pela grande parceria, por toda a ajuda na parte estatística e teórica, e por sempre “me incomodar” e me fazer ir além.

Aos meus pais, Claudionoura e José, por sempre acreditarem em mim e me motivarem a correr atrás dos meus sonhos e a nunca desistir.

Aos meus irmãos, Carla, Moacir e Vanderléia, por estarem sempre ao meu lado.

Ao meu amigo Matheus Almeida, por estar presente desde a seleção até a fase de conclusão desta pesquisa. Agradeço imensamente por todos os conselhos, leituras e por sempre ter me estimulado a prosseguir.

Às minhas grandes amigas que a pós-graduação me deu, Célia e Laiane, pelas conversas, almoços, risadas, e lágrimas compartilhadas durante todo esse tempo, tornando todo esse processo mais leve.

Aos meus também amigos da pós Ana Stela, Judithe e Jefferson, por também compartilharem tantos momentos de descontração e alegres.

Aos professores Ferrari e Elena pelas contribuições na qualificação e defesa.

A professora Marta Furtado (*in memoriam*), por ter me incentivado a buscar mais e mais conhecimento.

Aos colegas do LAPROL, pelas discussões, sugestões e críticas durante nossas reuniões.

A todos os professores e amigos que sempre acreditaram em mim.

Aos sujeitos que participaram do experimento.

À Capes, pelo apoio financeiro.

## LISTA DE GRÁFICOS

GRÁFICO 1 – Número total de respostas .....	48
GRÁFICO 2 – Distribuição do número de respostas por participante na escala Likert.....	48
GRÁFICO 3 – Média dos pares de palavras .....	49
GRÁFICO 4 – Distância de Levenshtein Normalizada para os pares de palavras .....	49
GRÁFICO 5 – Médias e desvios-padrão das pontuações da escala Likert dos pares de palavras .....	50
GRÁFICO 6 – Correlação e regressão linear das variáveis média de pontuação da escala Likert em função da Distância de Levenshtein Normalizada. As cores das palavras representam as divisões dos coeficientes da Tabela 3 .....	52
GRÁFICO 7 – Tempos de reação gerais dos participantes nas duas direções de <i>priming</i> .....	64
GRÁFICO 8 – Tempos de reação gerais das palavras cognatas e não cognatas em ambas as direções de <i>priming</i> . .....	65
GRÁFICO 9 – Tempos de reação gerais para a variável relação de <i>priming</i> em ambas as direções. ....	65
GRÁFICO 10 - diferenças nos tempos de reação de acordo com a direção de <i>priming</i> e do tipo de palavra.....	66

## **LISTAS DE FIGURAS**

FIGURA 1 - Representação do continuum do Modo de Língua.....	33
FIGURA 2 - Modelo Hierárquico Revisado.....	35
FIGURA 3 - Modelo de Ativação Interativa Bilíngue Mais (BIA+) .....	38
FIGURA 4 - Representação da arquitetura padrão do Multilink.....	42
FIGURA 5 – Instruções do experimento de julgamento de aceitabilidade no Google Forms .	47
FIGURA 6 - Exemplo de como o VLT é apresentado no Classmarker .....	59

## **LISTA DE ABREVIATURAS E SIGLAS**

**ANOVA** – *Analysis of Variance* (Análise da Variância)

**BIA** – *Bilingual Interactive Activation* (Modelo de Ativação Interativa Bilíngue)

**BIA+** – *Bilingual Interactive Activation Plus* (Modelo de Ativação Interativa Bilíngue Mais)

**DLR** – Distância de Levenshtein Relativa

**MHR** – Modelo Hierárquico Revisado

**NAR** – Níveis de Ativação em Repouso

**TR** - Tempo de reação

**VLT** – *Vocabulary Levels Test* (Teste de Níveis de Vocabulário)

## SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>9</b>
<b>1 PALAVRAS COGNATAS E NÃO COGNATAS: ESTADO DA ARTE.....</b>	<b>14</b>
1.1 BILINGUISMO E LÉXICO MENTAL.....	14
1.2 EFEITOS DE PRIMING.....	17
1.3 ASSIMETRIA NO PRIMING ENCOBERTO E REPRESENTAÇÃO DE COGNATOS NO LÉXICO.....	20
1.4 NÍVEL DE PROFICIÊNCIA E EXPERIÊNCIA LINGUÍSTICA.....	29
<b>2 MODELOS DE REPRESENTAÇÃO LEXICAL BILÍNGUES.....</b>	<b>31</b>
2.1 A HIPÓTESE DO MODO DE LÍNGUA.....	32
2.2 MODELO HIERÁRQUICO REVISADO.....	33
2.3 O MODELO DE ATIVAÇÃO INTERATIVA BILÍNGUE MAIS.....	37
2.4 MODELO MULTILINK.....	40
<b>3 EXPERIMENTO 1: JULGAMENTO DE ACEITABILIDADE DE PALAVRAS COGNATAS E NÃO COGNATAS .....</b>	<b>45</b>
3.1 PARTICIPANTES.....	45
3.2 MATERIAIS.....	46
3.3 PROCEDIMENTOS.....	47
3.4 RESULTADOS E DISCUSSÃO.....	48
<b>4 EXPERIMENTO 2: PRIMING ENCOBERTO DE PALAVRAS COGNATAS E NÃO COGNATAS .....</b>	<b>55</b>
4.1 PARTICIPANTES.....	57
4.2 TESTE DE PROFICIÊNCIA.....	58
4.3 MATERIAIS.....	60
4.4 PROCEDIMENTOS.....	62
4.5 RESULTADOS.....	63
4.6 DISCUSSÃO.....	68
<b>CONSIDERAÇÕES FINAIS.....</b>	<b>73</b>
<b>REFERÊNCIAS .....</b>	<b>75</b>
<b>APÊNDICES.....</b>	<b>81</b>
APÊNDICE A.....	82
APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.....	93
APÊNDICE C – QUESTIONÁRIO SÓCIOLINGUÍSTICO NO GOOGLE FORMS.....	95

## INTRODUÇÃO

A Psicolinguística tem como interesse principal investigar de que maneira o ser humano adquire a linguagem verbal, como a produz e como a compreende (LEITÃO, 2008, p. 220). Dentro desta perspectiva, encontra-se também o interesse em saber como o processamento linguístico ocorre em pessoas que falam apenas uma língua e aquelas que adquiriram duas ou mais línguas, ou seja, os bilíngues. Até a década de 80, as pesquisas voltadas para falantes de uma única língua, os monolíngues, eram prioridade no campo da Psicolinguística Experimental. Hoje em dia, no entanto, diversas investigações envolvendo falantes de mais de uma língua têm sido amplamente desenvolvidas no Brasil e no mundo.

Dentre os muitos aspectos do bilinguismo de interesse dos psicolinguistas, um que está sempre em discussão é a maneira como o léxico é organizado e como o acesso a este ocorre na mente. Como que nós guardamos, acessamos e processamos as palavras armazenadas na nossa mente? Em relação aos bilíngues, como as suas duas (ou mais) línguas são organizadas? Será que elas dividem o mesmo espaço, ou são elas armazenadas em locais diferenciados? Como que elas se relacionam? E quanto aos significados das palavras? As línguas dividem o mesmo significado ou cada uma as armazena separadamente? Diversos modelos e hipóteses foram formulados com o objetivo de explicar tais questionamentos, mas, mesmo que algumas perguntas já tenham sido respondidas, ainda há um longo caminho de descobertas a se percorrer.

O objetivo dessa dissertação é investigar como bilíngues falantes de português brasileiro e inglês como L2 em diferentes níveis de proficiência processam palavras de ambas as línguas. Mais especificamente, buscamos verificar através do paradigma experimental de *priming* encoberto com decisão lexical se há diferenças no processamento de palavras cognatas e não cognatas no léxico mental, assim como se há diferenças na forma como as línguas de um bilíngue acessa os significados e como o nível de proficiência pode influenciar esse acesso.

Concordamos com Lam e Dijkstra (2010) quando eles afirmam que não é surpresa que:

a questão básica da representação e uso das palavras pelos usuários de uma língua tem sido o foco de pesquisas com monolíngues e bilíngues há muito tempo, uma vez que se percebe que elas são os blocos básicos para a construção de sentenças, e que o tempo de reconhecimento de palavras explica a maior parte da variação temporal no processamento linguístico (LAM; DIJKSTRA, 2010, p. 688, tradução do autor).

Entender como as palavras são representadas e acessadas na mente é de grande importância se queremos compreender como sentenças e as línguas de maneira geral são processadas.

No que diz respeito ao bilinguismo, um dos questionamentos que é levantado é se palavras ortograficamente semelhantes entre as duas línguas de uma pessoa são ativadas ao mesmo tempo a partir de um determinado estímulo. Ou seja, se um sujeito falante de português brasileiro e inglês ao ver a palavra “vaso”, ele também ativará não só palavras do português como “caso” e “raso”, mas também palavras do inglês, como “vase”. De acordo com Dijkstra (2005), a resposta para essa pergunta varia basicamente de três formas: se um modelo defende que o acesso às informações lexicais se dá de forma seletiva, então a resposta seria “não”. Por outro lado, caso se assuma um acesso não seletivo para tais informações, a resposta seria “sim”. Ainda, é possível que a resposta para tal pergunta seja “depende”, uma vez que o acesso lexical pode variar entre seletivo e não seletivo com base nas circunstâncias que o envolvem.

As duas primeiras concepções acima têm relação direta com a forma como o léxico bilíngue pode ser organizado. Na primeira, os itens lexicais, ou seja, as palavras, são armazenados na memória separadamente. Sendo assim, o acesso a palavras de uma das línguas não influencia a outra, o que é conhecido como acesso seletivo. Em contrapartida, a segunda concepção afirma que as palavras de ambas as línguas de um bilíngue são armazenadas em um mesmo lugar, e que o acesso a elas acontece de forma não seletiva, ou seja, pode haver influência entre itens de línguas distintas.

Meyer, Schvaneveldt e Ruddy (1972) apontam ainda a possibilidade de que o léxico mental pode ser organizado por localizações distintas e organizadas semanticamente. Dessa forma, palavras como “pão” e “manteiga” podem estar juntas, enquanto palavras sem nenhuma associação, como “enfermeira” e “manteiga”, estão distantes uma da outra. Com base nisso, quando uma palavra é ativada, diversas outras que estão localizadas próximas a ela também são ativadas. Dessa forma, se a palavra “pão” for ativada, manteiga também será, o que facilitará o seu reconhecimento caso ela seja requerida.

Com isso, as pesquisas sobre o acesso lexical em bilíngues têm também o interesse em verificar se a exposição a informação linguística, ou seja, se o contexto linguístico pode modular a ativação lexical de maneira que o processamento das palavras ocorra mais rapidamente (GROSJEAN, 1985). Por exemplo, a apresentação de um experimento em língua inglesa pode dar pistas para o participante bilíngue de que a tarefa será naquela língua. Isso pode fazer com que o sistema de identificação de palavras do sujeito utilize essas pistas contextuais para suprimir a ativação da sua outra língua. Além disso, menciona Dijkstra (2005), caso esta informação linguística seja carregada apenas pelo próprio item lexical, questiona-se se a mesma fica disponível a tempo para afetar o sistema de reconhecimento de palavras ou se ela só chega após a palavra já ter sido reconhecida.

O *priming* encoberto com decisão lexical trata-se de um paradigma experimental que busca verificar a facilitação de um estímulo apresentado rapidamente sobre um outro posterior (FORSTER; DAVIS, 1984). Nos estudos com bilíngues, é possível observar que este paradigma associado à tradução de palavras entre línguas pode ser constantemente observado na literatura (DUYCK; WARLOP, 2009; WANG, 2013; WEN; VAN HEUVEN, 2017, dentre outros). Um dos pontos analisados é que palavras cognatas normalmente apresentam um efeito de facilitação maior do que palavras não cognatas.

De acordo com Duñabeitia, Perea e Carreiras (2010), esse resultado poderia ser decorrente do fato de que a grande maioria dos estudos realizados com *priming* de cognatos e não-cognatos foram feitos com sujeitos que adquiriram a L2 após a L1, ou seja, bilíngues tardios, e que apresentam maior domínio sobre a L1 em comparação com a L2. Além disso, eles ainda adicionam que quando o *priming* é realizado na direção L1-L2, o efeito de facilitação normalmente é maior do que na direção L2-L1.

Há um consenso entre os diversos modelos de representação e processamento lexical que bilíngues em níveis baixos de proficiência levam mais tempo para acessar e processar palavras. O modelo Multilink de Dijkstra *et al.* (2019) destaca que os níveis de ativação em repouso das palavras da L2 de um iniciante/intermediário são mais baixos do que os da língua materna. Isso acontece porque a L1 é frequentemente usada em contextos ativos de linguagem, enquanto a L2 é utilizada com menos frequência. Dessa forma, recuperar os itens lexicais da L2 leva um pouco mais de tempo em níveis iniciais, uma vez que eles estão menos ativos na memória. Em sujeitos avançados e com frequente uso da língua, essa diferença entre as línguas passa a ser mínima.

Evidências indicam que existe uma assimetria no *priming* de palavras não-cognatas com bilíngues (DIMITROPOULOU; DUÑABEITIA; CARREIRAS, 2011), o que reforça o exposto acima. Essa assimetria diz respeito à língua em que o *priming* é exibido. Quando o experimento é realizado na direção L1-L2, efeitos de facilitação podem ser obtidos, mas quando ele é feito na direção L2-L1, pouco ou nenhum efeito de facilitação pode ser observado.

Por este motivo, Duñabeitia, Perea e Carreiras (2010) advogam em seu estudo que bilíngues tardios e não balanceados, ou seja, que têm mais domínio sobre uma das línguas, não conseguiriam apresentar efeitos de facilitação na L2 como na L1, a não ser que estivessem em níveis muito altos de proficiência. Contudo, estudos como o de basnight-Brown e Altarriba (2007), assim como o de Duyck e Warlop (2009), mostram que é possível obter efeitos de facilitação mesmo com populações de diferentes níveis de proficiência nas duas direções de *priming*.

Assim, como fora mencionado anteriormente, esta pesquisa se propõe verificar se bilíngues tardios e não-balanceados, conforme a definição de Duñabeitia, Perea e Carreiras (2010), em níveis intermediário e avançado de proficiência apresentarão efeitos de facilitação na tradução de *primings* encobertos de palavras cognatas e não cognatas, tanto na direção L1-L2, como na direção L2-L1.

Esperamos que os bilíngues intermediários apresentem efeitos de facilitação de palavras cognatas tanto na direção L1-L2, quanto na direção L2-L1, mas pouco ou nenhum efeito de facilitação significativo na direção L2-L1 para palavras não cognatas, tendo efeito significativo apenas na direção L1-L2. Por sua vez, acreditamos que os participantes em nível avançado irão apresentar efeito de facilitação em ambas as direções de *priming*, tanto para palavras cognatas como não cognatas, não existindo diferença significativa entre elas.

Além disso, também realizamos um Julgamento de Aceitabilidade (OLIVEIRA; SÁ, 2013) com palavras cognatas, com participantes com pouco ou nenhum conhecimento da língua inglesa, com o objetivo de avaliar o status cognato das palavras que posteriormente seriam usadas no experimento de *priming* encoberto. Ainda, aproveitamos a oportunidade para investigar se apenas a representação ortográfica facilita a identificação das palavras cognatas por pessoas que tem pouco ou nenhum conhecimento da língua inglesa.

Esta dissertação está organizada em quatro capítulos.

No capítulo 1, iremos revisar a origem e eficácia do paradigma de *priming* encoberto nos estudos sobre o processamento lexical bilíngue, bem como iremos fazer um levantamento dos estudos envolvendo o *priming* encoberto de palavras cognatas e não cognatas.

No capítulo 2, trataremos sobre alguns modelos de representação e processamento lexical bilíngues. Mais especificamente, discutimos as principais características da Hipótese do Modo de Língua, uma seria mais geral sobre o processamento bilíngue, assim como descreveremos o Modelo Hierárquico Revisado, modelo de Ativação Interativa Bilíngue Mais e o modelo Multilink, os quais nos ajudam a compreender melhor a maneira como o léxico mental pode ser organizado.

No capítulo 3, descreveremos um estudo experimental utilizando o julgamento de aceitabilidade de palavras, que teve como principal objetivo verificar como palavras cognatas podem ser reconhecidas por falantes de uma outra língua mesmo com pouco ou nenhum conhecimento sobre a língua estrangeira. Também, comparamos a aceitação dos participantes com a distância de Levenshtein Relativa das palavras investigadas com o objetivo de verificar se há relação entre a semelhança ortográfica entre os pares cognatos e a média de aceitação no julgamento.

No capítulo 4 descrevemos o estudo experimental envolvendo o paradigma de *priming* encoberto com decisão lexical de palavras cognatas e não cognatas, as hipóteses, resultados e análise estatística dos dados obtidos, os relacionando a nossa pergunta de pesquisa e o marco teórico.

Por fim, apresentaremos as considerações finais sobre a pesquisa, as referências bibliográficas, os apêndices, contendo as palavras usadas no experimento descrito no capítulo 4 e os anexos.

## 1 PALAVRAS COGNATAS E NÃO COGNATAS: ESTADO DA ARTE

### 1.1 BILINGUISMO E LÉXICO MENTAL

O termo bilinguismo por si só levanta diversas discussões entre pesquisadores, em razão da complexidade de se definir quem é ou não bilíngue. Em sua forma mais simplória, podemos definir bilíngue alguém que conhece duas línguas (VALDEZ; FIGUEORA, 1994). Contudo, afirmar que alguém que conhece duas línguas não nos permite concluir “o quão bilíngue” essa pessoa é. Ora, existem pessoas que apresentam alto nível de proficiência em duas línguas, enquanto outras apresentam um domínio maior sobre uma das duas línguas. Por isso é importante levar em consideração que existem níveis diferentes de bilinguismo ao se usar esse termo.

Conforme Gottardo e Grant (2008), o que se deve considerar ao caracterizar bilíngues não é necessariamente o quanto ele conhece das línguas, mas “quando” ele as adquiriu. Mais que isso, é importante analisar o porquê, para quê e como cada indivíduo aprende uma língua. Dentre as diversas definições disponíveis na literatura, consideraremos nesse trabalho os termos bilíngues simultâneos e não-simultâneos. Com isso em mente, são considerados bilíngues simultâneos aqueles que foram expostos a duas línguas desde o nascimento (DE HOUWER, 2005), mas o termo também é aplicado àqueles que adquiriram a segunda língua (L2) entre os primeiros 02 ou 03 anos de idade (COSTA; SEBASTIÃO-GALLÊS, 2014). Casos de bilinguismo simultâneo não são comuns, segundo Gottardo e Grant (2008), assim como não é comum que haja proficiência equivalente em ambas as línguas.

Por sua vez, os bilíngues suscetíveis ou bilíngues sequenciais são aqueles que adquiriram a L2 após ter adquirido a sua primeira língua (L1, GOTTARDO; GRANT, 2008, COSTA; SEBASTIÃO-GALLÊS, 2014). Dentro dessa perspectiva, é importante considerar o quão depois da aquisição da L1 ocorreu a aquisição da L2. Conforme os autores, bilíngues que adquirem a L2 até os 05 anos de idade são considerados bilíngues precoces. Já aqueles que adquirem a L2 após esse período são chamados de bilíngues tardios. Isso se deve a estudos que indicam que o ser humano tem maior facilidade de adquirir uma língua até os 05 anos de idade, período esse em que o cérebro ainda está em processo de lateralização. Contudo, essa demarcação ainda gera discussões por linguistas e neurocientistas.

As palavras de uma língua, as quais consideramos aqui como itens lexicais, são armazenadas no cérebro. O léxico mental corresponde ao banco de dados que armazena esses itens lexical na mente de um falante, como um “dicionário”. Já o acesso lexical é a atividade de

se entrar nesse banco de palavras e recuperar dele os dados necessários para o processamento linguístico, que por sua vez, vão desde informações fonológicas e ortográficas até as representações semânticas. Sendo assim, o reconhecimento de palavras é o processo realizado para se buscar essas informações no léxico com base em um estímulo linguístico (DIJKSTRA, 2005). Em outras palavras, é através do processo de reconhecimento de palavras que conseguimos identificar a palavra “casa” ao visualizarmos o conjunto de letras “c”, “a”, “s” e “a”. A princípio, essas informações representadas no léxico, ou seja, as informações ortográficas, fonológicas e semânticas, podem ser acessadas em diferentes situações e todas ao mesmo tempo. Porém, existem tarefas ou situações de uso linguístico que não exigem o acesso a todas elas de uma vez.

Um exemplo para o acima mencionado é a tarefa de nomeação de imagens (*Picture naming*). Nessa tarefa, o participante deve falar em voz alta o nome do item representado por uma imagem. Para isso, ele deve recuperar do seu léxico a representação semântica para aquela imagem, como também a informação fonológica para enviar ao sistema articulatório responsável pela produção oral. Dito de outra forma, ao visualizar a imagem de um avião, por exemplo, a informação semântica relacionada ao objeto em questão armazenada no léxico, como utilidade, formato, dentre outras, será acessada. Semelhantemente, a representação fonológica também será acessada, traduzindo-se na sua articulação fonética /e.vj'ẽw/ para que possa ser reproduzida pelo falante. Por outro lado, em uma tarefa de decisão lexical com base em input visual, ou seja, de palavras, não só informação fonológica e semântica são necessárias, mas também a representação ortográfica da palavra precisa ser recuperada.

O acesso a essas informações no léxico mental leva algumas centenas de milissegundos (DIJKSTRA, 2005). Além disso, há evidências de que a apresentação de uma palavra ative diversas outras possíveis candidatas com base na proximidade na representação ortográfica entre elas. Assim, a leitura da palavra “porta” pode levar a ativação de outras palavras como “corta” e “torta”. Essas palavras são chamadas de vizinhos ortográficos. Conforme o *N* de Coltheart (COLTHEART *et al.*, (1977), vizinhos ortográficos correspondem ao número de palavras que podem ser geradas ao se mudar uma letra em uma palavra do mesmo tamanho. Por exemplo, a palavra “casa” possui  $N = 52$ , o que significa que ela tem 52 vizinhos ortográficos na língua portuguesa, segundo o LexPorBR (ESTIVALET; MEUNIER, 2017). Em outras palavras, 52 palavras diferentes podem ser geradas ao se mudar uma das letras da palavra casa.

Em estudos com bilíngues, essas palavras normalmente são usadas como forma de controle. No estudo de Grainger e Dijkstra (1992), participantes bilíngues falantes de francês e

inglês realizaram uma tarefa de decisão lexical em que as palavras alvo estavam apenas na língua inglesa, as quais foram selecionadas considerando-se o número de vizinhos ortográficos que elas possuíam tanto no inglês quanto no francês. Os resultados mostraram que as palavras do inglês que possuíam muitos vizinhos ortográficos no francês, ou seja, vizinhos interlinguísticos, demandaram mais tempo dos participantes para a decisão lexical do que palavras que possuíam um número semelhante de vizinhos nas duas línguas. Já as palavras que possuíam um número semelhante em ambas as línguas, por sua vez, levaram mais tempo do que palavras que possuíam mais vizinhos ortográficos na própria língua inglesa. Com isso, os autores interpretaram que mesmo quando apenas palavras da língua inglesa estavam sendo requisitadas para a tarefa de decisão lexical, as palavras do francês estavam influenciando o tempo de reação (TR) dos participantes, sugerindo que as duas línguas estavam sendo processadas mesmo quando apenas uma era necessária.

Em um outro estudo realizado por Van Heuven, Dijkstra e Grainger (1998), o número de vizinhos ortográficos foi manipulado tanto na língua inglesa quanto na língua holandesa em uma série de experimentos envolvendo tarefas de desmascaramento progressivo com decisões lexicais. Na tarefa de desmascaramento progressivo, um par de estímulos compostos por uma máscara, normalmente um conjunto de marcadores (#), é apresentado ao participante várias vezes consecutivas, manipulando-se o tempo de exposição da máscara seguida por um alvo, de maneira que esse tempo diminui a cada repetição. Essas repetições acontecem até que o participante consiga identificar o alvo e dar uma resposta através de uma tecla determinada no teclado do computador.

Conforme os autores, ao aumentar o número de vizinhos ortográficos no holandês, o TR para palavras da língua inglesa aumentava. Ao se aumentar o número de vizinhos no inglês, que era a língua alvo, observou-se efeito inibitório sobre palavras do holandês e assim efeitos de facilitação sobre palavras da língua inglesa. Segundo Van Heuven, Dijkstra e Grainger (1998), esses efeitos inibitórios estão relacionados à hipótese de um léxico bilíngue integrado, e tratam-se de processos no léxico que restringem a ativação de itens de outra língua senão a do estímulo recebido. Interessantemente, monolíngues também apresentaram efeitos de facilitação por causa dos vizinhos ortográficos no inglês, mas nenhum efeito causado pelos vizinhos do holandês. Esses resultados colaboram com a ideia de que as palavras de ambas as línguas são ativadas durante o processamento lexical, corroborando com a hipótese de que o acesso lexical ocorre de forma não seletiva, conforme menciona Dijkstra (2005).

As palavras cognatas, por sua vez, são palavras de duas línguas que compartilham um alto nível de sobreposição ortográfica e/ou fonológica e um significado em comum

(SCHEPENS; DIJKSTRA; GROOTJEN, 2012a). Por exemplo, a palavra “*force*” do inglês e “força” do português compartilham pelo menos um significado em comum e possuem representações ortográficas semelhantes. Existem ainda pares de cognatas que possuem exatamente a mesma representação ortográfica em ambas as línguas, as quais são consideradas cognatas perfeitas, como é o caso da palavra “banana” para o inglês e o português. Segundo Schepens, Dijkstra e Grootjen (2012), para considerarmos uma palavra cognata, não é necessário que o par compartilhe todos os significados por eles representados em ambas as línguas a que pertencem. Neste sentido, as palavras “papel” e “*paper*” podem ser consideradas cognatas, uma vez que compartilham pelo menos um significado em comum, o de material utilizado para se escrever. O fato da palavra “*paper*” também carregar o significado de artigo científico, por exemplo, e a palavra papel não possuir tal significado não interfere a sua relação cognata.

Diversas variáveis influenciam diretamente a maneira como um bilíngue processa as suas duas línguas, as quais podem ser divididas em dois tipos, segundo Sánchez-Casas e García-Albea (2005): variáveis relacionadas ao usuário da língua, tais como nível de proficiência, experiência e contexto de aquisição da L2; e variáveis com relação ao tipo de palavra, como o “status cognato”, concretude e frequência da palavra. Neste capítulo, buscamos compreender no tópico 1.1 os aspectos principais do paradigma experimental de *priming* com decisão lexical. No tópico 1.2, discutiremos a relação das palavras cognatas com o léxico mental. Por fim, no tópico 1.3 faremos um levantamento de pesquisas utilizando o *priming* encoberto com palavras cognatas e não cognatas nas duas direções de língua.

## 1.2 EFEITOS DE PRIMING

Dentre as muitas formas experimentais utilizadas para estudar o léxico mental bilíngue, a tarefa de decisão lexical tem sido amplamente usada ao longo das últimas décadas. Na tarefa de decisão lexical, o participante deve indicar se um determinado conjunto de letras exibido corresponde a uma palavra ou não de uma certa língua. O paradigma experimental de *priming* com decisão lexical, por sua vez, consiste na apresentação de dois estímulos relacionados em um determinado espaço de tempo. O experimento de *priming* foi utilizado inicialmente por Meyer e Schvaneveldt (1971) com o objetivo de investigar como a frequência das palavras influencia o tempo de resposta na tarefa de decisão lexical.

Meyer e Schvaneveldt (1971) modificaram a tarefa de decisão lexical simples de maneira que dois estímulos relacionados fossem exibidos um após o outro para julgamento

pelos participantes. Sendo assim, o participante vê um primeiro estímulo (*prime*) e, posteriormente, podendo ser logo após ou não, outro estímulo relacionado (alvo) é exibido para que então a decisão lexical seja realizada. Existem versões em que outros itens são exibidos entre o *priming* e alvo, distanciando os dois, porém nos deteremos à anteriormente mencionada.

A ideia por trás desse paradigma é observar como o *prime* afeta o TR na decisão lexical do alvo. Nesse estudo, os estímulos utilizados por Meyer e Schvaneveldt (1971) poderiam ser tanto um par de palavras repetidas (bread-BREAD), um par de pseudopalavras repetidas (boctor-BOCTOR), ou uma palavra e uma pseudopalavra (bread-BOCTOR). Em seus resultados, apesar de encontrarem efeito de facilitação em todas as condições de palavras repetidas, sejam elas pseudopalavras ou palavras reais, eles observaram que os participantes levaram mais tempo quando a palavra alvo era uma pseudopalavra. Ou seja, palavras existentes na língua alvo dos *primes* apresentaram efeito de facilitação na decisão lexical maior do que pseudopalavras repetidas, fazendo com que os participantes levassem menos tempo para decidir se o item era ou não uma palavra. Isso acontece porque as pseudopalavras não possuem um significado e, conseqüentemente não fazem parte do léxico do participante, fazendo com que os *primes* de pseudopalavras não apresentassem efeitos significativos de facilitação. A partir de então, a técnica de *prime* consecutivo com tarefa de decisão lexical para investigar fenômenos como o processamento lexical em bilíngües passou a ser amplamente utilizada na psicolinguística experimental, passando a ser conhecida como o paradigma de *priming* (facilitação) com decisão lexical.

Posteriormente, ainda na década de 1980, reforçou-se o fato de que quando uma palavra era apresentada mais de uma vez ao mesmo participante, ocorria um aumento na velocidade e na precisão da escolha em tarefas de decisão lexical. Esse fenômeno ficou conhecido como efeito de repetição (FORSTER; DAVIS, 1984). De acordo com os autores, há basicamente duas abordagens distintas para o efeito de repetição. Na primeira, ele é visto como um produto de uma mudança temporária no processo de acesso lexical. Assim, uma palavra que foi anteriormente ativada permanece em um estado temporário de acesso facilitado, o que conseqüentemente permite que ela seja acessada mais rapidamente se for necessário em um curto período após sua primeira ativação. Na segunda, esse efeito pode ser entendido como um fenômeno que ocorre na memória. Com base nesta concepção, a primeira apresentação de uma palavra estabelece um traço de memória episódico que é contatado assim que a mesma palavra é apresentada novamente, facilitando assim sua retomada. De acordo com os autores,

uma vez concedida a possibilidade de envolvimento episódico, não podemos mais ter certeza de nenhuma das propriedades de um efeito de repetição puramente lexical,

porque não há como decidir quais propriedades são produzidas pelo componente episódico (FORSTER; DAVIS, 1984, p.682, tradução nossa).

O experimento de *priming* não consegue assegurar que os efeitos obtidos são decorrentes do processamento lexical. Segundo Sánchez-Casas e García-Albea (2005), isso é porque em experimentos de *priming* tradicionais, em que normalmente se usa tempos de exposição de *priming* maiores e que o tornam parcialmente visível para o participante, o mesmo pode ser gravado na memória episódica do sujeito, o que impossibilitaria a separação entre o processamento lexical e efeitos mais gerais da memória lexical.

Por este motivo, Foster e Davis (1984) realizaram uma série de experimentos em que as contribuições de fatores episódicos foram diminuídas ao máximo, com o objetivo de observar quais propriedades restariam. Para que isso fosse possível, era necessário que o estímulo do *priming* fosse encoberto (*masked*, em inglês) e o seu tempo de exibição fosse diminuído, assim os participantes seriam incapazes de identificá-lo, evitando o desenvolvimento de traços episódicos na memória, o que deu origem ao paradigma de *priming* encoberto com decisão lexical. Dessa forma, qualquer possível traço episódico estaria relativamente inacessível durante a tarefa de decisão, restando apenas o acesso lexical.

Por exemplo, no experimento 1, os autores comparam o tempo de reação entre palavras e pseudopalavras usando *primes* com duração de 60ms seguidos pela palavra repetida ou não. Como máscara, utilizaram uma palavra aleatória não relacionada com os itens experimentais. Além disso, Foster e Davis ainda compararam os efeitos para palavras com alta frequência e baixa frequência. Conforme reportado pelos autores, os *primings* de palavras com alta frequência causaram um efeito de 45ms no tempo de reação dos participantes. Semelhantemente, palavras com baixa frequência causaram um efeito de 38ms. Por outro lado, pseudopalavras causaram um efeito negativo no tempo de decisão lexical de -8ms. Os resultados obtidos pelos autores mostraram que é possível produzir efeitos de repetição mesmo com o *priming* encoberto para o participante, o que distancia a hipótese sobre a influência de traços episódicos para o acesso lexical.

Basicamente, os experimentos de *priming* encoberto consistem em o *prime* ser apresentado por um curto período de tempo (normalmente em uma janela de tempo entre 20-70ms), podendo ser mascarado tanto anterior quanto posteriormente à sua exibição, o que normalmente consiste em um conjunto de marcadores (#). Em seguida, a palavra alvo é exibida na tela para que o participante possa realizar a decisão lexical. Apesar da curta exibição do *priming*, comenta Forster (1998), a tarefa apresenta-se altamente eficiente, considerando-se até mais efetiva do que a técnica com *primes* visíveis (com tempos de exibição de 100ms ou mais).

O autor menciona ainda que é bastante improvável que um traço de memória episódica seja desenvolvido sem que o participante perceba o *prime* por conta da curta exibição, e que, ainda que isso fosse possível, seria pouco provável que tal traço tivesse alguma influência sobre a decisão lexical em tão pouco tempo.

Conforme mencionado anteriormente, estudos mostram que o *priming* de vizinhos ortográficos pode causar efeitos inibitórios quando na língua oposta à língua alvo e efeitos de facilitação quando relacionados a língua em questão da tarefa de decisão lexical. Além disso, ao longo dos anos, observou-se que palavras cognatas apresentam tempos de resposta menores que pseudopalavras em experimentos de *priming*, assim como palavras não cognatas apresentavam pouco ou nenhum efeito significativo de facilitação (KERKMAN, 1984 *apud* DE GROOT, 1993; CRISTOFFANINI *et al.*, 1986 *apud* DE GROOT, 1993; ALTARRIBA, 1992; DE GROOT; NAS, 1991). O mesmo se repete em experimentos de *priming* encoberto (DIMITROPOULOU; DUÑABEITIA; CARREIRAS, 2011; DUYCK; WARLOP, 2009; BASNIGHT-BROWN; ALTARRIBA, 2007; DE GROOT; NAS, 1991). Aparentemente, as palavras cognatas podem ser representadas de uma maneira diferente no léxico, como discutiremos no tópico seguinte.

### 1.3 ASSIMETRIA NO PRIMING ENCOBERTO E REPRESENTAÇÃO DE COGNATOS NO LÉXICO

Entender os fatores que determinam a organização lexical é um dos principais questionamentos levantados nos estudos sobre o léxico, sejam eles com bilíngues ou monolíngues. Por exemplo, modelos computacionais de processamento bilíngues como o BIA+ (DIJKSTRA; VAN HEUVEN, 2002) e Multilink (DIJKSTRA *et al.*, 2019) defendem que o acesso ocorre através de processos *bottom-up* e *top-down*, e que as representações no léxico vão desde traços visuais usados para a identificação de letras até nós linguísticos que evitam a confusão entre as línguas do falante na hora do uso.

Ao considerarmos que no léxico existem níveis ortográficos, fonológicos e semânticos, como seriam representadas as palavras cognatas? De acordo com Sánchez-Casas e García-Albea (2005), a importância de tal questão está no fato de que palavras morfologicamente relacionadas, semelhantemente às cognatas, compartilham uma raiz comum. Tais representações morfológicas tendem a ser reduzidas pelos modelos de processamento lexical, como os mencionados anteriormente, ao conjunto de características ortográficas, fonológicas e semânticas entre elas compartilhadas sem representações mais explícitas. Porém, de acordo

com os autores, há evidências de que as relações morfológicas são de alguma forma especiais e diferentes das relações de forma e sentido e que a sua representação no léxico pode ser atribuída a um nível representacional independente (GRAINGER; COLÉ; SEGUI, 1991; DREWS; ZWITSERLOOD, 1995; FROST; FORSTER; DEUTSCH, 1997).

Dessa forma, ao pensarmos no léxico bilíngue, a representação das palavras cognatas pode ter um status diferenciado também. Como mencionado anteriormente, elas possuem relação semântica, ortográfica e/ou fonológica, assim como as palavras morfológicamente relacionadas por uma raiz. Segundo Sánchez-Casas e García-Albea (2005), é possível que essa relação seja reduzida meramente à similaridade de forma e sentido entre elas, ou, por outro lado, pode acontecer que elas tenham algum tipo de representação morfológica especial e que estejam representadas juntas em uma única entrada no léxico.

Por exemplo, as palavras “*book*” e “*livro*”. Podemos dizer que estas duas palavras compartilham um significado em comum, porém não existe nenhuma semelhança ortográfica e/ou fonológica. Ou seja, elas compartilham apenas uma relação semântica. Já as palavras “*car*” e “*carro*” compartilham não só uma relação semântica, mas também ortográfica e fonológica. Neste sentido, conforme Sánchez-Casas e García-Albea (2005), é possível que estas duas palavras compartilhem uma única entrada no léxico, o que significa que ao se ativar uma, a outra também será ativada.

Davis *et al.* (2010) investigaram a facilitação ao acesso lexical através de experimentos de *priming* encoberto por cognatas com quatro grupos de bilíngues: i. bilíngues iniciantes falantes de espanhol como L1, ii. bilíngues balanceados falantes de espanhol e inglês, iii. bilíngues com espanhol como língua dominante e iv. bilíngues com inglês como língua dominante. Eles testaram através do *priming* encoberto com duração de 57ms e tarefa de decisão lexical em quatro condições, sendo elas apresentadas tanto na direção L1-L2 quanto na direção L2-L1, o que significa que participantes viram o *priming* na língua dominante com alvo na segunda língua e vice-versa. As condições experimentais foram: tipo de palavra (i.e., cognata ou não-cognata), uma condição de repetição (e.g., *clear/CLEAR*), uma condição de tradução cognata (e.g., *clear/CLARO*) e uma condição de controle ortográfico (e.g., *rict/RICO*). Eles observaram que as palavras cognatas causavam efeito de facilitação praticamente igual ao da condição de repetição em ambas as direções de *priming*. Por outro lado, as palavras não cognatas causaram pouco efeito de facilitação na direção L2-L1 e efeito negativo na direção L1-L2, assim como não houve efeito de facilitação ortográfica na condição controle. Conforme os autores, “o efeito de *priming* encoberto de cognatas ocorre porque traduções cognatas têm uma conexão especial com as suas traduções interlinguísticas equivalentes; uma conexão que

transcende qualquer especificidade linguística do léxico” (DAVIS *et al.*, 2010, p. 149, tradução nossa). Mesmo assim, não houve efeito principal entre o tipo de tradução (cognatos vs não-cognatos).

Preuss, Fontes e Finger (2015) realizaram um experimento envolvendo o paradigma de interferência desenho-palavra com bilíngues falantes de português e espanhol não balanceados, com proficiência intermediária-avançada na L2, com palavras cognatas, não cognatas e falso-cognatas. No paradigma em questão os participantes têm como tarefa nomear desenhos a eles apresentados em uma língua, enquanto ignoram palavras distratoras que podem ser apresentadas tanto visual, quanto sonoramente, seja na mesma língua de nomeação ou em outra língua.

As autoras puderam observar através de uma ANOVA efeitos significativos no que se refere ao *status* cognato das palavras ( $p < 0,05$ ), mais precisamente no que diz respeito ao efeito de facilitação fonológica/ortográfica entre palavras cognatas e falso-cognatas. Elas mencionam ainda que, mesmo não encontrando efeitos estatisticamente significativos na comparação entre as palavras cognatas e não cognatas, é possível observar que as primeiras apresentaram uma latência menor que estas últimas, mas que não foi possível observar isso em relação às falso-cognatas.

Ainda, o estudo desenvolvido por Preuss, Fontes e Finger (2015) mostra-se relevante para a compreensão do *status* cognato no que se refere às hipóteses envolvendo o acesso seletivo e não seletivo no léxico. Conforme reportado pelas autoras, o *status* cognato das palavras impacta funcionalmente o processo de produção de fala em bilíngues. Palavras que apresentam maior semelhança formal, como é o caso das cognatas e falso-cognatas, não necessariamente causam efeitos de interferência, como seria esperado baseando-se na hipótese do acesso seletivo. Pelo contrário, as palavras cognatas apresentaram efeito de facilitação, diminuindo o TR dos participantes, enquanto as palavras não cognatas e falso-cognatas apresentaram efeitos inibitórios/interferência. Com base nisso, é possível afirmar que as palavras cognatas, por compartilharem traços formais e semânticos, podem ter um acesso diferenciado de palavras que apresentam apenas traços formais semelhantes, como as falso-cognatas, e de palavras que apenas compartilham traços semânticos, como as não cognatas.

Toassi, Mota e Freitas (2020) investigaram o *status* cognato no multilinguismo a partir de uma tarefa de leitura de frases associada ao rastreamento ocular. Nesta tarefa, os participantes são convidados a fazer a leitura de um determinado número de frases enquanto um aparelho de rastreamento ocular (*eye-tracker*, em inglês) é usado para registrar os movimentos dos seus olhos. O uso dessa ferramenta oferece a oportunidade de observar onde

os olhos dos participantes passam mais tempo e quantas vezes eles retornam ao início de uma sentença/palavra, por exemplo, o que permite uma melhor compreensão do seu comportamento leitor. As autoras utilizaram esse método para verificar o efeito das cognatas no processamento lexical em trilingües de português brasileiro, alemão e inglês.

Três grupos participaram do estudo: um grupo de bilíngües falantes de português nativos e inglês como L2 (L2G), um grupo de trilingües falantes de português nativos, alemão como L2 e inglês como L3 (L3G) e um grupo controle de falantes nativos de inglês (L1G). As variáveis independentes controladas pelas autoras para a elaboração das sentenças foram: cognatas duplas inglês-português brasileiro, cognatas duplas inglês-alemão e cognatas triplas inglês-alemão-português brasileiro. As suas respectivas condições controles consistiram na utilização de uma palavra não cognata entre as línguas investigadas. Abaixo seguem exemplos apresentados pelas autoras para a condição de cognatos triplos (1) e seu respectivo controle (2), em que a palavra em negrito corresponde ao alvo da investigação:

(1) *Kate said that the **author** was inspired by the new book.*

(2) *Kate said that the **reader** was inspired by the new book.*

Os resultados apresentados por Toassi, Mota e Freitas (2020) indicaram um comportamento semelhante para os grupos L1G, L2G e L3G, em que o tempo da primeira passada (medida relacionada ao tempo que os participantes levam para ler a sentença completamente uma primeira vez) dos participantes em relação às palavras experimentais e controle nas condições envolvendo português brasileiro e inglês foi pequeno. Segundo as autoras, a média do tempo de fixação para a primeira passada para palavras cognatas 58ms menor que para as palavras controle para o grupo L2G, enquanto para o grupo L3G esse tempo foi de 31ms e 20ms para o grupo L1G. Por sua vez, nas condições envolvendo cognatas do inglês e alemão, o tempo de fixação para o grupo L2G foi 18ms maior para cognatas que para palavras controle e 12ms maior para o grupo L1G. Já para o grupo L3G, esse tempo foi 8ms maior para a condição envolvendo palavras cognatas.

Em contrapartida, na condição envolvendo palavras cognatas entre as três línguas, o grupo L3G mostrou um tempo de fixação maior para as palavras controle do que para as palavras cognatas (448ms e 407ms, respectivamente). Já para os demais grupos, as autoras indicaram que os tempos não mostraram nenhuma diferença significativa. Com relação à média do tempo de fixação, os resultados foram semelhantes, em que apenas o grupo L3G mostrou uma maior diferença na condição de palavras cognatas triplas, sendo 47ms menor que para

condição controle. Após análise estatística, as autoras confirmaram efeito significativo de facilitação para palavras cognatas apenas para o grupo L3G na condição de cognatas triplas ( $p < 0,05$ ) tanto para a média da primeira passada, quanto para a média do tempo de fixação. Com relação a análise entre os grupos, foi observado efeito significativo para o grupo L1G no que se refere às medidas da primeira passada, sendo ele o mais rápido em relação aos grupos não nativos. O mesmo ocorreu em relação às medidas da primeira fixação.

Resumidamente, as autoras apontam para um efeito de facilitação para a condição de cognatas triplas apenas no grupo L3G, como era esperado, indicando que tais palavras influenciam a compreensão de sentenças. Por outro lado, não foi possível observar efeito de facilitação para as condições que os cognatos pertenciam a apenas duas línguas, tanto para o grupo de bilíngues (português brasileiro e inglês) quanto para o grupo de trilingües (português brasileiro, alemão e inglês). Conforme Toassi, Mota e Freitas (2020), o fato de não encontrar facilitação nesta última condição não nega a hipótese do efeito de facilitação cognato, uma vez que o efeito contrário (interferência/inibição) não foi observado.

De acordo com Sánchez-Casas e García-Albea (2005), as palavras cognatas são representadas juntamente no léxico, como são consideradas as palavras morfologicamente relacionadas. Dessa forma, as informações de palavras como “família” e “*family*” estão localizadas em uma mesma entrada no léxico mental. Assim, a apresentação de uma das duas palavras pode fazer com que a entrada compartilhada por elas seja ativada e causar efeito de facilitação caso seja requisitado acesso a qualquer uma das duas em um curto espaço de tempo, como é comum nos experimentos de *priming*. Além disso, aponta Davis *et al.* (2010), isso explicaria por que palavras não cognatas não conseguem em grande parte um efeito de facilitação significativo. Uma vez que os pares de tradução não cognatas não são listados juntos, não seria possível a facilitação em tão pouco tempo considerando apenas a semelhança semântica.

Se assim fosse, indica Voga e Grainger (2007), poderíamos então descrever a existência de um mecanismo geral de aprendizagem, responsável por identificar correlações de forma e significado, explicando como palavras morfologicamente relacionadas são representadas em monolíngues e como palavras cognatas são representadas em bilíngues. Dessa forma, poderíamos dizer que uma forma de organização lexical baseada na morfologia das palavras além de existir, transcende através das línguas adquiridas por um bilíngue, não se limitando a relações de caráter intralingüístico apenas, como acontece com os monolíngues. Contudo, Voga e Grainger indicam ainda em seu estudo uma outra hipótese de que o efeito de facilitação

causado por palavras cognatas estaria relacionado a fonologia das palavras ao invés da morfologia.

Os autores realizaram 3 experimentos em que investigaram a facilitação de palavras cognatas por meio do *priming* encoberto em que os participantes falavam grego como L1 e o francês como L2. Uma vez que as duas línguas não compartilham um mesmo script, ou seja, não compartilharam uma mesma representação ortográfica, é pouco provável que haja efeitos puramente relacionados à forma ortográfica (GOLLAN; FORSTER; FROST, 1997). Voga e Grainger observaram que palavras morfologicamente relacionadas entre duas línguas (*κανονιά/kanonia/-canon*; tiro de canhão-canhão, em grego e francês, respectivamente), só geraram efeitos de facilitação em *primes* com duração de 66ms (34ms de efeito), mas nenhum efeito em *primes* com duração de 50ms (5ms, não significativo). Por outro lado, as palavras cognatas (*κανόνι/kanoni/-canon*: canhão em grego e francês, respectivamente) geraram efeitos significativos em ambas as condições de *priming* (36ms e 50ms, nas condições de *priming* de 50ms e 66ms, respectivamente). Já no experimento 2, eles viram que palavras cognatas com alto nível de sobreposição fonológica (*μετρό/metro/-métrô*: metrô em grego e francês) geravam efeitos de facilitação maiores que palavras cognatas com níveis de sobreposição fonológica menores (*χημεία/himia/-chimie*: química em grego e francês; 55ms vs 45ms, respectivamente). Esse resultado foi posteriormente reforçado no experimento 3.

Mais especificamente, no segundo experimento, foram utilizadas palavras não relacionadas como *prime* controle, mas que apresentavam um nível de sobreposição de forma, neste caso fonológica, semelhante ao observado pelas palavras cognatas. Dessa forma, a sobreposição de forma deixa de ser um fator que afeta o *priming* de traduções, visto que tanto as condições experimentais quanto controle compartilhavam um mesmo nível de sobreposição. Segundo Voga e Grainger (2007), os resultados observados sugerem que os efeitos observados são reflexo de dois mecanismos diferentes: um sensível a sobreposição semântica entre *primes* e alvos, e outro sensível a sobreposição de forma. Conforme os autores, “cognatos podem não ter nenhum status representacional especial, em relação às traduções não cognatas. A vantagem cognata pode ser devida simplesmente ao *priming* de forma adicional gerado por esses estímulos” (VOGA; GRAINGER, 2007, p. 944, tradução do autor).

Considerando-se que as palavras cognatas compartilhassem uma raiz comum entre as línguas de um bilíngue, deveria esperar-se que a condição de relação morfológica testada no experimento 1 de Voga e Grainger gerasse o mesmo efeito de facilitação que o de palavras cognatas, mas isso não aconteceu. Semelhantemente, era de se esperar que palavras cognatas tivessem vantagem sobre palavras não cognatas no que diz respeito a tradução de *primes*. De

fato, essa vantagem foi observada quando considerada a sobreposição de forma entre condições relacionadas e não relacionadas (26ms vs 48ms, respectivamente para palavras cognatas; 27ms vs 22ms, respectivamente para palavras não cognatas). Contudo, ao compararmos o efeito considerando apenas a sobreposição fonológica na condição de palavras relacionadas, a diferença entre palavras cognatas e não cognatas desaparece (26ms vs 27ms, respectivamente). Segundo os autores, a vantagem cognata no *priming* de tradução encoberto depende do tipo de linha de base usada para medir os efeitos de *priming* (VOGA; GRANGER, 2007, p. 946). Em outras palavras, esse efeito pode aparecer ou não dependendo de quais traços formais e/ou semânticos serão considerados durante a análise dos resultados, conforme observado acima.

No que diz respeito a tradução de palavras não cognatas, em seu estudo seminal De Groot e Nas (1991) realizaram dois experimentos de *priming* com decisão lexical com bilíngues falantes de holandês e inglês usando pares de tradução cognatos e não-cognatos (experimentos 3 e 4). Seu material apresentava quatro listas divididas em EE (inglês-inglês), DD (holandês-holandês), DE (holandês-inglês) e ED (inglês-holandês). Dentro dessas listas, as variáveis avaliadas eram identidade (GROUND-ground) associação semântica (CALF-cow) e palavra sem relação (BRIDE-task), bem como uma condição controle de palavra com pseudopalavra (SHELL-tosk). Eles observaram que tanto palavras cognatas como não cognatas geraram efeitos de facilitação, mas apenas palavras cognatas causavam facilitação significativo quando a palavra alvo estava na condição de associação semântica.

Em sua investigação, Post da Oliveira (2020) verificou a compreensão de como os aprendizes de uma segunda língua (L2) processam e reconhecem o acento tônico das palavras por meio da técnica de *priming*. O estudo mostra evidências de que os aprendizes de L2 conseguem recuperar informações sobre o acento tônico de palavras a partir da ortografia, mesmo quando não têm conhecimento prévio sobre a pronúncia correta. Isso pode indicar que a percepção do acento tônico pode ser influenciada por fatores contextuais, como o uso de palavras relacionadas ou a presença de dicas visuais, como é muito comum em palavras cognatas, mais uma vez destacando a importância da ortografia na facilitação do processamento e reconhecimento de palavras.

O *priming* encoberto com decisão lexical de palavras não cognatas com bilíngues avançados já foi indicado em estudos anteriores como um efeito assimétrico que depende da direção do estímulo para sua aparição. Essa assimetria do *priming* encoberto (DIMITROPOULOU; DUNABEITIA; CARREIRAS, 2011) é evidenciada por um número consistente de estudos que obtiveram efeitos significativos de *priming* encoberto de traduções na direção L1-L2, mas encontraram pouco ou nenhum efeito na direção L2-L1 (DE GROOT;

NAS, 1991; WILLIAMS, 1994; JIANG; FORSTER, 2001; KIM; DAVIS, 2003; VOGA; GRAINGER, 2007). Além disso, Dimitropoulou, Duñabeitia e Carreiras (2011) destacam que o nível de proficiência parece ser um fator importante para a aparição desse efeito. Isso significa que tal efeito varia bastante conforme o nível de proficiência dos participantes.

Como já mencionado anteriormente, são muitos os fatores que podem influenciar os efeitos de *priming*. Wen e Van Heuven (2017) realizaram uma meta-análise envolvendo 64 experimentos de *priming* encoberto de palavras não cognatas, retirados de 24 estudos, sendo 31 deles na direção L1-L2 e os outros 34 na direção L2-L1. Eles consideraram um total de sete moderadores que podem influenciar o efeito de *priming* encoberto de traduções: a) o número de participantes, b) a duração do *priming*, c) a assincronia de início do estímulo (SOA, *Stimuli Onset Asynchrony*, ou seja, o intervalo entre o início do *prime* e o início do alvo), d) o intervalo interestímulo (ISI, Inter-Stimulus Interval, ou seja, o tempo entre a saída do *prime* e o início do alvo), e) o tipo de script linguístico, ou seja, o tipo de representação ortográfica adotado pelas línguas, e.g. o alfabeto romano, f) o número de itens por célula, que se refere à quantidade de estímulos que compõem uma lista experimental e g) tempo de resposta.

Os autores calcularam o tamanho do efeito de *priming* (d) a partir dos valores de *t* ou *F* disponibilizados pelos estudos analisados, mais o número de participantes (*n*) a partir da fórmula proposta por Rosenthal (1991),  $d = \frac{t}{\sqrt{n}}$  ou  $d = \sqrt{\frac{F}{n}}$ . Foram observados efeitos significativos no tamanho do *priming* (d) de traduções nas duas direções: 0.86,  $z = 12.869$ ,  $p < 0.0001$  para a direção L1-L2 e 0.31,  $z = 6.3481$ ,  $p < 0.0001$ , para a direção L2-L1. De acordo com eles, a assimetria do *priming* de tradução deveria ser vista de forma quantitativa, ao invés de qualitativa. Assim sendo, essa diferença observada entre os efeitos de *priming* de palavras não cognatas não seria reflexo de uma diferença qualitativa na aquisição da L1 e da L2, mas que outros fatores de caráter quantitativo, como exposição ao idioma, frequência de uso, estariam envolvidos na diferença no processamento delas.

Além disso, após a análise dos sete moderadores anteriormente mencionados, Wen e Van Heuven (2017) observaram que o *priming* na direção L1-L2 parece ser moderado pelo ISI, enquanto o *priming* na direção L2-L1 é moderado pelo número de itens por célula. Isso significa que ao se aumentar o ISI, provavelmente haverá efeitos mais fortes na direção L1-L2 de *priming*. Já com relação a direção L2-L1, aumentar o número de itens que compõem a lista experimental pode aumentar a força do efeito de *priming* também. Segundo os autores, a escolha do número de itens por célula nos estudos analisados, com estudos em que apenas 12 estímulos por lista eram apresentados, enquanto outros apresentavam 80, é pouco justificada.

Eles sugerem que a busca pelo estabelecimento de um número de itens por lista experimental pode ajudar a compreender esses dados de forma mais precisa. Em outras palavras, a grande variação no número de itens experimentais usados nas listas dos estudos analisados dificultou a análise deste moderador. Talvez se diminuíssemos essa grande diferença entre a quantidade de itens que compõem as listas dos experimentos com *priming* encoberto de palavras não cognatas, obteríamos dados bem mais claros e teríamos maior propriedade para avaliar seus efeitos.

Em nosso estudo, utilizaremos o mesmo tempo de exibição de *priming* e alvo para as duas direções, assim como usaremos o mesmo número de traduções tanto para cognatas quanto para não cognatas. A ausência de efeito na direção L2-L1 pode significar que o número de itens utilizados foi insuficiente para gerar um efeito significativo, corroborando com a ideia de que quanto mais itens por célula, mais forte fica o efeito na direção L2-L1. Como não teremos diferenças no ISI, não poderemos avaliar o efeito desse moderador.

Para Kroll e Tokowicz (2001), a assimetria na direção do *priming* está relacionada a necessidade de mediação do acesso aos significados para a L2. Nesse sentido, bilíngues dependem da mediação realizada pela língua materna para garantir acesso aos significados para palavras da sua L2. Já Finkbeiner *et al.* (2004) sugere que há uma assimetria no nível representacional entre as representações semânticas da L1 e da L2. Segundo ele, a língua materna possui um nível bem maior e mais rico de representações semânticas em relação a L2. Por essa razão, o *priming* em L1 consegue causar efeitos de facilitação, uma vez que esta é capaz de ativar muito mais “sentidos” que a L2, deixando-os disponíveis para acesso posterior temporariamente. Já a L2, por não ter um número abrangente de sentidos, esta não consegue ativar representações léxico-semânticas suficientes para causar facilitação para a L1. Ambas as hipóteses concordam que ao passo que o indivíduo aumenta seu nível de proficiência, essa assimetria entre as línguas diminui. Tais questões serão retomadas no capítulo 2.

Em nosso estudo, utilizaremos bilíngues tardios nos níveis intermediário e avançado de proficiência. Assim, esperamos que os bilíngues com níveis mais altos de proficiência apresentem efeitos de facilitação simétricos para ambas as direcionalidades da língua e independente do tipo de tradução usado, corroborando com a hipótese de que quanto maior os níveis de proficiência, menor a diferença no efeito de *priming* com relação a direção e tipo de tradução.

#### 1.4 NÍVEL DE PROFICIÊNCIA E EXPERIÊNCIA LINGUÍSTICA

Como mencionado no início deste capítulo, o termo bilíngue conforme utilizado aqui, considera diversos fatores como idade de aquisição, nível de proficiência em cada uma das línguas do bilíngue, dentre outros. O nível de proficiência também parece ter um papel crucial no surgimento do efeito de *priming* com não-cognatos. Basnight-Brown e Altarriba (2007) obtiveram efeitos significativos de *priming* para palavras não-cognatas com bilíngues com alto nível de proficiência. Semelhantemente, Duñabeitia, Perea e Carreiras, (2010a) obtiveram com bilíngues simultâneos e balanceados de basco e espanhol efeitos de *priming* tanto para palavras cognatas quanto para palavras não cognatas, em ambas as direções de *priming*, reforçado mais tarde no estudo feito por Duñabeitia *et al.*, (2010b) utilizando bilíngues simultâneos e balanceados com *priming* e decisão lexical associados a ERPs para as mesmas línguas.

De acordo com a Weaker Links hypothesis de Gollan et al. (2008), as conexões entre as informações semânticas, fonológicas e ortográficas são mais fracas em bilíngues não muito proficientes. Segundo Litkofsky et al. (2016), a idade de aquisição (IdA) também pode ser considerada um fator crucial no desenvolvimento do desempenho linguístico em L2: quanto mais alta a IdA, menor o desempenho linguístico na L2. Eles mencionam ainda que o tempo de imersão à L2 para bilíngues está diretamente associado a diminuição da exposição à L1. Ou seja, quanto mais tempo o sujeito é exposto à L2, menos exposição a sua L1 acontecerá, o que indicaria um processo inverso entre as duas línguas. Em seu estudo, Litkofsky et al. (2016) associaram melhor produção em L1 ao menor tempo de vivência em espaços em que a L2 é utilizada e a melhor produção na L2 à menor idade de aquisição.

Van Hell, Adams e Abdollahi (2019) Fizeram uma análise de diversos estudos envolvendo o processamento bilíngue e que consideraram a proficiência enquanto fator importante nas suas análises. Dentre eles, o estudo de Van Hell e Dijkstra (2011) em particular mostra que o nível de proficiência não só afeta a magnitude do efeito cognato no processamento lexical, mas também que bilíngues menos proficientes na L2 tendem a ser mais vulneráveis a fatores contextuais, como a composição da lista experimental. Em outras palavras, a seleção dos estímulos experimentais e das palavras controles (homógrafas, falso-cognatos, não cognatos, pseudopalavras) afeta a forma como os participantes irão se comportar na tarefa experimental em questão.

Segundo Van Hell, Adams e Abdollahu, “a coativação de línguas alvo e não-alvo é modulada por variações na proficiência linguística dos bilíngues, de modo que o processamento lexical na língua alvo só é afetado por informações da língua não-alvo se a proficiência no

idioma não-alvo for suficientemente alta (tradução do autor, 2019)”. Dito de outra forma, palavras da L2 só afetam o processamento da L1 se o bilíngue apresentar um nível de proficiência suficientemente alto na L2, o que se relaciona diretamente nos efeitos e direcionalidade de *priming* investigados nesta pesquisa.

Buscamos aqui analisar qualitativa e quantitativamente de que forma o nível de proficiência e a IdA da L2 influencia o tempo de resposta dos participantes na tarefa de decisão lexical com *priming* encoberto. Aplicamos um questionário sociolinguístico, além de um teste de proficiência com o objetivo de verificar as particularidades na aquisição e no uso da língua inglesa pelos participantes.

Como mencionado anteriormente, as palavras cognatas parecem ter algum tipo de vantagem sobre as não cognatas no que diz respeito à facilitação por *priming* encoberto. A não existência de efeitos de facilitação com palavras não cognatas mencionada por Sánchez-Casas e García-Albea (2005) para propor um status diferenciado para as palavras cognatas pode ser um pouco equivocada, uma vez que estudos mais recentes mostram que é possível sim obter tais efeitos. Ainda, conforme destacado por Schoonbaert *et al.*, (2009) e posteriormente reforçado por Wen e Van Heuven (2017) em seu estudo, a assimetria no *priming* de palavras não cognatas parece ser um efeito quantitativo, ao invés de qualitativo, e diferentes fatores podem influenciar a aparição desse efeito. Como veremos a seguir, diversos modelos de representação e processamento lexical em bilíngues propõem hipóteses para responder tais questões.

## 2 MODELOS DE REPRESENTAÇÃO LEXICAL BILÍNGUES

A busca por compreender-se o processamento lexical levou ao desenvolvimento de diversos modelos que tentam explicar como as palavras são acessadas e organizadas na memória de indivíduos/falantes/pessoas bilíngues. Weinreich (1968) *apud* Dijkstra (2005) propôs que os bilíngues podem ser classificados de três formas: coordenativos, que armazenariam suas duas línguas em lugares distintos em todos os níveis representacionais: ortográfico, fonológico e semântico/conceitual; compostos, que, ao contrário dos coordenativos, seriam capazes de diferenciar as palavras das duas línguas, porém compartilhariam um único significado; e por fim os bilíngues subordinativos, que apresentariam diferenças na força de cada língua, sendo a L2 mais fraca e subordinada a mediação de significados realizada pela língua mais forte, que é a L1.

Conforme Dijkstra (2005), essas definições foram posteriormente utilizadas para, ao invés de caracterizar os tipos de bilíngues, sugerir como o léxico mental bilíngue poderia ser organizado. Com isso, os modelos poderiam ser basicamente elaborados com os pressupostos de que o acesso lexical acontece através da mediação conceitual ou por associação de palavra. Desta forma, ao visualizar as letras do alfabeto correspondentes à palavra em português “casa”, essa representação ortográfica poderia ser ativada no léxico, por sua vez, ativaria a representação semântica a ela correspondente. Em seguida, a ativação desse significado seria capaz de ativar a palavra em inglês “house”. Esse procedimento de tradução então pode ser chamado de mediação conceitual. Por outro lado, outra possibilidade seria a que a própria palavra no português “casa” ativasse a palavra em inglês “house” diretamente pela ativação semântica. Essa abordagem se refere à associação de palavra.

Atualmente, os principais modelos de acesso lexical bilíngue são distinguidos pelo tipo de conexão desenvolvida pelos componentes da memória lexical. De acordo com Soto (2010), esses modelos são: conexionistas distribuídos, os quais consideram o caráter dinâmico de redes neurais interconectadas diferenciadas pelo peso de ativação a elas atribuídas por meio da experiência, sem uma separação clara entre cognições diferentes; conexionistas locais, que postulam diferentes camadas de representações com identidades discretas e interligadas por um tipo de corrente de ativação; e os modelos modulares, os quais isolam as operações linguísticas em diferentes módulos encapsulados que trabalham de forma serial, de maneira que o produto de saída de um módulo trata-se do produto de entrada para o módulo seguinte.

A seguir, discutiremos sobre alguns modelos relevantes para os estudos sobre o processamento lexical bilíngue. Iniciamos com a hipótese do modo de língua, um modelo mais

geral que considera as diferentes possibilidades de ativação para o léxico bilíngue e que está em consonância com os demais modelos. Em seguida, apresentamos o modelo Hierárquico Revisado, o qual foi de grande importância principalmente por considerar a maneira como a proficiência influencia o acesso lexical por bilíngues. Por fim, analisamos dois modelos conexionistas locais, o modelo de Ativação Interativa Bilíngue Mais e o modelo Multilink, que atualmente mostram-se como referências no que se refere à representação lexical bilíngue.

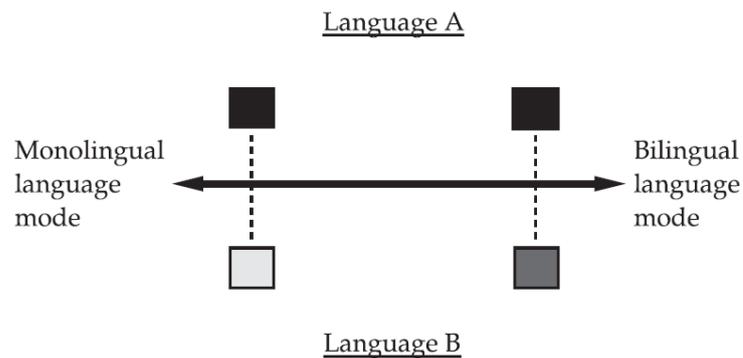
## 2.1 A HIPÓTESE DO MODO DE LÍNGUA

O estado de ativação de uma das línguas do bilíngue pode variar de acordo com o contexto linguístico que ele está inserido em um determinado momento. Essa hipótese foi desenvolvida por Grosjean (1985) e ficou conhecida como a Hipótese do Modo de Língua (*The Language Mode Hypothesis*, em inglês). De acordo com ela, um bilíngue pode entrar em um “modo de língua” com base na influência contextual externa, como por exemplo, a visualização de uma frase em um livro ou em uma página na internet em uma das línguas, ou até um encontro com um amigo na rua.

Para isso, Grosjean (1985) estabelece um continuum, que vai desde o “modo monolíngue”, em que a língua materna permanece com altos níveis de ativação enquanto a L2 é suprimida pelo sistema, até o “modo bilíngue”, em que ambas as línguas se encontram em altos níveis de ativação. Ainda considerando o exemplo mencionado anteriormente, se tal amigo encontrado na rua só fala a sua língua materna, por exemplo, você automaticamente estará no primeiro modo mencionado. Em contrapartida, se você encontra um amigo com quem você costuma utilizar a sua L2, você provavelmente passará para o modo bilíngue.

Observe na Figura 1 que em ambas as posições, a língua A, que se refere à língua materna, está altamente ativada, representada pelos quadrados escuros. Em contrapartida, no lado esquerdo inferior do continuum, a língua B, que se refere à L2, está com níveis de ativação consideravelmente baixos, representados pelo quadrado mais claro. Isso significa que o bilíngue está no modo monolíngue. Por outro lado, a língua B está altamente ativada no outro extremo do continuum, ainda que não tanto quanto a língua A, sinalizado pelo quadrado mais escuro, indicando que o modo bilíngue do sujeito está ativado.

FIGURA 1 - Representação do continuum do Modo de Língua



Fonte: Grosjean (1985)

Conforme Grosjean (2013), a forma como essas línguas são ativadas é bastante dinâmica. Normalmente, bilíngues permanecem no modo monolíngue até que o uso da sua L2 se faça necessário. Em situações em que a L2 é o principal instrumento de comunicação, o que o autor chama de “língua base”, pode ser que a L2 tenha níveis mais altos de ativação que a língua materna do indivíduo. Ainda, há a possibilidade que uma das línguas do indivíduo, podendo ser tanto a L1 quanto a L2, permaneça parcialmente ativada em um nível intermediário do continuum, por exemplo, quando o interlocutor é um bilíngue, mas não se sente confortável em fazer alternâncias entre as línguas durante a interação. Na figura 1, essa possibilidade é representada pela L2, ou seja, a língua B.

Diversos fatores influenciam a intensidade que essas línguas são ativadas no continuum. Segundo Tokowics (2014), a hipótese do modo de língua leva em consideração a importância do conhecimento *top-down*, ou seja, conhecimento de níveis mais elevados no controle da ativação do léxico bilíngue. Esses conhecimentos podem ser o contexto em que uma conversação é iniciada, o conhecimento sobre o background linguístico de um interlocutor, o tópico da conversa, dentre outros. Essas informações, assim, podem ser usadas para modular o nível de ativação das línguas do bilíngue, conforme mencionado anteriormente.

## 2.2 MODELO HIERÁRQUICO REVISADO

Proposto por Kroll e Stewart (1994), o Modelo Hierárquico Revisado (MHR) foi o modelo que melhor respondeu às questões que dizem respeito a representação da organização do léxico bilíngue durante a década de 1990. Em meados da década de 1980, Potter *et al.* (1984) desenvolveram dois modelos envolvendo uma hierarquia na organização lexical em bilíngues.

O primeiro modelo foi chamado de modelo de Associação da Palavra<sup>1</sup>, que sugere que as palavras, ou seja, os itens lexicais da L2 são associados aos seus vocábulos correspondentes, ou seja, suas traduções, na L1. Uma alternativa seria a possibilidade dos itens lexicais da L2 serem associados aos conceitos não-linguísticos comuns a ambas as línguas, ao invés de serem associadas às palavras da L1. Esses conceitos não-linguísticos se referem ao conhecimento de mundo dos falantes, a informação obtida através do contato do indivíduo com o mundo. Esse modelo foi chamado de Modelo de Mediação Conceitual<sup>2</sup>. Tanto o primeiro como o segundo estão de acordo com a ideia de que há uma separação entre itens lexicais e os conceitos na memória, ou seja, existiria uma hierarquia entre esses componentes.

Com base nesses pressupostos, Kroll e Stewart (1994) montaram três experimentos com o objetivo de verificar as hipóteses propostas por Potter *et al.* (1984). Os dois primeiros foram tarefas de nomeação de figuras e de palavras; e o terceiro foi uma tarefa de tradução. Nas duas primeiras tarefas, o participante deve nomear em alta voz uma imagem ou uma palavra a ele apresentada. Já na segunda tarefa, o participante deve falar a tradução de uma determinada palavra a ele apresentada. Os autores verificaram que os participantes conseguiam nomear palavras com mais facilidade do que figuras, bem como as pessoas tinham mais dificuldade em nomear figuras agrupadas semanticamente (*hat-belt-suit*) do que de forma aleatória (*lion-ambulance-lemon*). Segundo Kroll e Stewart, esse tempo adicional para nomear as imagens pode estar relacionado à necessidade das imagens de acesso ao nível conceitual, enquanto os itens lexicais não o precisariam. Além disso, os autores observaram através da tarefa de tradução que os participantes demoravam mais para traduzir na direção L1-L2 (119ms a mais) do que na direção L2-L1. Com base nesses dados, eles propuseram uma revisão para o modelo hierárquico de Potter *et al.* (1984).

Conforme proposto, as palavras armazenadas na memória de um bilíngue são organizadas em dois léxicos separados, ou seja, as palavras da L2 são organizadas em um módulo diferente das palavras da L1, ao passo que os conceitos são armazenados em um sistema de memória abstrato, mais profundo, separado e compartilhado por ambas as línguas. Para Brysbaert e Duyck (2010), a característica mais importante dessa revisão está na interação entre esses diversos componentes. A saber, a proposta de Kroll e Stewart (1994) combina o modelo de Associação de Palavras e o modelo de Mediação Conceitual propostos por Potter *et al.*

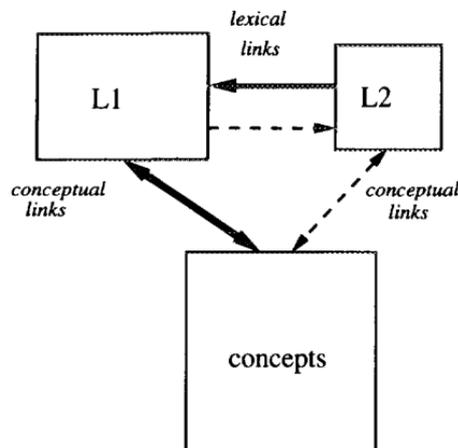
---

<sup>1</sup> *Word Association Hypothesis*, Potter *et al.* (1984).

<sup>2</sup> *Concept Mediation Hypothesis*, idem (1984).

(1984). Destaca-se que a diferença se dá na existência de ligações mais fortes ou mais fracas entre esses módulos, conforme observamos na Figura 2.

FIGURA 2 - Modelo Hierárquico Revisado



Fonte: Kroll e Stewart (1994)

Conforme a Figura 2 acima do modelo proposto pelos autores, o módulo lexical da L1 seria maior do que o módulo lexical da segunda língua pelo fato de que, mesmo em indivíduos com alto nível de proficiência, mais palavras são conhecidas na L1 do que na L2. Observa-se ainda que a ligação entre a L2 e a L1 é mais forte do que o inverso. Isso se dá devido ao fato que normalmente a L2 é adquirida fazendo-se relação com as palavras equivalentes na L1. Em contrapartida, a ligação da L1 dos bilíngues com o sistema conceitual seria mais forte do que a da L2, uma vez que a aquisição da L1 se dá por meio da associação das palavras aos conceitos apreendidos do mundo. Desta forma, os autores definem o que ficou conhecido como o Modelo Hierárquico Revisado (MHR, KROLL; STEWART, 1994).

Brysbaert e Duyck (2010) destacam as principais contribuições do MHR para os estudos sobre organização lexical e o processamento bilíngue. Além da separação entre as representações lexicais e conceituais, ressalta-se a hipótese do acesso seletivo aos componentes lexicais distintos.

Em vistas a coexistência de módulos lexicais distintos para cada língua adquirida pelo bilíngue, tornou-se possível a hipótese de que o acesso a esses sistemas se dá de forma seletiva, uma vez que eles funcionam de forma independente. De acordo com essa perspectiva, o bilíngue pode ativar ou inibir uma de suas línguas, a depender do contexto (VAN HEUVEN; DIJKSTRA; GRAINGER, 1998). Dessa forma, uma pessoa que está processando palavras em sua L1 consegue desativar o léxico da L2, bem como alguém com alto nível de proficiência

processando palavras na sua L2 consegue inibir o léxico da sua L2 (BRYLSBAERT; DUYCK, 2010).

Além disso, outra contribuição de grande importância para o bilinguismo e que ainda hoje é amplamente discutida é a assimetria entre o processamento da L1 e da L2. Como pode ser observado na **Erro! Fonte de referência não encontrada.** e fora mencionado anteriormente, esse modelo assume que as conexões da L1 de um bilíngue com o módulo conceitual são muito mais fortes do que as da L2. Da mesma forma, a ligação entre o módulo lexical da L2 com o sistema da L1 é muito mais forte do que o contrário. Sendo assim, afirmam Kroll e Stewart (1994), ainda que haja uma relação de bidirecionalidade entre as línguas, a assimetria entre essas conexões implicaria em diferenças no nível em que a tradução nas duas direções é afetada por fatores de nível conceitual.

Em dois dos seus experimentos, nos quais os participantes deveriam nomear imagens da mesma categoria ou de categorias diferentes, Kroll e Stewart (1994) evidenciaram a existência de um efeito de interferência de categoria. Foi observado que palavras organizadas em grupos da mesma categoria levavam mais tempo para serem nomeadas do que de categorias diferentes. A partir disso, um terceiro experimento foi realizado com o objetivo de identificar se o mesmo ocorria para ambas as línguas dos participantes. Os autores identificaram que esse efeito de interferência só acontecia com tradução de palavras na direção L1-L2 e não na inversa. Com isso, os resultados apontaram para a hipótese de que, “para bilíngues relativamente fluentes, mas não-balanceados, existe uma assimetria entre as duas direções da tradução que reflete a dependência diferencial da ativação lexical e conceitual durante o processo de tradução” (KROLL e STEWART, 1994, p.168, tradução do autor). Em outras palavras, como a ligação entre o léxico da L2 e o sistema conceitual são mais fracas do que na L1, a tradução na direção L2-L1 depende menos dos conceitos e mais dos itens correspondentes na L1, com a qual tinha uma conexão mais forte. Por outro lado, a tradução na direção L1-L2 depende mais dos conceitos do que dos itens da L2, dado que a sua ligação com o sistema conceitual era mais forte.

Estudos utilizando o paradigma do *priming* interlinguístico como observado no Capítulo 1, estão em consonância com a hipótese da assimetria entre L1 e L2 propostas pelos autores em seu modelo.

Por fim, outro destaque do MHR feito por Brysbaert e Duyck (2010) refere-se aos aspectos de desenvolvimento do bilinguismo e que é consoante com diversos outros modelos, como o BIA/BIA+ (VAN HEUVEN; DIJKSTRA; GRAINGER, 1998; DIJKSTRA; VAN HEUVEN, 2002) e o Sense Model (FINKBEINER *et al.*, 2004). A proposta é de que o léxico

bilíngue muda com base no nível de proficiência. Assim, uma pessoa adquirindo uma L2 só desenvolveria conexões com as suas respectivas traduções na sua L1. Ao passo que a proficiência aumenta, essa pessoa começa a criar ligações com o sistema conceitual, de maneira que indivíduos com alto nível de proficiência teriam conexões entre a L2 e os conceitos quase tão fortes quanto da L1.

A contribuição do MHR para o bilinguismo foi indubitavelmente valiosa, especialmente no que se refere à produção e tradução bilíngue. Contudo, diversas críticas ao mesmo surgiram desde a sua apresentação (BRYSSBAERT; DUYCK, 2010), o que levou, posteriormente, ao desenvolvimento de novas hipóteses e modelos compatíveis com as descobertas mais recentes.

### 2.3 O MODELO DE ATIVAÇÃO INTERATIVA BILÍNGUE MAIS

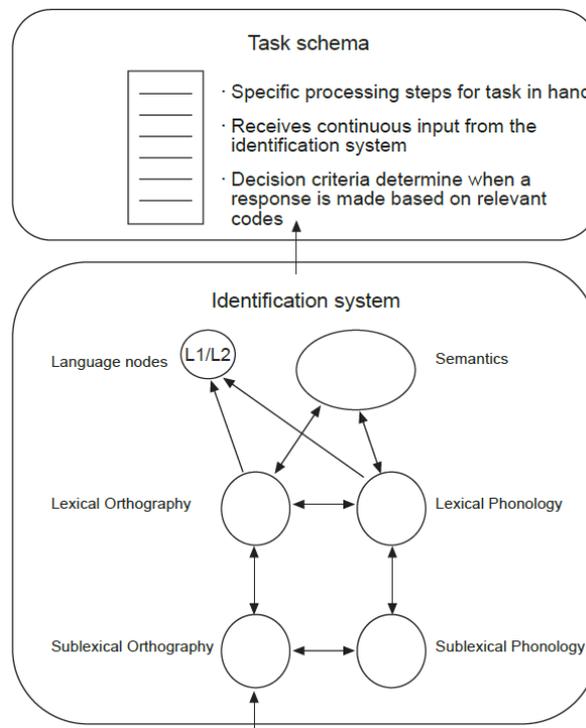
De acordo com Lam e Dijkstra (2010), o número de variáveis que podem influenciar no processamento lexical são um problema para quem busca investigar o fenômeno. Um modelo computacional capaz de simular a velocidade que uma palavra é reconhecida, bem como quais variáveis podem interagir de maneira quantitativa e qualitativa nesse tempo de seleção com base nos pressupostos básicos dos processos de reconhecimento de palavras, pode contribuir significativamente para o avanço nas investigações sobre o tema.

O Modelo de Ativação Interativa Bilíngue Mais (BIA+, Bilingual Interactive Activation Model em inglês), desenvolvido por Dijkstra e Van Heuven (2002), propõe-se a explicar e simular, através de modelagens computacionais, como os padrões assimétricos observados no processamento bilíngue ocorrem. Diferente do MHR, o BIA+ incorpora os léxicos das línguas do bilíngue em um só componente, assim como defende o seu acesso não seletivo, ou seja, sem separação entre as línguas durante o processamento.

Como o próprio nome indica, o BIA+ é uma versão ampliada de um modelo existente, o BIA (DIJKSTRA; VAN HEUVEN, 1998; VAN HEUVEN; DIJKSTRA; GRAINGER, 1998) o qual, por sua vez, baseia-se no modelo monolíngue de Ativação Interativa de McClelland e Rumelhart (1981). A diferença entre eles está na simplificação do modelo, assim como na adição de componentes anteriormente não contemplados pelo BIA. A saber, o modelo se preocupava tão somente com reconhecimento de representações ortográficas, deixando de lado fatores não-linguísticos externos como contexto, tipo de tarefas, dentre outros, assim como aspectos fonológicos e semânticos. Em razão das diversas limitações apresentadas pelo modelo, especialmente relacionadas à descrição e funcionalidades dos nós linguísticos, que seriam responsáveis pela identificação da língua a que um determinado estímulo correspondia, a falta

de representações morfológicas e semânticas, dentre outras, se viu a necessidade de uma modificação no mesmo. O BIA+ surge então com o objetivo de preencher tais lacunas. Ainda assim, afirmam Dijkstra e Van Heuven (2002), muitos desses problemas foram parcialmente resolvidos, enquanto outros puderam apenas ser analisados verbalmente.

FIGURA 3 - Modelo de Ativação Interativa Bilíngue Mais (BIA+)



Fonte: Dijkstra e Van Heuven (2002)

Conforme pode ser observado na Figura 3, o BIA+ é composto por um sistema de reconhecimento que se inicia em um nível sublexical, que por sua vez se divide na identificação de letras e fonemas. Esses são seguidos pelo nó lexical que também apresenta a mesma separação. Por fim, existe o nó semântico, que é responsável por armazenar os significados para ambas as línguas e o nó referente ao parentesco linguístico (*linguistic membership representations*) que é responsável por identificar a que língua o estímulo recebido corresponde. Como representado pelas setas, a interação entre esses componentes acontece de forma mútua com ativação tanto *bottom-up* quanto ativação *top-down*, com exceção do nó linguístico.

Uma adição importante para o BIA+ em relação à versão anterior é a distinção entre o sistema de identificação de palavra e o sistema de tarefa/decisão. Fortemente influenciados pelo trabalho de Green (1998) e Dijkstra e Van Heuven (2002) complementam ao BIA+ o modelo de Controle Inibitório que estabelece diferenças entre o tipo de ativação que irá ocorrer a depender do tipo de tarefa a ser realizada. Essa mudança no modelo faz com que os nós

linguísticos fiquem responsáveis apenas pelas representações de parentesco linguístico dentro do sistema de identificação de palavra, ou seja, eles conseguem identificar a que língua as palavras pertencem, mas não funcionam mais como filtros, os quais, por sua vez, tornam-se um sistema distinto denominado de sistema de tarefa/decisão. De acordo com Lam e Dijkstra (2010), esse sistema é responsável pelas diversas tarefas cognitivas que precisam ser tomadas, assim como decisões que precisam ser feitas para que o leitor consiga chegar ao seu objetivo final. Agora, os efeitos relacionados a contextos não-linguísticos e disparidades nos resultados entre experimentos está associado a maneira como os parâmetros das tarefas de tomada de decisão são elaborados e organizados, refletindo em diferenças nas respostas obtidas por meio do sistema de tarefa/decisão.

Além disso, o modelo BIA+ ainda faz a incorporação de dois importantes nós não presentes no BIA: os nós fonológico e semântico. A partir disso, o modelo compreende que o reconhecimento de palavras em bilíngues não é afetado apenas pelas representações ortográficas, mas também por representações fonológicas e relações semânticas entre os itens. Por essa razão, a representação do modelo, como pode ser observado na Figura 3, recebeu um novo grupo de itens sublexicais, tanto ortográfico, quanto fonológico, que interagem entre si e que ativam ou inibem os itens lexicais propriamente ditos.

Dois experimentos realizados por Dijkstra, Hilberink-Schulpen e Van Heuven (2010) foram capazes de replicar o efeito de *priming* ortográfico segundo os princípios do processamento monolíngue de Grainger e Jacobs (1999 *apud* LAM; DIJKSTRA, 2010). Particularmente no segundo experimento, cujas línguas eram o holandês e o inglês, as palavras alvo eram repetidas quatro vezes, cada uma delas acompanhadas por uma condição de *prime* encoberto diferente (i. vizinhos ortográficos do inglês com alta frequência, ii. palavras do inglês sem relação com os alvos, iii. pseudopalavras com relação ortográfica e iv. pseudopalavras sem relação ortográfica). Cada participante viu todas as condições, porém com palavras alvo diferentes.

Os resultados mostraram que tanto palavras da língua inglesa sem relação quanto pseudopalavras sem relação apresentaram efeitos semelhantes. Já as pseudopalavras relacionadas apresentaram efeito de facilitação em todas as quatro condições de *priming*, enquanto as palavras da língua inglesa relacionadas, por outro lado, apresentaram inibição na primeira condição, o de vizinhos ortográficos, mas facilitação foi observada nas condições seguintes. Conforme os autores, os resultados estão em consonância com os pressupostos de Grainger e Jacobs (1999 *apud* LAM; DIJKSTRA, 2010). De acordo com eles, a facilitação de

pseudopalavras relacionadas se deu em razão de sobreposição sublexical. Já o efeito de inibição apresentado na condição de vizinhos ortográficos deve-se à competição lexical entre as línguas.

Com base nesses dados, Lam e Dijkstra (2010) realizaram simulações computacionais com o BIA+ na tentativa de replicar os resultados obtidos. Segundo eles, o modelo apresentou forte inibição lexical para a primeira condição de palavras relacionadas em relação ao de palavras não relacionadas. Além disso, os autores também simularam os efeitos de facilitação na condição de palavras sem relação nos demais blocos. Para isso, eles diminuíram os níveis de ativação em repouso (NAR) dos vizinhos ortográficos e aumentaram os das palavras sem relação, o que fez com que as pseudopalavras, tanto na condição relacionada, quanto na condição sem relação apresentassem praticamente o mesmo tempo de processamento em ciclos. Esses dados se mostraram muito coerentes com os dados empíricos mencionados anteriormente, confirmando a eficiência do BIA+ em simular e replicar tais efeitos.

É importante fazermos aqui uma distinção entre os modelos até então apresentados. O MHR busca oferecer uma visão de como o léxico mental de um bilíngue pode ser organizado. Já o BIA+, por sua vez, busca simular, através de um modelo computacional, a maneira como o processamento bilíngue ocorre. Veja que esses modelos não se sobrepõem, mas explicam coisas distintas com base naquilo a que se propõem.

O modelo Multilink, como veremos a seguir, busca fazer uma junção entre o MHR e o BIA+, apresentando assim uma proposta que tenta explicar não só a forma como a produção e tradução bilíngue ocorre e como a memória lexical é organizada, mas também como o processamento ocorre com base nessa organização.

## 2.4 MODELO MULTILINK

Passados anos do desenvolvimento inicial de modelos como os já mencionados MHR e BIA/BIA+, viu-se a limitação que estes apresentam sobre determinados aspectos do processamento lexical, o que motivou pesquisadores a desenvolver um novo modelo capaz de superar essas limitações. O modelo computacional Multilink é definido por Dijkstra *et al.* (2019, p.659, tradução do autor) como “uma representação precisa, operacionalizada e quantitativa da realidade em um domínio restrito do processamento de informações humanas”. Conforme os autores, o modelo busca dispor uma solução geral para a recuperação de significado e forma das palavras no léxico durante tarefas de reconhecimento e produção.

Um dos aspectos que distingue o Multilink do MHR é que, ao invés de realizar a mediação do acesso lexical da L2 por meio da L1 aos conceitos, define que as representações

semânticas são ativadas pelas representações ortográficas a elas ligadas, e que as representações semânticas fazem a ativação das representações fonológicas, sendo todo o processo não seletivo. Em outras palavras, não há separação entre os léxicos do falante bilíngue, ou algum tipo de processo serial entre as línguas para que o acesso ocorra. Isso quer dizer que o input linguístico irá ativar diversas representações em diferentes níveis, a saber: ortográfico, fonológico e semântico, não dependendo da língua alvo. No que se refere ao BIA+, o modelo apresenta a limitação de simulação de processamento lexical com palavras de até 4 ou 5 letras. Além disso, ele não implementa a relação entre as representações lexicais e as representações semânticas. O Multilink, no entanto, é capaz de simular computacionalmente o processamento de palavras entre 3 e 8 caracteres e permite a simulação de efeitos relacionados a fatores léxico-semânticos, simulando o processamento de palavras consideradas especiais, como as cognatas, o que não era possível no BIA+.

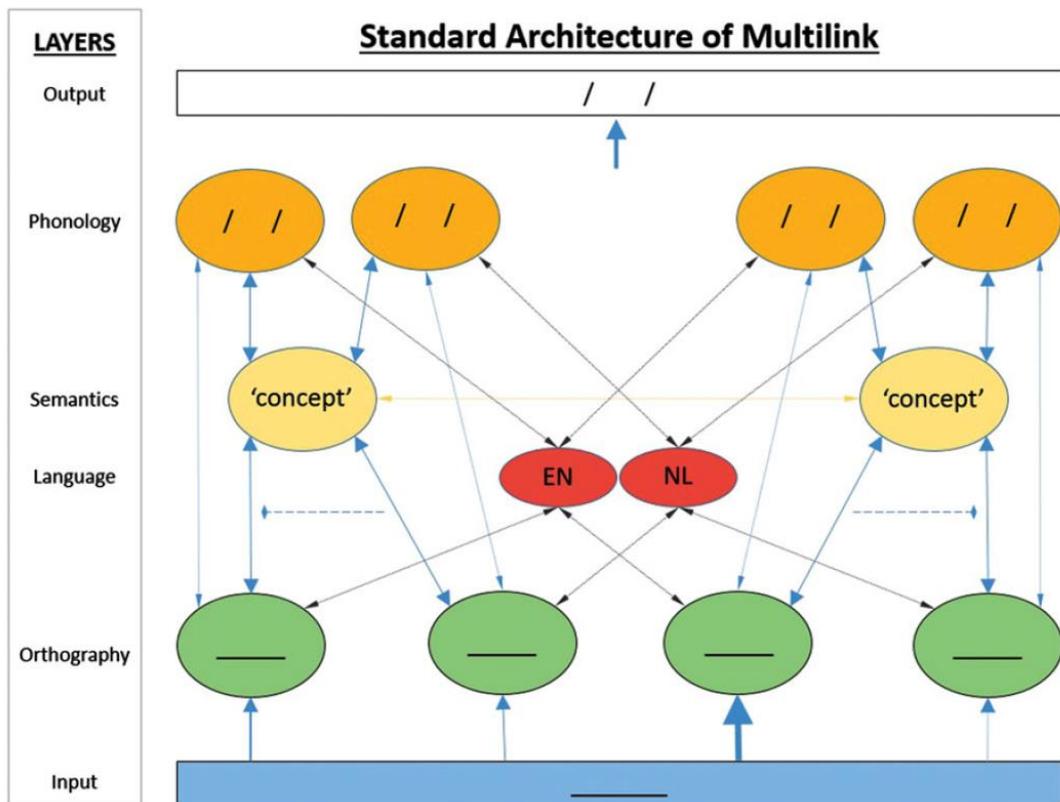
Além do mencionado, o modelo Multilink também permite a simulação do processamento tanto em monolíngues quanto em bilíngues, considerando diversos fatores como a frequência de uso e semelhanças interlinguísticas. Desta forma, ele consegue simular não só o processamento de palavras cognatas, como “*music*” e “*música*”, mas também palavras que possuam apenas semelhanças semânticas, como pares de tradução não-cognatos (*book* e *livro*), bem como permite o ajuste de parâmetros importantes tais como nível de proficiência na L2 e frequência das palavras. Ainda, o modelo utiliza como medida para semelhança ortográfica a distância de Levenshtein relativa (SCHEPENS *et al.*, 2012).

Como já mencionado anteriormente, essa medida mostra-se bastante efetiva no que se refere à verificação da sobreposição ortográfica entre itens lexicais de diversos tamanhos e, para o modelo, ela é usada para determinar o estímulo correto quando múltiplos vizinhos ortográficos ou, no caso dos bilíngues, palavras cognatas são ativadas simultaneamente. Em Post da Oliveira e Van Leussen (2015), os autores investigaram o uso da Distância de Levenshtein Normalizada Interlínguas (INLD) levando em consideração o conhecimento linguístico dos falantes bilíngues para criar o corpus lexical bilíngue baseado na fonologia das palavras.

Os autores usaram um conjunto de textos em inglês e espanhol e calcularam as INLDs entre as palavras das duas línguas. Em seguida, eles usaram um algoritmo de agrupamento para agrupar as palavras em clusters de acordo com suas INLDs. Por fim, os pesquisadores examinaram os clusters e selecionaram aqueles que continham pares de palavras com uma alta INLD, o que indicava uma alta semelhança entre as palavras em ambas as línguas. Esses pares foram então usados para criar o corpus lexical bilíngue.

As conexões realizadas pelo Multilink podem ser observadas na Figura 4 Conforme Dijkstra *et al.* (2019), ao apresentar um estímulo escrito, várias alternativas léxico-ortográficas são ativadas, as quais ativam suas respectivas representações semânticas e fonológicas, assim como as representações referentes às línguas a elas associadas. Por se tratar de um modelo interativo, todo o fluxo de ativação acontece de forma bidirecional, ou seja, tanto há ativação *bottom-up* quanto *top-down*. Após a verificação de todas essas representações, um output é gerado com o resultado que melhor corresponde ao input, ou seja, ao estímulo recebido. Os autores ainda destacam que a versão atual do modelo considera as representações semânticas como unidades holísticas e que não há propagação de ativação entre representações associadas, como “*doctor*” e “*nurse*”.

FIGURA 4 - Representação da arquitetura padrão do Multilink



Fonte: Dijkstra *et al.* (2019)

Assim como no BIA+, o Multilink apresenta um sistema de tarefa/decisão responsável pela seleção da resposta mais adequada para um estímulo recebido. Ou seja, é esse sistema que irá definir os parâmetros e especificar a resposta para um input a depender do estímulo linguístico. Segundo os autores, o sistema de tarefa/decisão pode verificar a família linguística do input e output, bem como o grau de ativação ortográfica, fonológica ou semântica, assim como requisitos para a liberação de uma resposta (DIJKSTRA *et al.*, 2019).

Além disso, o modelo apresenta um sistema de níveis de ativação em repouso (NAR). A diferença entre ele e o BIA+ está no fato de que o NAR apresentado no BIA+ é organizado por um ranqueamento baseado na frequência de cada palavra. Isso quer dizer que os níveis de ativação em repouso para cada palavra serão determinados pela frequência, mas apenas de forma indireta. Isso acontece porque nesse tipo de organização por rank não existe diferenciação para palavras adjacentes. Neste sentido, não há diferença para o BIA+ entre um léxico composto por palavras com frequências de 1, 10 e 100 por milhão de palavras e um léxico composto por palavras com frequências de 1, 2 e 3 por milhão de palavras. Além disso, a adição de apenas uma palavra afeta diretamente a organização de todas as outras, o que não é o ideal, assim como, segundo os autores do Multilink, essa classificação não reflete os tempos de reação obtidos em diversas tarefas experimentais. Conforme Dijkstra *et al.* (2019), duas equações foram aplicadas com o objetivo de mapear a frequência das palavras ao domínio dos NAR no Multilink, levando em consideração a frequência da própria palavra e das palavras com maior e menor frequência dentro do corpus em uso.

Nosso interesse está principalmente nas questões do modelo envolvendo o léxico bilíngue e a representação de palavras cognatas e não cognatas. Para determinar a eficácia do Multilink em simular o processamento bilíngue, Dijkstra *et al.* (2019) utilizaram dados empíricos como base para suas simulações, como por exemplo o estudo de Dijkstra *et al.* (2010) envolvendo uma tarefa de reconhecimento de palavras com decisão lexical.

Em uma das simulações, os autores comparam a performance no reconhecimento de palavras do holandês e do inglês com bilíngues balanceados e não-balanceados com palavras de até 4 letras através do Multilink, do modelo de Ativação Interativa (IA) e do BIA+. Conforme os resultados encontrados, as correlações com os dados empíricos foram mais baixas para o IA e BIA/BIA+ do que para o Multilink. Segundo os autores, isso se deve a forma como o Multilink calcula os níveis de ativação em repouso (NAR). Ainda assim, o modelo se sai melhor que seus antecessores mesmo em tarefas usando palavras com até 4 letras na simulação de performance com bilíngues balanceados e não balanceados (DIJKSTRA *et al.*, 2019).

Em outra série de simulações, os autores utilizaram o estudo de Dijkstra *et al.* (2010) como base, o qual tratava-se de uma decisão lexical envolvendo palavras cognatas por bilíngues não-balanceados de holandês-inglês. Um fator importante para esse estudo é que Dijkstra *et al.* (2010) utilizaram a distância de Levenshtein como uma variável para a relação entre as palavras do inglês e do holandês. Desta forma, eles consideraram palavras que tinham sobreposição ortográfica perfeita, ou seja, cognatas perfeitas, palavras com alto nível de sobreposição ortográfica e palavras com sobreposição ortográfica mínima. Conforme os resultados da

simulação, foi possível observar que o Multilink é capaz de distinguir palavras cognatas e não cognatas, apresentando tempos de reação menores para as primeiras. Segundo Dijkstra *et al.* (2019):

A apresentação de uma palavra cognata ativa duas representações ortográficas similares em ambas as línguas. Essas representações enviam ativação para o nó semântico por elas compartilhado, fazendo com que este tenha uma ativação mais rápida do que quando recebe estímulo de um único nó ortográfico. Por essa razão, a ativação recebida da representação semântica pela camada ortográfica é mais forte e permite que estes fiquem disponíveis mais rápido para palavras cognatas do que não cognatas (DIJKSTRA *et al.*, 2019, p.669).

Esses resultados estão de acordo com os dados empíricos obtidos por Dijkstra *et al.* (2010). Semelhantemente, fora realizada uma simulação com base nos dados de Vanlangendonck *et al.* (*apud* DIJKSTRA *et al.*, 2019) com as mesmas condições do estudo acima e que apresentou resultados semelhantes, reforçando a capacidade do Multilink em simular a facilitação de palavras cognatas em tarefas de decisão lexical.

Em suma, o Multilink é um modelo computacional capaz de simular consideravelmente bem resultados obtidos em estudos empíricos envolvendo o processamento de palavras cognatas e não cognatas, além de diversos outros aspectos em monolíngues e bilíngues. É importante levar em consideração que a versão apresentada por Dijkstra *et al.* (2019) ainda está em fase inicial e que aspectos morfológicos não foram abordados, assim como as unidades semânticas são consideradas como itens holísticos, ou seja, que são completamente compartilhadas ou completamente separadas através das línguas, desconsiderando a interação entre itens semanticamente relacionados. Além disso, o sistema de tarefa/decisão ainda precisa de maiores atualizações. Contudo, ele apresenta-se como uma boa proposta de representação do léxico bilíngue, assim como consegue dar conta de aspectos não considerados por seus antecessores. Em nosso estudo experimental buscamos avaliar com base nos resultados obtidos qual dos modelos (ou quais) poderia explicar melhor os nossos dados.

### 3 EXPERIMENTO 1: JULGAMENTO DE ACEITABILIDADE DE PALAVRAS COGNATAS E NÃO COGNATAS

O objetivo deste experimento foi investigar a partir de uma tarefa *offline* de julgamento de aceitabilidade como falantes do português brasileiro como língua materna não-bilíngues associam palavras cognatas do inglês às suas respectivas traduções. A motivação para a sua realização surgiu do interesse de se verificar quanto a semelhança ortográfica entre pares de palavras cognatas influencia na apreensão de seu significado. Além disso, desejamos correlacionar os resultados do experimento com a Distância de Levenshtein Normalizada (DLN) (SCHEPENS; DIJKSTRA; GROOTJEN, 2012), com o intuito de identificar um possível padrão para a qualificação de palavras cognatas com base no grau de semelhança ortográfica. Ainda, buscamos através deste experimento verificar o status cognato das palavras investigadas para o experimento de *priming* encoberto com decisão lexical.

Conforme Oliveira e Sá (2013), a seleção de estímulos para estudos (psico)linguísticos baseados tão somente na introspecção do autor pode enviesar os seus resultados e diminuir assim a confiabilidade dos seus dados. Por esse motivo, é importante o uso de técnicas para a validação do material selecionado. Na Psicolinguística, conforme Leitão (2008), existem duas formas experimentais distintas denominadas *off-line* e *on-line*. Os experimentos *off-line* são aqueles em que as respostas obtidas são obtidas após os participantes verem ou escutarem um texto ou uma frase, por exemplo. Isso significa que os dados são obtidos após o processamento do estímulo. Já os experimentos *on-line*, tem como característica aferir dados no momento em que o estímulo, seja texto, som, dentre outros, está sendo processado. O julgamento de aceitabilidade é um experimento *off-line* que tem como tarefa a avaliação de palavras ou frases pelos participantes, comumente se usando uma escala Likert, usada para seleção de estímulos linguísticos (OLIVEIRA; SÁ, 2013). Portanto, o participante deve indicar a pontuação na escala de uma palavra ou sentença com uma variação de 1 a 5 (ainda, alguns utilizam uma escala de 1 a 7).

#### 3.1 PARTICIPANTES

Os participantes do presente experimento foram selecionados através do compartilhamento de um link por meio das redes sociais. Antes de realizar o experimento, cada participante foi solicitado a consentir sua participar no experimento de forma voluntária e anônima por meio de um Termo de Consentimento Livre e Esclarecido (TCLE). Além disso,

eles foram orientados a apenas darem continuidade com o experimento se eles autodeclarassem conhecimento nenhum ou básico da língua inglesa. Assim, 117 participantes realizaram o experimento de julgamento de aceitabilidade, sendo a maioria do sexo feminino, entre 18 e 45 anos de idade, onde 93,3% dos participantes declarou não ter nenhum conhecimento de inglês e 6,7% dos participantes declarou ter nível básico de inglês.

Uma das nossas hipóteses foi que mesmo pessoas que não tiveram nenhum contato ou instrução formal do inglês seriam capazes de inferir o significado das palavras na língua inglesa, considerando apenas sua representação formal e semelhança com o português brasileiro.

### 3.2 MATERIAIS

Utilizamos o Léxico do Português Brasileiro (ESTIVALET; MEUNIER, 2017) e o SUBTLEX-UK (VAN HEUVEN *et al.*, 2014) como corpora para seleção de palavras. Primeiramente, as palavras do português brasileiro foram selecionadas como estímulos de base, uma vez que é consideravelmente difícil manter a simetria entre as frequências de ambas as línguas, assim como devido às frequências das palavras do inglês corresponderem a frequências consideradas por nativos da língua e não refletindo o uso de um falante de inglês como L2.

Para escolha das palavras do português, filtramos todos os substantivos singulares entre a frequência 1 e 1000 por milhão (f/M) de palavras. O objetivo dessa filtragem foi encontrar o maior número de palavras candidatas a cognatas com uma frequência semelhante, mas que não apresentassem nem alta e nem baixa frequência, minimizando possíveis efeitos desta métrica sobre os itens experimentais. Foram selecionadas 60 palavras com média de 260f/M e desvio padrão de 58, considerando-se o maior número de palavras desejadas no menor intervalo de frequência possível. Em seguida, controlamos o número de caracteres, sendo a média de caracteres do português de 6,67 e desvio padrão de 1,32, sendo a menor palavra composta por 5 caracteres e a maior palavra composta por 9 caracteres. Após essa seleção, verificamos as traduções das palavras do português brasileiro para o inglês, de maneira que a média de caracteres de todas as 60 palavras foi 6,52 e desvio padrão de 1,57, sendo a menor palavra composta por 5 caracteres e a maior palavra composta por 10 caracteres.

Realizamos um Teste-t para verificar possíveis diferenças entre o número de caracteres das palavras do português e do inglês  $t(59) = 1,21$ ,  $p = 0,1$ , não evidenciando diferenças significativas. Além disso, recuperamos as frequências das palavras do inglês a partir do SUBTLEX-UK, calculamos a distância de Levenshtein (DAVIS; PEREA; ACHA, 2009) e a

DLN (SCHEPENS; DIJKSTRA; GROOTJEN, 2012) entre todos os pares de palavras do português e do inglês selecionados para os nossos materiais.

### 3.3 PROCEDIMENTOS

O experimento foi realizado através da Internet e de forma remota por meio de um formulário elaborado no Google Forms. Nas instruções para o experimento, conforme a Figura 5, os participantes foram esclarecidos sobre a definição de palavra cognatas perfeitas “banana-banana” e palavras não-cognatas “casa-house”. Ainda, as demais cognatas foram descritas como palavras que possuem ortografia semelhante entre as duas línguas e com o mesmo significado entre elas, como as palavras “computador-computer”.

FIGURA 5 – Instruções do experimento de julgamento de aceitabilidade no Google Forms

## JULGAMENTO DE ACEITABILIDADE DE PALAVRAS COGNATAS

Universidade Federal da Paraíba  
Programa de Pós-Graduação em Linguística (PROLING)

\*Obrigatório

---

**Obrigado por colaborar com essa pesquisa!**

Você verá uma sequência de 60 palavras da Língua Inglesa e deverá avaliar o quão cognata elas são. Caso tenha dúvida, palavras cognatas tratam-se de itens lexicais que tem uma escrita idêntica ou muito semelhante a outra palavra de um outro idioma, e que compartilham o mesmo significado entre esses idiomas. Tome por exemplo as palavras abaixo:

(1) banana - banana.  
(2) region - região.  
(3) strength - força.

A palavra (1) pode ser considerada uma cognata perfeita, uma vez que a escrita entre os dois idiomas é idêntica e o significado é o mesmo. A palavra (2) também é uma cognata, pois nós conseguimos deduzir que region significa região. No entanto, não podemos chamá-la de perfeita, uma vez que a escrita difere entre os idiomas. Por fim, (3) não pode ser considerada uma cognata, uma vez que a escrita de strength não nos permite identificar seu equivalente no Português, que é força.

Entendida a noção do que é uma palavra cognata e avaliando em uma escala de 1(pouco cognata) a 5 (muito cognata), julgue as palavras que lhes forem apresentadas. Não existem cognatas perfeitas neste teste.

**OBSERVAÇÃO!** É muito importante para nós que você não seja falante de Inglês. Se você considera seu nível de inglês pelo menos intermediário, por favor, não responda a esse teste.

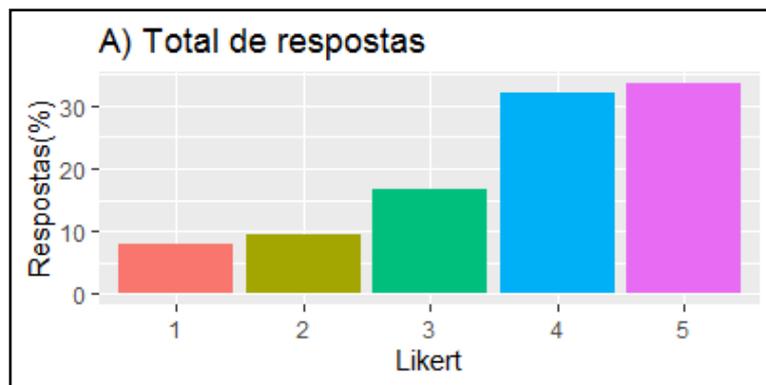
Fonte: Elaborado pelo autor (2022)

Os participantes foram informados que todos os pares de palavras português-inglês do experimento possuíam o mesmo significado e que sua tarefa era avaliar em uma escala Likert de 1 a 5 o quão cognatas essas palavras eram, sendo 1 para muito pouco cognata, 2 para pouco cognata, 3 para cognata, 4 para bastante cognata e 5 para muito cognata.

### 3.4 RESULTADOS E DISCUSSÃO

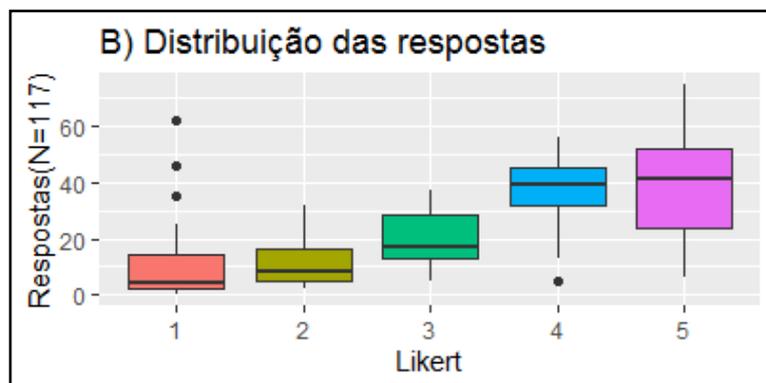
Os dados foram analisados através do programa R (R CORE TEAM, 2014). Destaca-se que a análise realizada possui um caráter de mineração de dados com três objetivos específicos: i. análise descritiva da distribuição dos pares selecionados e dados coletados; ii. análise inferencial das diferenças das respostas em função da DLN; e iii. análise exploratória para determinação do status de palavra cognata. Inicialmente, foram contabilizados o número total de respostas, assim como a distribuição do número de respostas para cada um dos cinco pontos da escala Likert, conforme o Gráfico 1 e o Gráfico 2, respectivamente.

GRÁFICO 1 – Número total de respostas



Fonte: Elaborado pelo autor (2022)

GRÁFICO 2 – Distribuição do número de respostas por participante na escala Likert



Fonte: Elaborado pelo autor (2022)

Em seguida, com objetivo de descrever as características lexicais dos pares de palavras utilizados no experimento de julgamento de aceitabilidade, foram calculadas as médias e desvios-padrão da média, mediana e modo da pontuação na escala Likert das palavras, assim como da distância de Levenshtein, da DLN, do número de caracteres e da frequência das palavras do português e do inglês, conforme a Tabela 1. As médias da escala Likert também

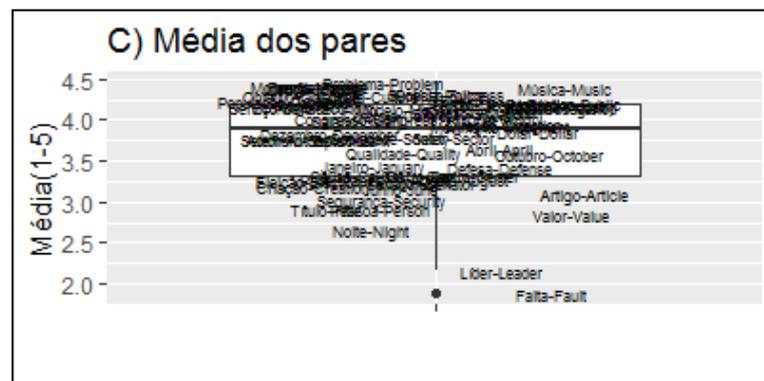
podem ser observadas no Gráfico 3 e no Gráfico 5, e, as DLN também podem ser observadas no Gráfico 4.

TABELA 1 - Média e desvio padrão das respostas dos participantes e das características dos pares de palavras do experimento

<b>Categoria</b>	<b>Média</b>	<b>Desvio-padrão</b>
Média	3,74	0,58
Mediana	3,82	0,81
Modo	4,08	1,06
Distância de Levenshtein	2,30	1,06
Distância de Levenshtein Normalizada	0,67	0,14
Número de caracteres português	6,67	1,32
Número de caracteres inglês	6,52	1,53
Frequência português	5,40 <sup>3</sup>	0,09
Frequência inglês	4,77	0,61

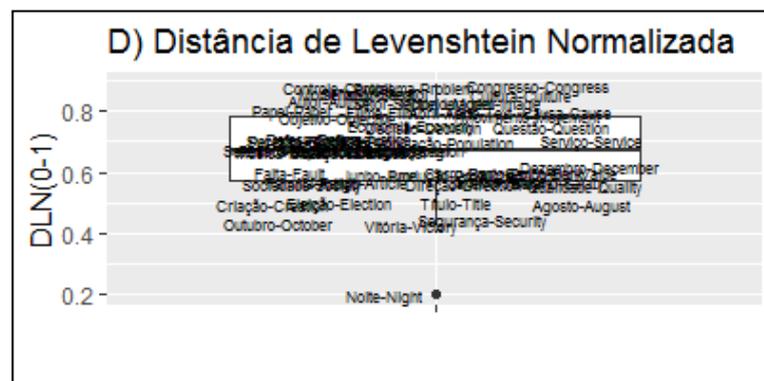
Fonte: Elaborado pelo autor (2022)

GRÁFICO 3 - – Média dos pares de palavras



Fonte: Elaborado pelo autor (2022)

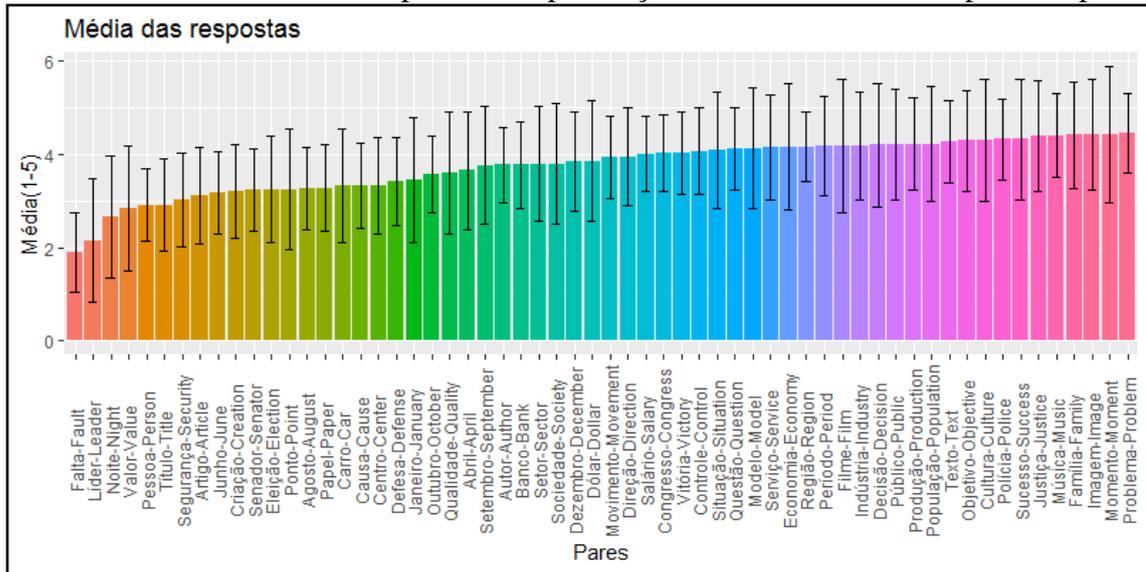
GRÁFICO 4 – Distância de Levenshtein Normalizada para os pares de palavras



Fonte: Elaborado pelo autor (2022)

<sup>3</sup> As médias e os desvios padrão das frequências das palavras do português e do inglês nesta tabela são expressas através da escala Zipf (VAN HEUVEN *et al.*, 2014).

GRÁFICO 5 – Médias e desvios-padrão das pontuações da escala Likert dos pares de palavras



Através da Tabela 2, fica clara a média de 0,67 da DLN dos pares de palavras. Destaca-se a pequena diferença das médias e desvios padrão do número de caracteres das palavras do português e do inglês, assim como o pequeno desvio padrão das frequências do português, evidenciando um bom controle dos materiais escolhidos para o experimento. A média e desvio padrão das frequências das palavras do inglês não puderam ser perfeitamente controladas justamente porque estes dados foram trazidos pelas traduções em inglês do SUBLEX-UK das palavras do português selecionadas primeiramente (VAN HEUVEN *et al.*, 2014). Caso as palavras tivessem sido selecionadas do inglês para o português, é possível que a diferença no nível de ativação das palavras na língua portuguesa e o fato do corpus do inglês ser baseado na proficiência de falante nativos influenciasse os resultados.

TABELA 2 – Análise inferencial das simulações da divisão dos dados em não-cognatas e cognatas a partir da Distância de Levenshtein Normalizada

DLN	Pares de palavras	Teste-t	Kruskal-Wallis
0,4	1/59	-	$X^2(1) = 2,52, p = 0,11$
0,5	8/52	$t(11) = 3,47, p < 0,01$	$X^2(1) = 8,75, p < 0,01$
<b>0,6</b>	<b>23/37</b>	<b><math>t(40) = 3,38, p &lt; 0,01</math></b>	<b><math>X^2(1) = 11,55, p &lt; 0,001</math></b>
0,7	34/26	$t(55) = 3,87, p < 0,001$	$X^2(1) = 11,17, p < 0,001$
0,8	50/10	$t(19) = 2,61, p < 0,05$	$X^2(1) = 3,74, p = 0,05$

Fonte: Elaborado pelo autor (2022)

Tendo em vista que a escala Likert de cinco pontos utilizadas no presente experimento sugere um contínuo na escala de aceitabilidade e semelhança das palavras, utilizou-se como coeficiente da variável dependente a média das respostas dos participantes, ainda, a média foi

o coeficiente que apresentou menor desvio padrão de resposta. Este coeficiente foi calculado a partir da soma de pontuações das respostas da escala Likert (1-5) dividida pelo número total de participantes do experimento (117) (HARPE, 2015).

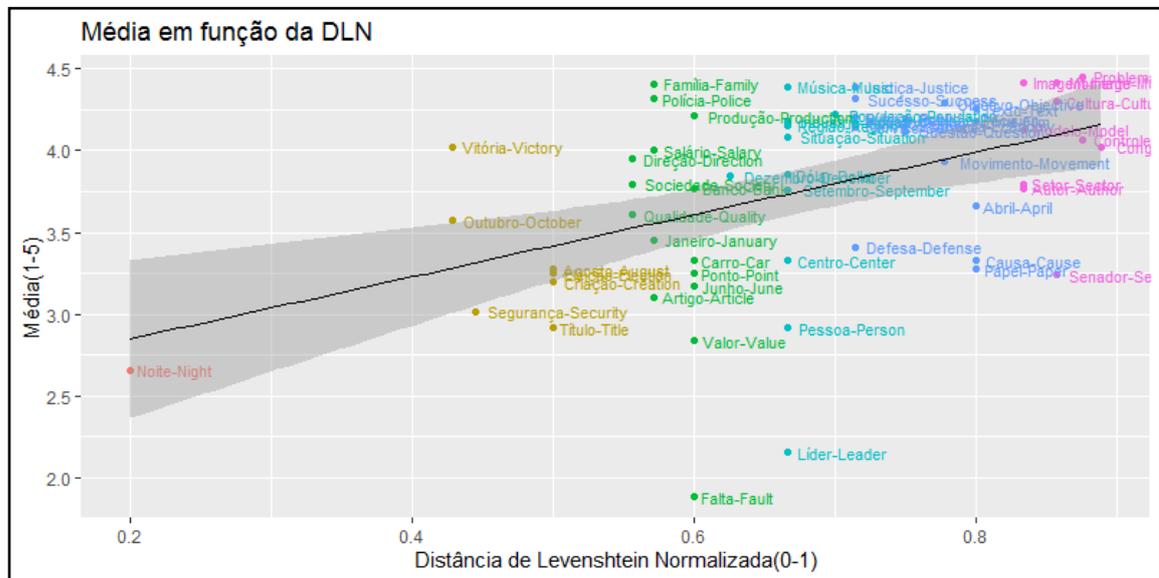
Em seguida, observou-se a distribuição da média das respostas, conforme o Gráfico 3 e o Gráfico 5. O Gráfico 3 evidencia que os pares “falta-*fault*” e “líder-*leader*” foram *outliers* em relação às demais médias e são analisados qualitativamente posteriormente. Portanto, através do Gráfico 5, fica evidente que a maior parte (90%) dos pares de palavras obtiveram médias de pontuação entre 3 e 4,5, ou seja, acima da metade superior da escala Likert utilizada. Estes resultados corroboram nossa hipótese de que mesmo participantes não-bilíngues podem reconhecer e classificar as palavras cognatas com boa assertividade baseados na semelhança formal ortográfica.

Em seguida, o teste de Kolmogorov-Smirnov ( $D = 0,14$ ,  $p = 0,18$ ) evidenciou que os dados apresentam uma distribuição normal, logo, as análises subseqüentes foram realizadas a partir de testes paramétricos. Entre todas as variáveis apresentadas acima (média, mediana, modo, distância de Levenshtein, DLN, número de caracteres e frequência do português e do inglês), o par de variáveis que apresentou maior correlação e resultado mais significativo foi justamente a correlação entre a média das pontuações da escala Likert e a DLN  $t(58) = 3,84$ ,  $p < 0,001$ ,  $r = 0,45$ ,  $IC(95\%) = 0,22-0,63$ , conforme o Gráfico 6.

Em seguida, foi aplicado um modelo generalizado de regressão linear multifatorial com a média das pontuações da escala Likert como variável dependente, e, distância de Levenshtein, DLN, número de caracteres e frequência do português e do inglês como variáveis independentes. A ANOVA do modelo linear apresentou resultados significativos apenas para as variáveis DLN  $F(1,53) = 18,67$ ,  $p < 0,001$  e distância de Levenshtein  $F(1,53) = 12,17$ ,  $p < 0,001$  (DAVIS; PEREA; ACHA, 2009). Este resultado já era esperado justamente porque ambas as variáveis se correlacionam fortemente, tendo em vista que a primeira é a normalização da segunda em função do número de caracteres das palavras comparadas. Enfim, o modelo linear evidenciou como único resultado significativo a DLN  $t(53) = 2,09$ ,  $p < 0,05$ ,  $r^2 = 0,36$ .

Estes resultados corroboram nossa hipótese de trabalho da correlação positiva significativa entre a média do julgamento de aceitabilidade das palavras da escala Likert e a DLN dos pares de palavras  $t(58) = 12,97$ ,  $p < 0,001$ ,  $r = 0,86$ ,  $IC(95\%) = 0,77-0,91$ , como no trabalho de Post e Leussen (2015). Ou seja, quanto menor o número de modificações entre um par de palavras, mais elas são parecidas e, conseqüentemente, mais facilmente elas são identificadas como palavras cognatas, explicando as médias no julgamento de aceitabilidade.

GRÁFICO 6 – Correlação e regressão linear das variáveis média de pontuação da escala Likert em função da Distância de Levenshtein Normalizada. As cores das palavras representam as divisões dos coeficientes da Tabela 3



Fonte: Elaborado pelo autor (2022)

Logo após, foi realizada uma análise da variável numérica independente DLN com o objetivo de se definir o coeficiente que permite a interpretação do status de palavra cognata (SCHEPENS; DIJKSTRA; GROOTJEN, 2012b). Para tanto, realizaram-se análises inferenciais com o teste paramétrico Teste-t e o teste não-paramétrico Kruskal-Wallis de cinco simulações de divisão do conjunto dos 60 pares de palavras em dois (i.e., não-cognatas e cognatas), a partir dos diferentes coeficientes de DLN possíveis, conforme a Tabela 2.

A análise da Tabela 2 aponta que pares de palavras com DLN de 0,5 ou mais já podem ser consideradas cognatas, ou pelo menos, já há algum grau de semelhança formal de semelhança entre as palavras. Contudo, destaca-se que o coeficiente de 0,6 parece ser mais seguro na determinação do status de palavras cognatas, especialmente observando-se os resultados do teste não-paramétricos. Portanto, além da análise formal indicando que DLN maiores que 0,5 deveriam ser consideradas palavras cognatas (SCHEPENS; DIJKSTRA; GROOTJEN, 2012b), os resultados da Tabela 2 corroboram esta hipótese através de um teste de percepção e julgamento de aceitabilidade através da semelhança formal ortográfica.

Finalmente, aplicou-se o teste Tukey a partir as divisões 0-0,4-0,5-0,6-0,7-0,8-0,9 da DLN (conforme as cores do Gráfico 6) para se definirem as diferenças significativas dos diferentes graus de pares de palavras cognatas. Conforme a Tabela 3, observa-se que as palavras cognatas com DLN entre 0,6 e 0,7 são as que possuem diferença mais significativa, seguidas pelas cognatas com DLN entre 0,5 e 0,6. Isso indica mais uma vez que existe uma diferença

significativa no quão cognatas as palavras são entre as escalas de 0,5 a 0,7 em uma escala crescente (quanto maior, mais cognata). Contudo, essa diferença não existe nas escalas entre 0,7 e 0,8, nem entre 0,8 e 0,9, sugerindo que a DLN de 0,6 seria a melhor escolha para se determinar o status cognato das palavras.

Em relação à DLN, o Gráfico 4 e o Gráfico 6 destacam que a DLN de 0,2 de “noite-*night*” foi um *outlier* e que todos os demais pares de palavras possuem uma distância entre 0,4-0,9. Mesmo sendo uma palavra frequentemente encontrada no Brasil, “*night*” recebeu uma pontuação baixa no teste, sobretudo, em função da baixa DLN. Logo, é possível que os participantes reconheçam esta palavra, mas não a consideram cognata da palavra “noite”, tendo em vista sua baixa semelhança formal ortográfica, sugerindo o bom funcionamento e confiabilidade do experimento aplicado, assim como das hipóteses propostas.

TABELA 3 – Teste Tukey (IC 95%) dos diferentes níveis de Distância de Levenshtein Normalizada

NLD	Diferença	Limite inferior	Limite superior	Valor p
0,4-0,5	0,66	-0,99	2,31	0,81
0,5-0,6	0,68	-0,04	1,38	0,05
<b>0,6-0,7</b>	<b>0,74</b>	<b>-0,02</b>	<b>1,51</b>	<b>0,03</b>
0,7-0,8	0,28	-0,32	0,89	0,73
0,8-0,9	0,06	-0,56	0,68	0,91

Fonte: Elaborado pelos autores

Por outro lado, o par de palavras “congresso-*congress*” apresentou a maior DLN de 0,89 em função de ter apenas uma modificação e serem palavras grandes, recebendo média de 4,03. De forma semelhante, o par de palavras “problema-*problem*” também apresenta alta DLN de 0,88 e apresentou a maior média de 4,45 em função de apresentar a modificação apenas na última letra da palavra. Ainda, destacamos que o par de palavras “senado-*senator*” também apresenta DLN alta de 0,86, mas recebeu média de 3,24 em função de apresentar a modificação no meio da palavra (DAVIS; PEREA; ACHA, 2009).

Já em relação às médias da escala Likert, o par de palavras fonologicamente semelhantes “líder-*leader*” com DLN de 0,67 e média baixa de 2,15 indica que o julgamento de aceitabilidade foi feito exclusivamente na forma ortográfica. Diferentemente do inglês, o português apresenta transparência entre ortografia e fonologia, logo, este resultado sugere que não houve influência da forma fonológica semelhante desse par de palavras, uma vez que os participantes não-bilíngues não apresentam nível de inglês suficiente para tal assimilação. Assim, é provável que os participantes leram a palavra “*leader*” usando o padrão fonológico do

português, dificultando a identificação da semelhança formal com a palavra “líder” (COSTA; CARAMAZZA; SEBASTIAN-GALLES, 2000).

Ainda, o par de palavras “falta-*fault*” apresentou tanto uma baixa DLN de 0,60 quanto uma baixa média de 1,88. A diferença formal entre essas palavras parece pequena, mas ao se considerar o número de modificações, assim como o seu tamanho pequeno, justifica-se a dificuldade dos participantes em identificar a semelhança ortográfica. Portanto, corrobora-se que o tamanho das palavras é um fator determinante na medida da semelhança entre pares, onde palavras maiores são menos afetadas do que palavras menores por poucas modificações.

Enfim, destacamos o par de palavras “família-*family*”, recebendo DLN baixa de 0,57 e média alta de 0,41. Interessantemente, sua média foi maior que palavras como “modelo-*model*” que apresenta apenas uma modificação. Esse resultado pode estar relacionado a fatores como a familiaridade dos participantes com o termo. Outra possibilidade é que, apesar da letra “y” ser incomum em palavras do português, a palavra “*family*” já se tornou conhecida entre falantes do português não-bilíngues e, portanto, foi julgada como altamente semelhante, além do fato da letra “y” apresentar a mesma pronúncia da vogal “i” do português.

#### 4 EXPERIMENTO 2: *PRIMING* ENCOBERTO DE PALAVRAS COGNATAS E NÃO COGNATAS

Neste capítulo, investigamos por meio do paradigma de *priming* encoberto com decisão lexical o acesso lexical em bilíngues brasileiros falantes de inglês nos níveis intermediário e avançado de proficiência. Buscamos verificar o efeito de facilitação na tradução de palavras cognatas e não cognatas tanto na direção L1-L2, quanto na direção L2-L1, e como tal efeito pode ser associado aos modelos apresentados no capítulo 2.

Conforme mencionado anteriormente no capítulo 1, o paradigma de *priming* encoberto com decisão lexical se mostra de grande valia para os estudos de acesso e processamento lexical (DIMITROPOULOU; DUÑABEITIA; CARREIRAS, 2011). Além do baixo custo e simplicidade de montagem e execução do experimento de decisão lexical, diversas variáveis podem ser ajustadas com base no que se deseja investigar, tal como o tempo de apresentação do *prime*, tipo e tempo da máscara usada, direcionalidade linguística dos estímulos.

As variáveis independentes que utilizamos para esse experimento foram o tipo de palavra (palavras cognatas e palavras não cognatas), direcionalidade do *priming* (L1-L2 e L2-L1), a variável tipo de *priming* (palavras relacionadas e sem relação) e o nível de proficiência em L2 dos participantes (intermediário ou avançado). Como variáveis dependentes, temos o tempo de reação para a decisão lexical e a acurácia da resposta. As perguntas de pesquisa que buscamos responder com esse experimento foram: o *priming* de palavras cognatas causa mais efeito de facilitação do que palavras não cognatas em bilíngues tardios falantes de português brasileiro como L1 e de inglês como L2? Existe diferença no efeito de *priming* a depender da direcionalidade da língua alvo? Bilíngues avançados reconhecem palavras em ambas as línguas mais rapidamente do que bilíngues intermediários?

Neste experimento, nossa hipótese em relação aos grupos é que existe uma diferença significativa no tempo de decisão lexical entre os bilíngues intermediários e avançados, sendo os últimos mais rápidos do que os primeiros. No que diz respeito à variável de *priming*, nossa hipótese é que as palavras alvo precedidas por *primings* relacionados apresentarão tempos de resposta menores do que os sem relação, uma vez que o *priming* de palavras relacionadas ocasiona a facilitação no acesso lexical, diminuindo o tempo de decisão (FORSTER; DAVIS, 1984).

Para a variável tipo de palavra, esperamos uma diferença significativa entre palavras cognatas e não cognatas, baseados em Dimitropoulou et al (2011) e Wen e Van Heuven (2017), considerando que palavras não cognatas precisam de um tempo maior de exposição do *priming*

encoberto para maiores efeitos de facilitação. Além disso, acreditamos que *primings* na direção L1-L2 apresentarão tempos de reação menores do que *primings* na direção L2-L1, tendo em vista que palavras da L1 apresentam efeitos de facilitação maior sobre a L2 do que a direção contrária, conforme Dunãbeitia *et al.* (2010).

No que diz respeito às interações, referente às palavras cognatas, esperamos que tanto intermediários quanto avançados apresentem tempos de reação menores, conforme Dunãbeitia *et al.* (2010), Dimitropoulou *et al.* (2011) e demais estudos mencionados no capítulo 1 e 2. Mais especificamente, acreditamos que os falantes avançados apresentem tempos de semelhantes para palavras cognatas e não cognatas. Essa hipótese baseia-se no exposto pelos autores supracitados, dentre outros mencionados no capítulo 2, que observaram efeitos de facilitação semelhante em palavras cognatas e não cognatas em bilíngues com nível avançado de proficiência, não havendo distinção para o tipo de palavra do *priming*.

Além disso, acreditamos que haverá efeito de facilitação de *priming* para os participantes em nível avançado de proficiência, tanto na direção L1-L2, quanto na direção L2-L1, reforçando a hipótese de que a L2 passa a ter um status semelhante à língua materna do bilíngue quando atingimos níveis avançados de proficiência. Já para os falantes intermediários, esperamos que haja uma interação entre as condições tipo de palavra, direcionalidade. Acreditamos que haverá efeitos significativos de facilitação apenas na direção L1-L2 no que se refere às palavras não cognatas, conforme Dimitropoulou, Duñabeitia & Carreiras (2011). Como mencionado anteriormente, existe uma assimetria no *priming* de palavras não cognatas sobre a direção do estímulo. Quando o *priming* é apresentado na direção L1-L2, se obtém efeito significativo de facilitação. Porém, quando o *priming* é apresentado na direção L2-L1, normalmente os efeitos são obtidos apenas com falantes avançados da L2.

Como é possível observar, todas as nossas hipóteses envolvem interação entre o nível de proficiência e as demais variáveis, com exceção da condição de tipo de *priming*, em que esperamos resultados semelhantes para ambos os grupos investigados. Mais especificamente, esperamos que haja interação entre o nível de proficiência e o tipo de palavra, em que palavras cognatas e não cognatas apresentarão resultados semelhantes para os avançados, mas diferentes para os intermediários. Por fim, acreditamos que haja interação entre a direcionalidade do *priming* e o tipo de palavra, em que palavras cognatas causarão efeito de facilitação independente da direção do *priming*, mas o mesmo não ocorra para as palavras não cognatas.

Quadro 1 - Hipóteses por variável e interações para o experimento 2.

Variável	Efeito	Explicação
Proficiência	*	Falantes avançados são mais rápidos porque os itens lexicais da L2 tornam-se mais acessíveis no léxico em razão da maior frequência de uso e exposição à língua.
Tipo de priming	*	Alvos com <i>primes</i> relacionados são mais rápidos porque os <i>primes</i> compartilham traços fonológicos/orotográficos e/ou semânticos.
Tipo de palavra	*	Palavras cognatas são mais rápidas porque possuem um status diferenciado no léxico mental.
Direção	*	L1-L2 é mais rápido porque as palavras da L1 apresentam NAR maiores e conseqüentemente ativam mais rapidamente outros itens lexicais, sejam da L1, sejam da L2, causando facilitação no reconhecimento das palavras alvo.
Proficiência x tipo de priming	?	Não temos uma hipótese.
Proficiência x tipo de palavra	*	Apenas intermediários apresentarão efeito de tipo de palavras porque palavras cognatas causam facilitação mesmo em níveis mais baixos de proficiência, enquanto palavras não cognatas causam efeitos de facilitação apenas em níveis altos de proficiência.
Proficiência x direção	*	Apenas intermediários apresentarão efeito de direção porque os itens lexicais da L2 não estão tão consolidados no léxico mental quanto os itens da L1, fazendo com que apenas a L1 gere efeitos de facilitação significativos.
Tipo de priming x tipo de palavra	?	Não temos uma hipótese.
Tipo de priming x direção	?	Não temos uma hipótese.
Tipo de palavra x direção	*	Palavras cognatas causarão efeitos de facilitação em ambas as direções de <i>priming</i> , enquanto as palavras não cognatas causarão efeitos de facilitação apenas na direção L1-L2.

Fonte: Elaborado pelo autor (2022)

#### 4.1 PARTICIPANTES

Todos os participantes consentiram sua participação no experimento de forma voluntária através de um termo de consentimento livre e esclarecido. No total de 33 pessoas participaram do experimento, todos brasileiros, falantes de inglês como LE, sendo que deste número, 20 participantes eram de nível intermediário e os demais 13 de nível alto/avançado.

Para melhor controlarmos o grupo que estava participando da pesquisa, aplicamos um questionário sociolinguístico. Dentre os participantes, 21 eram mulheres (63,64%) e 12 eram homens (36,36%). Todos os participantes eram maiores de idade, entre 18 e 45 anos, com uma idade média de 27,26 anos para o grupo de mulheres e 29,09 anos para homens. Todos estavam

no mínimo cursando graduação. Mais especificamente, 40,63% dos participantes indicaram já possuir ensino superior completo, 37,5% concluíram o ensino médio e os demais declararam possuir mestrado ou doutorado.

Todos os participantes declararam ter nascido e vivido a maior parte da vida no Brasil. Oito dos 33 participantes já moraram fora do Brasil por no mínimo seis meses e no máximo um ano e onze meses, não tendo visitado um país falante de inglês nos últimos dois anos. Três participantes afirmaram ter estudado outro idioma além do próprio português, sua língua materna, antes de iniciar o ensino formal, todos se referindo ao inglês. Além disso, sete dos 33 participantes declararam ter estudado inglês antes dos 10 anos de idade. Após o início da educação formal, todos os participantes indicaram que estudaram língua inglesa, porém o português continuou sendo o idioma usado com mais frequência tanto em ambientes formais quanto informais de comunicação. No total, 84,85% dos participantes indicaram usar a língua inglesa semanalmente ou diariamente, além de declararem sua proficiência na língua inglesa como sendo boa ou excelente.

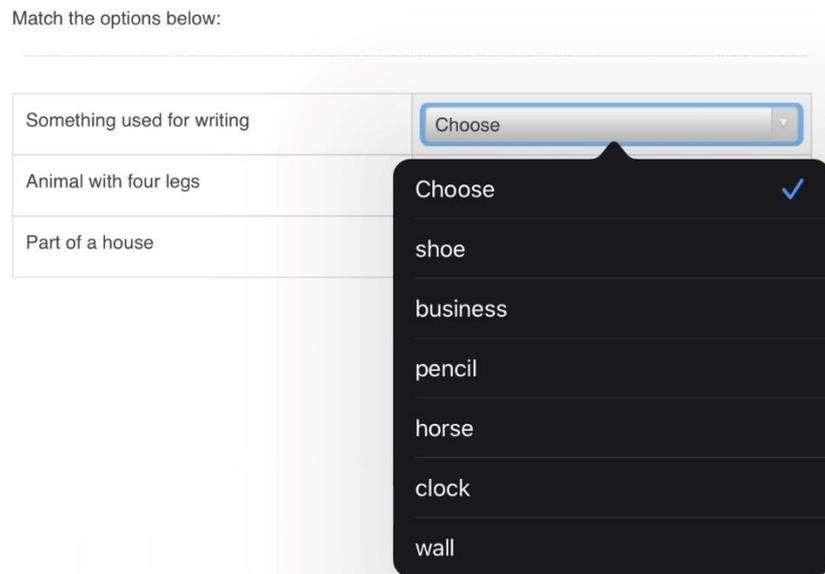
#### 4.2 TESTE DE PROFICIÊNCIA

Com o objetivo de separar os participantes entre os níveis intermediário e avançado de proficiência, realizamos um teste de nivelamento chamado *Vocabulary Levels Test* (NATION, 1990), doravante VLT, adotando as alterações propostas por Souza e colegas (SOUZA; DUARTE; BERG, 2015; SOUZA; SOARES-SILVA, 2015). Segundo esta proposta, o VLT é aplicado em um período máximo de 10 minutos, o que implementa ao teste um componente de automaticidade, conforme os autores, denominado de velocidade de execução de tarefa.

O objetivo do VLT é verificar o tamanho do vocabulário de falantes da língua inglesa. Neste sentido, o teste é composto por cinco níveis, os quais são divididos em seis conjuntos de palavras. Cada conjunto também dispõe de três definições, que o participante deve associar a uma palavra dentre as seis palavras mencionadas anteriormente. Sendo assim, ao todo existem 18 definições para cada nível do teste. Conforme Nation (1990), cada nível avalia o tamanho do vocabulário de um falante de forma progressiva baseada na frequência das palavras que compõe cada um (THORNDIKE; LORGE, 1944). Desta maneira, o nível 1 indica que o participante conhece as 2000 palavras mais frequentes, o nível 2 corresponde às 3000 palavras mais frequentes, o nível 3 às 5000 palavras mais frequentes. O nível 4 trata-se de um nível especial que corresponde ao vocabulário acadêmico e científico. Por fim, o nível 5 se refere ao conhecimento das 10.000 palavras mais frequentes da língua inglesa. Ainda segundo o autor, o

ponto de corte para cada nível é o acerto mínimo de 12 definições dentre as 18 que compõem cada nível. Na Figura 6, é possível observar uma seção de um dos níveis do VLT aplicado de forma online.

FIGURA 6 - Exemplo do VLT no Classmarker



Fonte: Elaborado pelo autor (2022)

De acordo com Souza, Duarte e Berg (2015), em testes com brasileiros universitários falantes de inglês, o nível 4, que corresponde a vocabulário de nível acadêmico e científico, não mostrou diferença significativa para discriminar o conhecimento linguístico dos participantes em relação aos demais níveis do teste. Segundo o argumento dos autores, isso pode ser interpretado como influência da predominância de itens de origem latina neste nível, o que torna tais palavras cognatas para falantes do português brasileiro, que por sua vez causa facilitação durante a leitura.

Souza e Soares-Silva (2015) investigaram a relação entre o tamanho do vocabulário e o nível de proficiência ao comparar o VLT com o *Oxford Placement Test* (OPT). Esse teste trata-se de um nivelamento composto por 60 perguntas de múltipla escolha sobre gramática, compreensão textual e uso do idioma com duração máxima de 30 minutos. Conforme os autores, foi possível observar que os resultados dos participantes que atingiram o nível mais alto do VLT se correlacionaram com os resultados de nível alto no OPT, o qual é mencionado por avaliar de forma muito precisa o nível de proficiência segundo o Quadro de Referência Comum Europeu (CEFR). Da mesma forma, houve correlação entre os resultados de participantes no nível 3 do VLT e as pontuações mais baixas do OPT. Desta forma, conforme

os autores, o VLT se mostra como um teste prático e confiável para verificar o nível de proficiência em língua inglesa.

Com base no exposto acima, utilizamos o VLT para a seleção dos participantes para esta pesquisa. Para que o candidato se classificasse para a tarefa experimental, o mesmo deveria acertar, no mínimo, 12 definições em cada nível. Classificamos o participante no nível intermediário com 12 acertos nos níveis 1, 2 e 3 dentro da janela temporal de 10 minutos. Já para o nível avançado, foram considerados os participantes que obtiveram 12 acertos em todos os níveis do VLT. Uma vez que não foi possível a realização presencial do teste em razão das medidas de distanciamento social por conta da pandemia causada pelo Covid-19, o mesmo foi adaptado para a sua realização de forma remota através do site Classmarker ([www.classmarker.com](http://www.classmarker.com); AQUINO *et al.*, 2020).

#### 4.3 MATERIAIS

O material utilizado foi constituído de 60 palavras cognatas e 60 palavras não cognatas da língua portuguesa, totalizando 120 palavras experimentais. O mesmo foi feito com palavras da língua inglesa. As palavras da língua portuguesa foram traduções de palavras da língua inglesa e vice-versa (cognatas: família-*family*;; não cognatas: terra-*earth*).

As listas de estímulos para esse estudo foram organizadas de maneira que o experimento se dividisse em duas partes: na primeira, que chamamos de experimento 2A, os participantes sempre viram *primes* de palavras da língua inglesa com alvos da língua portuguesa. No segundo, o qual chamamos de experimento 2B, os participantes sempre viram *primes* de palavras da língua portuguesa com alvos da língua inglesa. Sendo assim, quatro listas foram criadas: duas listas para o experimento 2A e duas para o experimento 2B. Essas listas foram organizadas para que todos os participantes fossem apresentados à mesma palavra *alvo*, porém, precedida de *primes* diferentes. Por exemplo, na lista 1 do experimento 2A, o participante foi apresentado ao *prime* “*region*” seguido pela palavra alvo REGIÃO, ou seja, condição de *priming* com relação de palavra cognata. Por outro lado, ao participante que viu lista 2 do experimento 2A, foi apresentado o *prime* “*january*” seguido pela palavra alvo REGIÃO, ou seja, condição de *priming* sem relação. Portanto, definimos a condição controle de palavra sem relação como o *prime* de palavras sem nenhuma relação fonológica, ortográfica e/ou semântica sobre a palavra alvo que o segue.

Além das palavras da língua portuguesa e língua inglesa utilizadas, 2 conjuntos de 120 pseudopalavras cada foram elaborados para a condição de decisão lexical negativa. Cento e

vinte destas pseudopalavras foram baseadas no padrão fonotático da língua portuguesa para o experimento 2A, retiradas do Léxico do Português Brasileiro (ESTIVALET; MEUNIER, 2017). Já as demais 120 pseudopalavras foram baseadas no padrão fonotático da língua inglesa retiradas de Rastle, Harrington e Coltheart (2002) para o experimento 2B. Ainda as pseudopalavras foram apresentadas com dois *primings* diferentes: relacionado (vegiro/VEGIRO) e sem relação (maisão/VEGIRO). Assim, nós pudemos investigar também se os efeitos de facilitação gerados pelas palavras cognatas se tratavam ou não de fatores puramente ortográficos.

Sendo assim, além das condições experimentais geradas a partir da combinação das variáveis independentes, ainda temos a condição controle de pseudopalavras, resultando nas condições experimentais apresentadas na tabela 1, correspondente ao experimento 2A. As mesmas condições foram usadas no experimento 2B, com a diferença na direcionalidade do *priming*, sendo este na direção português-inglês, ou seja, L1-L2.

QUADRO 2 - Condições experimentais referentes ao estudo 2A, na direção L2-L1.

<b>Cognato</b>	<b>Prime ã-rel.</b>	<b>Prime relacionado</b>	<b>Alvo</b>
<b>Cognata</b>	<i>image</i>	<i>region</i>	REGIÃO
<b>Não-cognata</b>	<i>answer</i>	<i>body</i>	CORPO
<b>Pseudopalavra</b>	<i>chilaro</i>	<i>mevisia</i>	MEVISIA

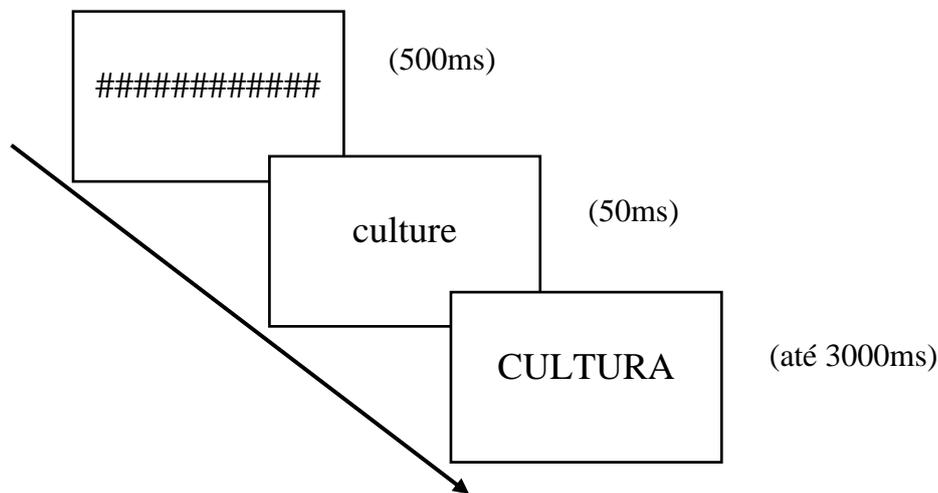
Fonte: Elaborado pelo autor (2022)

Após a criação das quatro listas, realizamos a pseudorandomização das mesmas a partir do software *Mix* (VAN CASTEREN, 2006). A pseudorandomização foi elaborada de maneira que i) nunca fossem exibidas mais que três palavras ou pseudopalavras seguidamente; ii) nunca fossem exibidas mais que três palavras cognatas ou não cognatas seguidamente; iii) nunca houvesse três palavras com relação ou sem relação sendo exibidas seguidamente. Assim, foram criadas 15 listas pseudorandomizadas para cada uma das 4 listas experimentais, garantindo que cada participante fosse apresentado a uma sequência diferente, mas que todos vissem os mesmos itens experimentais. Todos os itens experimentais, assim como as pseudopalavras estão disponíveis no apêndice.

#### 4.4 PROCEDIMENTOS

Os participantes foram apresentados a uma palavra *prime* em caixa-baixa, fonte Courier New, com altura de 0,5 conforme as unidades do Psychopy, com duração de 50ms, precedida por uma máscara composta por 10 marcas *hash* (#) com duração de 500ms. Após o *priming*, a palavra alvo permanecia na tela em caixa-alta por 3000ms ou até que o participante realizasse a decisão lexical, conforme a figura. Para isso, os participantes foram orientados a apertar a tecla P do computador para responder SIM ou a tecla Q do computador para responder NÃO, sendo logo em seguida apresentado ao próximo estímulo. Inicialmente utilizamos o software Psychopy (PEIRCE, 2007; PEIRCE *et al.*, 2019), versão 2020.1, para a elaboração do experimento. Em seguida sincronizamos o mesmo com a plataforma online<sup>4</sup> para aplicação através do Psychopy para web versão 2020.1.

Figura 7 - Representação da apresentação do priming encoberto.



Fonte: elaborado pelo autor (2022)

Durante a aplicação do experimento, uma chamada de vídeo foi realizada com cada um dos participantes. O objetivo dela foi explicar a tarefa para os sujeitos e tirar quaisquer possíveis dúvidas, além de monitorar possíveis interrupções durante a aplicação do experimento. Durante toda a chamada, os participantes utilizaram suas câmeras desligadas, mantendo apenas o microfone ligado. Após a explicação da tarefa, o aplicador manteve seu microfone e vídeo

<sup>4</sup> [www.pavlovia.org](http://www.pavlovia.org)

desligados para evitar quaisquer interrupções. Os participantes realizaram o experimento com um tempo médio de dez minutos e responderam a 240 estímulos experimentais. Após finalizar, uma mensagem de agradecimento era exibida e a janela do experimento se fechava automaticamente após seis segundos.

#### 4.5 RESULTADOS

Um resumo dos resultados do experimento 1A, cujo a direção de *priming* foi L2-L1, pode ser observado na Tabela 4. Conforme a tabela, os participantes apresentaram TRs mais baixos para palavras que para pseudopalavras. Os participantes foram mais rápidos nas condições que a palavra alvo se travava de uma cognata em relação às palavras não cognatas. Ainda, é possível observar que palavras precedidas por um *priming* relacionado levaram os participantes a respostas mais rápidas do que quando precedidas por um *priming* não relacionado, com exceção da condição de palavras não-cognatas por participantes com nível avançado de proficiência. Por fim, destaca-se que em todas as condições os participantes com proficiência avançada foram mais rápidos que os participantes com nível intermediário.

TABELA 4 - Tempos de reação dos participantes para o experimento 1A (L2-L1)

<b>Tipo de palavra</b>	<b>Tipo de relação</b>	<b>Avançados (TR)</b>	<b>Intermediários (TR)</b>
<b>Cognata</b>	<b>Não relacionada</b>	623	716
<b>Cognata</b>	<b>Relacionada</b>	617	696
<b>Não cognata</b>	<b>Não relacionada</b>	627	748
<b>Não cognata</b>	<b>Relacionada</b>	634	736
<b>Pseudopalavra</b>	<b>Não relacionada</b>	715	913
<b>Pseudopalavra</b>	<b>Relacionada</b>	704	917

Fonte: Elaborado pelo autor (2022)

No que diz respeito ao experimento 1B, cujo a direção de *priming* foi L1-L2, podemos observar um resumo dos dados obtidos na Tabela 5 logo abaixo. De forma semelhante ao experimento 1A, os participantes apresentaram TR menores para palavras do que para pseudopalavras. Além disso, os participantes também foram mais rápidos na decisão lexical de palavras cognatas do que não cognatas. Ainda, eles foram mais rápidos quando uma palavra *prime* relacionada era exibida em contrapartida aos *primes* não relacionados. Enfim, os participantes avançados foram mais rápidos do que os participantes no nível intermediário em todas as condições apresentadas.

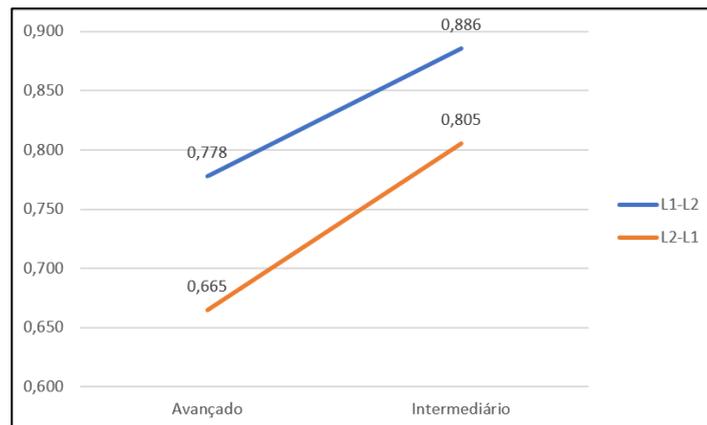
Tabela 5 - Tempos de reação dos participantes para o experimento 1B (L1-L2)

Tipo de palavra	Tipo de relação	Avançados (TR)	Intermediários (TR)
Cognata	Não relacionada	663	801
Cognata	Relacionada	643	746
Não cognata	Não relacionada	698	827
Não cognata	Relacionada	682	818
Pseudopalavra	Não relacionada	899	1052
Pseudopalavra	Relacionada	905	1041

Fonte: Elaborado pelo autor (2022)

No gráfico 7, observa-se que participantes com proficiência avançada apresentam TR mais rápidos que participantes intermediários em todas as direções de *priming*.

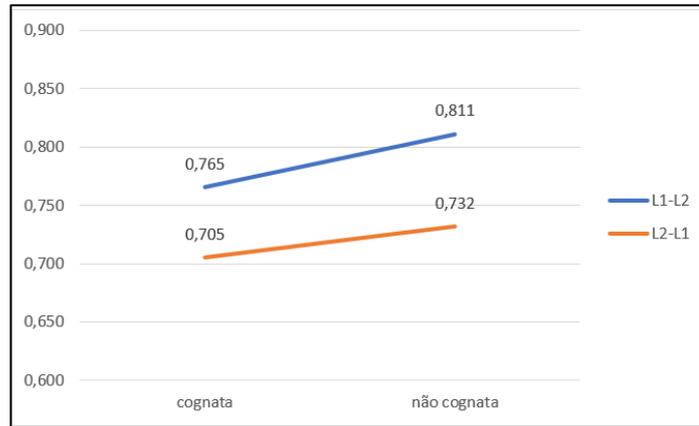
GRÁFICO 7 – Tempos de reação gerais dos participantes nas duas direções de *priming*. Na vertical são apresentados os TRs e na horizontal o nível de proficiência. As cores representam a direção do *priming*.



Fonte: Elaborado pelo autor (2022)

Em seguida, no gráfico 8, pode-se verificar que as palavras cognatas foram mais rápidas do que as palavras não-cognatas, tanto na direção L1-L2, quando na direção L2-L1.

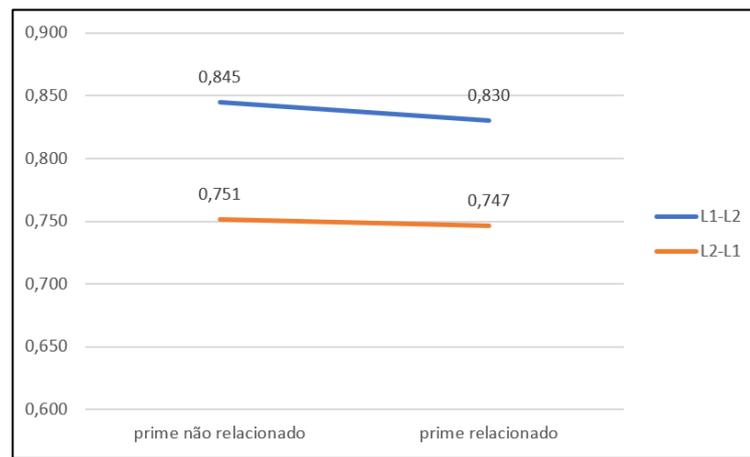
GRÁFICO 8 – Tempos de reação gerais das palavras cognatas e não cognatas em ambas as direções de *priming*. Na vertical estão os TRs e na horizontal o tipo de palavra. As cores representam a direção do *priming*.



Fonte: Elaborado pelo autor (2022)

Já no gráfico 9, observa-se que o TR das palavras precedidas por um *prime* relacionado foi menor do que palavras precedidas por um *prime* não relacionado.

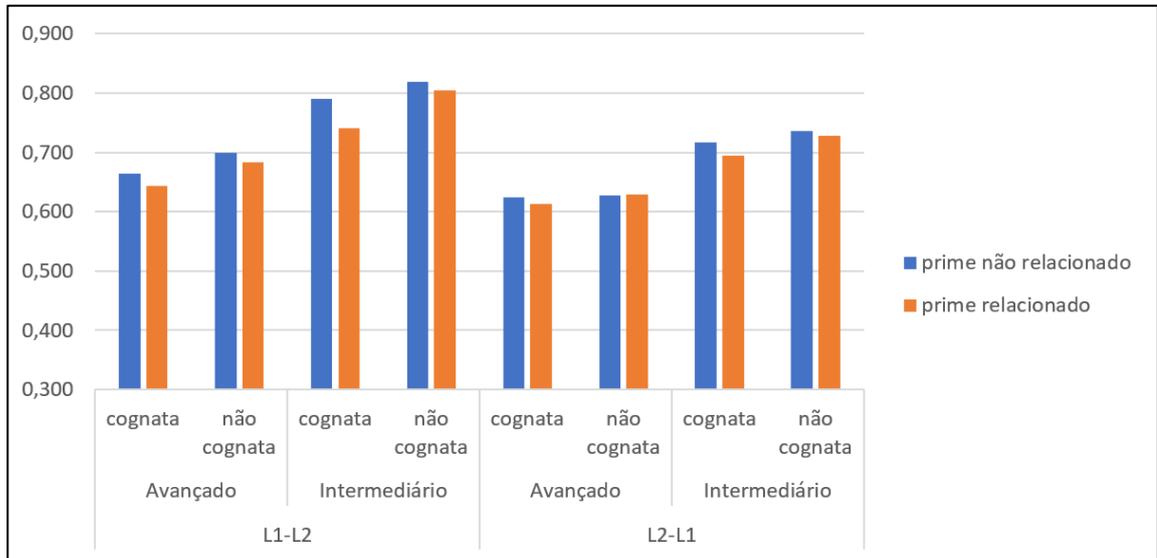
GRÁFICO 9 – Tempos de reação gerais para a variável relação de *priming* em ambas as direções. Na vertical estão os TRs e na horizontal o tipo de *priming*. As cores representam a direção do *priming*.



Fonte: Elaborado pelo autor (2022)

Ainda, é possível verificar no Gráfico 10 como ficaram os tempos de reação para cada variável em todas as condições utilizadas na tarefa experimental. Fica claro que em praticamente todo o experimento, independente do nível de proficiência, do tipo de palavra e a direção, o tempo de reação dos participantes foi menor quando o *priming* era realizado por palavras relacionadas.

GRÁFICO 10 - Diferenças nos tempos de reação de acordo com a direção de priming e do tipo de palavra. Na vertical estão os TRs e na horizontal as condições experimentais de direção de *priming*, proficiência e tipo de palavra. As cores representam o tipo de *priming*.



Fonte: Elaborado pelo autor (2022)

A Tabela 7 abaixo apresenta a média de erros para cada uma das condições experimentais do experimento 2A para os dois níveis de proficiência aqui investigados.

TABELA 7 - Média de erros por condição experimental e grupo de proficiência no experimento 2A

Proficiência	Direção	Palavra	Priming	Média de erros
Avançados	L2-L1	cognata	relacionado	1,03%
Avançados	L2-L1	cognata	Não relacionado	1,29%
Avançados	L2-L1	Não cognata	relacionado	1,54%
Avançados	L2-L1	Não cognata	Não relacionado	0,77%
Intermediários	L2-L1	cognata	relacionado	1,17%
Intermediários	L2-L1	cognata	Não relacionado	2,67%
Intermediários	L2-L1	Não cognata	relacionado	2,17%
Intermediários	L2-L1	Não cognata	Não relacionado	2,00%

Fonte: Elaborado pelo autor (2022)

A tabela 8, por sua vez, apresenta a média de erros dos participantes no experimento 2B para os dois níveis de proficiência.

TABELA 8 - Média de erros por condição experimental e grupo de proficiência no experimento 2B

<b>Proficiência</b>	<b>Direção</b>	<b>Palavra</b>	<b>Priming</b>	<b>Média de erros</b>
Avançados	L1-L2	cognata	relacionado	1,28%
Avançados	L1-L2	cognata	Não relacionado	0,52%
Avançados	L1-L2	Não cognata	relacionado	1,03%
Avançados	L1-L2	Não cognata	Não relacionado	2,31%
Intermediários	L1-L2	cognata	relacionado	3,53%
Intermediários	L1-L2	cognata	Não relacionado	1,97%
Intermediários	L1-L2	Não cognata	relacionado	5,10%
Intermediários	L1-L2	Não cognata	Não relacionado	4,12%

Fonte: Elaborado pelo autor (2022)

Foi realizada uma análise de variância (ANOVA) de efeito fixo em função dos participantes (F1). O resultado indicou que a proficiência foi um fator significativo para a taxa de erros apresentadas pelos participantes, indicando que bilíngues avançados erraram significativamente menos do que os intermediários,  $F(1, 8) = 13,81$ ,  $p < 0,01$ . Os fatores direção, tipo de palavra e relação de *priming* não apresentaram efeitos significativos em relação à média de erro dos participantes.

Uma ANOVA de medidas repetidas foi realizada com o objetivo de verificar a significância dos dados obtidos através da tarefa de decisão lexical com *priming* encoberto. Em uma primeira análise, foi verificado se a diferença nos TRs para palavras era significativa em relação aos TRs para pseudopalavras. Foi observado um efeito principal para palavras existentes,  $F(1, 32) = 147,76$ ,  $p < 0,001$ , em que palavras foram significativamente mais rápidas que pseudopalavras. Após essa análise, realizamos uma segunda ANOVA de efeitos repetidos considerando os TRs como variável dependente e as seguintes variáveis independentes: tipo de palavra, relação de *priming*, direção de *priming* e nível de proficiência.

Foi observado um efeito principal para o nível de proficiência dos participantes,  $F(1, 30) = 14,59$ ,  $p < 0,001$ , indicando que os participantes com proficiência avançada foram significativamente mais rápidos que os participantes com nível intermediário. Também foi identificado um efeito principal significativo de direção de *priming*,  $F(1, 32) = 135,54$ ,  $p < 0,001$ , sendo os TRs na direção L2-L1 significativamente mais rápidos do que na direção L1-L2. Para a variável tipo de palavra, observou-se um efeito principal significativo com os participantes apresentando TRs mais rápidos para identificar as palavras cognatas em relação às palavras não cognatas,  $F(1, 32) = 33,75$ ,  $p < 0,001$ . Por fim, foi identificado um efeito significativo de relação de *priming*,  $F(1, 32) = 11,01$ ,  $p < 0,01$ , indicando que os participantes foram mais rápidos na decisão lexical quando a palavra alvo era precedida por um *priming* relacionado.

#### 4.6 DISCUSSÃO

Conforme esperado, os resultados indicaram que as palavras tiveram TRs significativamente menores que pseudopalavras. Destacamos que usamos uma condição controle dentro das pseudopalavras em que metade delas eram precedidas por um *priming* não relacionado e a outra metade era precedida por um *priming* relacionado, em que a mesma palavra do alvo era exibida como *priming*. Nosso objetivo era identificar se meramente a semelhança ortográfica iria gerar algum efeito de facilitação, se assemelhando às palavras cognatas. Conforme os dados indicados na Tabela 2 e 3, ainda que as pseudopalavras precedidas por um *priming* idêntico tenham apresentado TR menor que as precedidas por um *priming* não relacionado no grupo de avançados, o TR para pseudopalavras é significativamente maior do que para palavras.

É possível observar que as duas maiores médias de erros foram apresentadas pelos participantes no nível intermediário, com *priming* na direção L1-L2 e tendo como alvo palavras não cognatas, sendo a média mais alta a que envolvia um *prime* relacionado a palavra alvo (renda-*income*). Por sua vez, as médias mais baixas foram apresentadas pelos participantes no nível avançado de proficiência. Isso é um indicativo que os sujeitos com alto nível de proficiência apresentam maior precisão na identificação das palavras, refletindo na acurácia apresentada nos dados mencionados anteriormente. Conforme Van Hell, Adams e Abdollahi (2019), bilíngues menos proficientes são mais afetados por fatores contextuais do que bilíngues mais proficientes, corroborando para os dados indicando maior taxa de erros para os intermediários.

Conforme os resultados obtidos, pudemos confirmar algumas das nossas hipóteses principais. Os participantes avançados foram significativamente mais rápidos do que os intermediários. Conforme Dimitropoulou *et al.* (2011), o processamento linguístico de uma segunda língua torna-se mais rápido ao passo que o aprendiz aumenta o seu nível de proficiência. Os modelos BIA+ e Multilink indicam que quanto maior o nível de proficiência, mais rápido se torna o processo do acesso lexical, uma vez que a diferença nos níveis de ativação em repouso (NAR) para as palavras de ambas as línguas do bilíngue consequentemente diminui (DIJKSTRA *et al.* 2019).

Nesse sentido, segundo Dijkstra *et al.* (2019), o processamento lexical de uma língua está relacionado a quão ativos estão seus itens lexicais na memória. Desta maneira, a língua materna apresenta maiores níveis de ativação em repouso no léxico mental, uma vez que está em frequente uso pelo falante. Por esta razão, itens lexicais da L2 tendem a apresentar

representações menos frequentes e consistentes do que os itens lexicais da L1, refletindo no tempo de processamento desses itens durante o acesso lexical. Contudo, ao passo que o bilíngue avança na proficiência, é possível identificar uma aproximação contínua entre o processamento da L1 e L2. Essa aproximação pode ser considerada como um reflexo do uso que os bilíngues fazem da L2. Em outras palavras, quanto mais se avança na proficiência de uma L2, mais frequentes os itens lexicais dela tendem a ser no léxico mental, como reflexo do uso que é feito dessa língua. Como observado nos questionários sociolinguísticos mencionados anteriormente, os bilíngues avançados informam estar mais imersos no uso da L2, mesmo que esta não seja uma língua falada oficialmente no Brasil, por diferentes formas – filmes e séries, música, trabalho, relacionamentos, dentre outros. Assim, as palavras da L2 conseqüentemente tendem a estar mais ativas no léxico mental do que dos bilíngues intermediários, que apresentam uma exposição inferior à língua.

Conforme a Weaker Links hypothesis de Gollan et al. (2008), a diferença no processamento de uma L2 está relacionada à força das conexões dos aspectos semânticos, ortográficos e fonológicos. Desta maneira, os bilíngues com nível intermediário têm conexões mais fracas entre esses aspectos na L2 do que os de nível avançado de proficiência.

De acordo com Gollan *et al.* (1997), Jiang e Forster (2001), Voga e Grainger (2007), Duñabeitia, Perea e Carreiras (2010), dentre outros, *primings* na direção L2-L1 causam pouco ou nenhum efeito de facilitação, o que ficou conhecido como a assimetria de *priming*. Conforme o modelo computacional Multilink, a facilitação causada pelo *priming* de palavras está relacionada ao número de fatores lexicais que estão sendo ativados pelo *input*, a saber, ortográficos, fonológicos e/ou semânticos e o NAR, como mencionado anteriormente. As palavras da L2 tendem a apresentarem NAR menores do que das palavras da L1, principalmente em bilíngues com níveis de proficiência mais baixos. Dessa forma, seria mais difícil para tais palavras da L2 causarem efeitos de facilitação igual ou superior a palavras da L1.

Logo após, foi possível observar um efeito significativo para a variável tipo de palavra, indicando que as palavras cognatas são identificadas mais rapidamente do que as palavras não cognatas. As cognatas apresentaram TRs menores que as palavras não cognatas para ambos os níveis de proficiência, conforme nossas hipóteses. As palavras cognatas, conforme mencionado anteriormente por Sánchez-Casas e García-Albea (2005), possuem um status diferenciado no léxico, o qual é conhecido como status cognato. Por esse motivo, elas são acessadas mais rapidamente do que outras palavras que não sejam cognatas.

Uma vez que palavras não cognatas compartilharam apenas alguns traços semânticos em comum, elas irão competir com diversos outros itens lexicais que compartilham esses traços

assim como outras palavras, como é o caso de vizinhos ortográficos. Por exemplo, a palavra “fogo” compartilha o mesmo significado da palavra “fire”. Porém, ela também compartilha traços semânticos com palavras como “fogão”, “fogueira”, assim como traços ortográficos e fonológicos com palavras como “rogo”, “jogo”, “figo”, dentre outros. Por essa razão, essas palavras podem competir entre si no léxico, o que pode influenciar no tempo necessário para seu reconhecimento. Além disso, no que diz respeito ao NAR, palavras da L2 tendem a ter NARs mais baixos do que palavras da L1. Com base nisso, as palavras da L2 levam mais tempo para serem ativadas do que palavras da L1 usadas com mais frequência. Em tarefas de decisão lexical com *priming* encoberto em que normalmente o tempo ISI é consideravelmente curto, o item lexical da L2 não tem tempo suficiente para ser selecionado e causar o efeito de facilitação sobre a L1.

De acordo com o Modelo Multilink de Dijkstra *et al.* (2019, p. 658), “a apresentação de uma palavra cognata ativa duas representações ortográficas semelhantes”. Assim, a ativação semântica no léxico mental acontece de forma ainda mais rápida, visto que o nó em questão recebe mais de um estímulo ortográfico ao mesmo tempo. Davis *et al.* (2010) afirmam que as palavras cognatas possuem representações que transcendem qualquer outra especificidade no léxico mental, o que explicaria o porquê de elas apresentarem TRs menores em decisões lexicais do que outras palavras. Enfim, a distinção das cognatas pode ser atribuída ao seu status diferenciado no léxico mental, o que daria a este tipo de item lexical uma representação distinta dos demais.

No experimento 1, foi observado uma forte relação a DLR e ao julgamento do status cognato de palavras mesmo por participantes com pouco ou nenhum conhecimento da língua inglesa. A DLR mostra-se como um parâmetro bastante eficaz no que se refere a medição de quão cognata uma palavra é. É possível que haja níveis de cognitude (ou níveis de status cognato) distintos no léxico mental, assim como conseguimos estabelecer com a DLR.

Ainda, identificamos efeitos significativos no que diz respeito a variável relação de *priming*. Ou seja, quando as palavras alvo eram precedidas por *primes* relacionados, a decisão lexical foi realizada mais rapidamente pelos participantes. Isso quer dizer que, quando uma palavra como “*family*” era precedida pela palavra “família”, o tempo de decisão dos participantes era menor do que se a palavra “*family*” fosse precedida por “aeroporto”, por exemplo.

Ao observarmos os resultados para cada direção de *priming* para cada grupo de participantes, verificamos a facilitação dos *primings* em ambas as direções, ou seja, tanto no experimento 2A quanto 2B, além de efeitos significativos para o tipo de palavra no grupo

intermediário de proficiência. Contudo, não houve efeito significativo de facilitação para os participantes no nível avançado em nenhuma das direções de *priming*. Os resultados dos participantes intermediários estão de acordo com nossas hipóteses, em que esperávamos encontrar efeitos de facilitação de palavras cognatas em ambas as direções de *priming*. Interessantemente, o efeito de facilitação foi maior na direção L2-L1 do que na direção L1-L2, como esperávamos. Os participantes, tanto intermediários quanto avançados, são influenciados pela língua alvo no momento da decisão lexical. Dessa forma, explica-se o fato de todos os participantes apresentarem TRs menores quando as palavras alvo estavam na L1.

No que diz respeito às palavras cognatas e os avançados, esperávamos que elas causassem efeito de facilitação tanto para os participantes no nível avançado, quanto intermediário, porém, não obtivemos tal resultados. Apesar de estudos relatarem com mais frequência a existência de efeitos de facilitação por palavras cognatas na direção L2-L1 que para palavras não cognatas (DE GROOT; NAS, 1991; DAVIS *et al.*, 2010), Duñabeitia, Perea e Carreiras (2010) mencionam que normalmente esses efeitos são consideravelmente menores do que na direção L1-L2. Mesmo assim, não obtivemos resultados significativos na direção L1-L2 para este grupo.

Mesmo não obtendo efeitos de facilitação estatisticamente significativos para os participantes avançados, é possível observar uma diferença interessante entre o TR de palavras cognatas e palavras não cognatas. Além disso, as análises considerando os dois grupos apontaram efeitos principais de facilitação de palavras e de *priming*. Mesmo sem efeitos significativos, as médias obtidas a partir da análise descritiva dos dados estão na direção das nossas hipóteses, em que as palavras precedidas por *primings* relacionados tiveram TRs um pouco menores.

No seu experimento, Duñabeitia, Perea e Carreiras (2010) compararam bilíngues simultâneos e balanceados. Conforme os autores, bilíngues tardios ou bilíngues não balanceados dificilmente poderiam apresentar efeitos de facilitação em ambas as direções de *priming*, visto que existiria uma desproporção no nível de proficiência dos falantes e consequentemente isso afetaria a forma como eles processam as duas línguas. O modelo Multilink, de Dijkstra *et al.* (2019) traz a proposta de que a ativação dos itens lexicais se dá conforme os níveis de ativação em repouso deles no léxico mental. Assim, isso não teria uma relação direta com o tempo de aquisição da língua ou exatamente com quão proficiente os sujeitos são, mas o quão ativas esses itens lexicais estão no léxico mental deles.

Em nosso experimento, verificamos que os sujeitos com nível intermediário de proficiência foram capazes de apresentar efeitos de facilitação de *priming*, tanto de palavras

cognatas, como de palavras não cognatas, em ambas as direções de *priming*. Semelhantemente, observamos tempos de reação menores para os bilíngues avançados quando os *primes* eram palavras relacionadas, em ambas as direções, para os dois tipos de palavra. Interessantemente, diferente do que foi proposto por Duñabeitia, Perea e Carreiras (2010), os participantes com nível intermediário, não simultâneos e não balanceados apresentaram efeitos de facilitação para as duas direções de *priming*, enquanto os avançados não tiveram efeito de facilitação significativo.

## CONSIDERAÇÕES FINAIS

Este trabalho teve como proposta investigar o processamento lexical de palavras cognata em falantes bilíngues do português brasileiro como L1 e do inglês como L2 nos níveis intermediário e avançado de proficiência, através do paradigma experimental de *priming* encoberto com decisão lexical.

Para isso, realizamos o experimento 1 de julgamento de aceitabilidade com o objetivo de compreender melhor o status cognato de pares de tradução português/inglês que compartilham semelhanças ortográficas, fonológicas e semânticas. A partir dos dados obtidos, comparamos os resultados com a DLR para identificar o grau de correlação entre a semelhança ortográfica e a aceitação das palavras utilizadas no que diz respeito ao seu status cognato. Conseguimos identificar que existe uma correlação significativa entre a semelhança da DLR das palavras e a pontuação atribuída pelos participantes no julgamento de aceitabilidade.

Em seguida, utilizamos as palavras usadas no julgamento de aceitabilidade no experimento 2 de *priming* encoberto com decisão lexical de palavras, em que os participantes foram expostos a palavras cognatas e não cognatas com *priming* em uma língua e palavras-alvo em outra. Os resultados indicaram que os participantes no nível avançado de proficiência são significativamente mais rápidos do que os participantes no nível intermediário. Além disso, verificamos que os participantes foram mais rápidos quando o experimento foi realizado com *primes* na L2 e palavras alvo na L1, diferente do que alguns dos trabalhos analisados apontam (DAVIS *et al.*, 2010). Ainda, observamos efeitos significativos de facilitação para participantes em nível intermediário em ambas as direções de *priming*, porém não para os participantes com alto nível de proficiência.

Salientamos que, apesar dos resultados confirmarem parte das nossas hipóteses referentes ao tipo de palavra e o tipo de *priming*, eles estão de acordo com estudos realizados em outros idiomas (GOLLAN *et al.*, 1997; JIANG; FORSTER, 2001; KIM; DAVIS, 2003; VOGA; GRAINGER, 2007). A assimetria no *priming* encoberto de traduções é um fenômeno comumente observado. Buscamos verificar, no entanto, se bilíngues tardios com alto nível de proficiência na L2 conseguiriam replicar os resultados apresentados por bilíngues simultâneos balanceados, conforme apresentado por Duñabeitia, Perea e Carreiras (2010), em que essa assimetria não acontece. Sendo assim, nossos resultados corroboram com a afirmação dos autores no que diz respeito aos bilíngues tardios e não balanceados. Contudo, foi observado efeitos de facilitação por parte dos bilíngues no nível intermediário de proficiência, diferindo com o que os autores descreveram. A saber, os participantes da pesquisa realizada por

Duñabeitia e colaboradores utilizaram ambas as línguas, espanhol e basco, cotidianamente e não apresentavam distinção na proficiência das duas.

É importante destacar que nosso experimento foi realizado de forma diferente dos experimentos presenciais tradicionais, isto é, ele foi realizado remotamente através de uma plataforma online e um navegador Web. Por se tratar de um experimento realizado de forma remota, não foi possível controlar por exemplo o espaço em que o participante o realizou, sendo possível a distração por meio de barulhos, pessoas, dentre outros. Ainda, questões envolvendo a atenção, conforto, cansaço e estado emocional dos participantes pode ter influenciado de alguma forma os dados, visto que muitas pessoas transformaram seus lares em seu ambiente de trabalho, além de terem sido afetadas pela pandemia causada pelo covid-19 de diversas maneiras. Mesmo com todos os controles e instruções realizadas para a realização do experimento, o monitoramento e aplicação ficaram bastante limitados por conta da mudança para a forma remota.

Acreditamos que pesquisas futuras com um número de participantes maior para cada grupo de proficiência, especialmente os avançados, pode trazer dados ainda mais precisos com relação aos efeitos de *priming* de palavras cognatas e não cognatas com brasileiros falantes de inglês como L2. Ainda, experimentos em que as palavras cognatas estejam inseridas dentro de sentenças, ao invés de apresentadas isoladamente, podem contribuir para não só o estudo sobre o status cognato a nível ortográfico, fonológico e semântico, mas também no nível sintático. Além disso, é possível que os diferentes níveis da DLR tenham relação com o tamanho do efeito de *priming* causado pelas palavras cognatas. Um experimento verificando essas questões pode trazer informações bastante importantes para a psicolinguística e os estudos envolvendo essas palavras.

Com essa pesquisa, verificamos que é possível obter efeitos de facilitação com bilíngues mesmo que intermediários e não balanceados, em diferentes direções de *priming*. Isso abre novamente a discussão sobre a relação entre o nível de proficiência de bilíngues e a assimetria do *priming* encoberto, já tão conhecida no escopo da psicolinguística experimental. Além disso, traz dados sobre os efeitos de *priming* de palavras cognatas e não cognatas em duas direções de *priming* para o português brasileiro e o inglês, preenchendo uma lacuna existente há anos para os estudos de bilinguismo.

## REFERÊNCIAS

- AQUINO, Estela Maria Motta Lima Leão de *et al.* Social distancing measures to control the COVID-19 pandemic: potential impacts and challenges in Brazil. *Ciência & Saúde Coletiva*, Rio de Janeiro, v. 25, supl. 1, p. 2423-2446, Jun. de 2020.
- ALTARRIBA, Jeanette. The representation of translation equivalents in bilingual memory. *In: Cognitive processing in bilinguals*. Oxford, England: North-Holland, 1992. (Advances in psychology, 83.).p. 157–174. *E-book*. Disponível em: [https://doi.org/10.1016/S0166-4115\(08\)61493-4](https://doi.org/10.1016/S0166-4115(08)61493-4)
- BASNIGHT-BROWN, Dana M.; ALTARRIBA, Jeanette. Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory and Cognition*, [S. l.], v. 35, n. 5, p. 953–965, 2007. Disponível em: <https://doi.org/10.3758/BF03193468>
- BRYLSBAERT, Marc; DUYCK, Wouter. Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? *Bilingualism*, [S. l.], v. 13, n. 3, p. 359–371, 2010. Disponível em: <https://doi.org/10.1017/S1366728909990344>
- COLTHEART, M. *et al.* Access to internal lexicon. *In: Attention and performance IV*. [S. l.: s. n.]. *E-book*.
- COSTA, Albert; CARAMAZZA, Alfonso; SEBASTIAN-GALLES, Nuria. The Cognate Facilitation Effect: Implications for Models of Lexical Access. *Journal of Experimental Psychology: Learning Memory and Cognition*, [S. l.], v. 26, n. 5, p. 1283–1296, 2000. Disponível em: <https://doi.org/10.1037/0278-7393.26.5.1283>
- DAVIS, Chris *et al.* Masked translation priming: Varying language experience and word type with Spanish-English bilinguals. *Bilingualism*, [S. l.], v. 13, n. 2, p. 137–155, 2010. Disponível em: <https://doi.org/10.1017/S1366728909990393>
- DAVIS, Colin J.; PEREA, Manuel; ACHA, Joana. Re(de)fining the Orthographic Neighborhood: The Role of Addition and Deletion Neighbors in Lexical Decision and Reading. *Journal of Experimental Psychology: Human Perception and Performance*, [S. l.], v. 35, n. 5, p. 1550–1570, 2009. Disponível em: <https://doi.org/10.1037/a0014253>
- DE GROOT, Annette M. B. Word-Type Effects in Bilingual Processing Tasks. *In: [S. l.: s. n.]. p. 27. E-book*. Disponível em: <https://doi.org/10.1075/sibil.6.04gro>
- DE GROOT, Annette M. B.; NAS, Gerard L. J. Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, [S. l.], v. 30, n. 1, p. 90–123, 1991. Disponível em: [https://doi.org/10.1016/0749-596X\(91\)90012-9](https://doi.org/10.1016/0749-596X(91)90012-9)
- DIJKSTRA, T. O. N. *et al.* Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism*, [S. l.], v. 22, n. 4, p. 657–679, 2019. Disponível em: <https://doi.org/10.1017/S1366728918000287>
- DIJKSTRA, Ton. Bilingual Visual Word Recognition and Lexical Access. *In: Handbook of bilingualism: Psycholinguistic approaches*. Dijkstra, Ton: Nijmegen Institute for Cognition and Information, Radboud University of Nijmegen, P.O. Box 9104, Nijmegen, Netherlands,

NL-6500 HE, t.dijkstra@nici.ru.nl: Oxford University Press, 2005. p. 179–201. *E-book*.

DIJKSTRA, Ton *et al.* How cross-language similarity and task demands affect cognate recognition. **Journal of Memory and Language**, [*S. l.*], v. 62, p. 284–301, 2010. Disponível em: <https://doi.org/10.1016/j.jml.2009.12.003>

DIJKSTRA, Ton; VAN HEUVEN, Walter J. B. The BIA model and bilingual word recognition. *In: Localist connectionist approaches to human cognition*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 1998. (Scientific psychology series.)p. 189–225. *E-book*.

DIJKSTRA, Ton; VAN HEUVEN, Walter J. B. The architecture of the bilingual word recognition system: From identification to decision. **Bilingualism: Language and Cognition**, [*S. l.*], v. 5, n. 3, p. 175–197, 2002. Disponível em: <https://doi.org/10.1017/s1366728902003012>

DIMITROPOULOU, Maria; DUÑABEITIA, Jon Andoni; CARREIRAS, Manuel. Masked translation priming effects with low proficient bilinguals. **Memory and Cognition**, [*S. l.*], v. 39, n. 2, p. 260–275, 2011. Disponível em: <https://doi.org/10.3758/s13421-010-0004-9>

DREWS, Etta; ZWITSERLOOD, Pienie. Morphological and Orthographic Similarity in Visual Word Recognition. **Journal of Experimental Psychology: Human Perception and Performance**, [*S. l.*], 1995. Disponível em: <https://doi.org/10.1037/0096-1523.21.5.1098>

DUÑABEITIA, Jon Andoni *et al.* Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. **Brain Research**, [*S. l.*], v. 1359, p. 142–154, 2010. Disponível em: <https://doi.org/10.1016/j.brainres.2010.08.066>

DUÑABEITIA, Jon Andoni; PEREA, Manuel; CARREIRAS, Manuel. Masked translation priming effects with highly proficient simultaneous bilinguals. **Experimental Psychology**, [*S. l.*], v. 57, n. 2, p. 98–107, 2010. Disponível em: <https://doi.org/10.1027/1618-3169/a000013>

DUYCK, Wouter; WARLOP, Nele. Translation priming between the native language and a second language new evidence from dutch-french bilinguals. **Experimental Psychology**, [*S. l.*], v. 56, n. 3, p. 173–179, 2009. Disponível em: <https://doi.org/10.1027/1618-3169.56.3.173>

ESTIVALET, Gustavo Lopez; MEUNIER, Fanny. Corpus psicolinguístico Léxico do Português Brasileiro. **Revista SOLETRAS**, [*S. l.*], v. 0, n. 33, p. 212.229, 2017. Disponível em: <https://doi.org/10.12957/soletras.2017.29702>

FINKBEINER, Matthew *et al.* The role of polysemy in masked semantic and translation priming. **Journal of Memory and Language**, [*S. l.*], v. 51, p. 1–22, 2004. Disponível em: <https://doi.org/10.1016/j.jml.2004.01.004>

FORSTER, Kenneth I. The Pros and Cons of Masked Priming. **Journal of Psycholinguistic Research**, [*S. l.*], v. 27, n. 2, p. 203–233, 1998. Disponível em: <https://doi.org/10.1023/A:1023202116609>

FORSTER, Kenneth I.; DAVIS, Chris. Repetition priming and frequency attenuation in lexical access. **Journal of Experimental Psychology: Learning, Memory, and Cognition**,

[*S. l.*], v. 10, n. 4, p. 680–698, 1984. Disponível em: <https://doi.org/10.1037/0278-7393.10.4.680>

FROST, Ram; FORSTER, Kenneth I.; DEUTSCH, Avital. What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. **Journal of Experimental Psychology: Learning Memory and Cognition**, [*S. l.*], 1997. Disponível em: <https://doi.org/10.1037/0278-7393.23.4.829>

GOLLAN, Tamar H.; FORSTER, Kenneth I.; FROST, Ram. Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. **Journal of Experimental Psychology: Learning Memory and Cognition**, [*S. l.*], 1997. Disponível em: <https://doi.org/10.1037/0278-7393.23.5.1122>

GRAINGER, Jonathan; COLÉ, Pascale; SEGUI, Juan. Masked morphological priming in visual word recognition. **Journal of Memory and Language**, [*S. l.*], 1991. Disponível em: [https://doi.org/10.1016/0749-596X\(91\)90042-I](https://doi.org/10.1016/0749-596X(91)90042-I)

GRAINGER, Jonathan; DIJKSTRA, Ton. On the Representation and Use of Language Information in Bilinguals. **Advances in Psychology**, [*S. l.*], v. 83, n. C, p. 207–220, 1992. Disponível em: [https://doi.org/10.1016/S0166-4115\(08\)61496-X](https://doi.org/10.1016/S0166-4115(08)61496-X)

GROOTSWAGERS, Tijn. A primer on running human behavioural experiments online. **Behavior Research Methods**, [*S. l.*], v. 52, n. 6, p. 2283–2286, 2020. Disponível em: <https://doi.org/10.3758/s13428-020-01395-3>

GROSJEAN, François. The bilingual as a competent but specific speaker-hearer. **Journal of Multilingual and Multicultural Development**, [*S. l.*], 1985. Disponível em: <https://doi.org/10.1080/01434632.1985.9994221>

HARPE, Spencer E. **How to analyze Likert and other rating scale data**. [*S. l.*]: Elsevier Inc., 2015. Disponível em: <https://doi.org/10.1016/j.cptl.2015.08.001>

HELL, Janet G. van; ADAMS, Katharine Donnelly; ABDOLLAHI, Fatemeh. 18. Individual variation in bilingual lexical processing: the impact of second language proficiency and executive function on cross-language activation. **Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK)**, [*S. l.*], n. August, p. 210–222, 2019. Disponível em: <https://doi.org/10.1515/9783110435351-018>

JIANG, Nan; FORSTER, Kenneth I. Cross-Language Priming Asymmetries in Lexical Decision and Episodic Recognition. **Journal of Memory and Language**, [*S. l.*], 2001. Disponível em: <https://doi.org/10.1006/jmla.2000.2737>

KIM, Jeesun; DAVIS, Chris. Task effects in masked cross-script translation and phonological priming. **Journal of Memory and Language**, [*S. l.*], 2003. Disponível em: [https://doi.org/10.1016/S0749-596X\(03\)00093-7](https://doi.org/10.1016/S0749-596X(03)00093-7)

KROLL, Judith F.; STEWART, Erika. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. **Journal of Memory and Language**, [*S. l.*], 1994. Disponível em:

<https://doi.org/10.1006/jmla.1994.1008>

KROLL, Judith F.; TOKOWICZ, Natasha. The Development of Conceptual Representation for Words in a Second Language. *In: One mind, two languages: Bilingual language processing*. Kroll, Judith F.: Department of Psychology, Pennsylvania State University, 641 Moore Building, University Park, PA, US, 16802, isjfk7@psu.edu: Blackwell Publishing, 2001. (Explaining linguistics.).p. 49–71. *E-book*.

LAM, Kevin J. Y.; DIJKSTRA, Ton. Word repetition, masked orthographic priming, and language switching: Bilingual studies and BIA+ simulations. **International Journal of Bilingual Education and Bilingualism**, [S. l.], v. 13, n. 5, p. 487–503, 2010. Disponível em: <https://doi.org/10.1080/13670050.2010.488283>

MCCLELLAND, James L.; RUMELHART, David E. An interactive activation model of context effects in letter perception: I. An account of basic findings. **Psychological Review**, US, v. 88, n. 5, p. 375–407, 1981. Disponível em: <https://doi.org/10.1037/0033-295X.88.5.375>

MEYER, David E.; SCHVANEVELDT, Roger W. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. **Journal of Experimental Psychology**, [S. l.], 1971. Disponível em: <https://doi.org/10.1037/h0031564>

MEYER, David E.; SCHVANEVELDT, Roger W.; RUDDY, Margaret G. **Activation of Lexical Memory**. [S. l.: s. n.]

OLIVEIRA, Cândido Samuel Fonseca de; SÁ, Thaís Maíra Machado de. Métodos off-line em psicolinguística: julgamento de aceitabilidade. **Revele: Revista Virtual dos Estudantes de Letras**, [S. l.], v. 5, p. 77, 2013. Disponível em: <https://doi.org/10.17851/2317-4242.5.0.77-96>

PEIRCE, Jonathan *et al.* PsychoPy2: Experiments in behavior made easy. **Behavior Research Methods**, [S. l.], v. 51, n. 1, p. 195–203, 2019. Disponível em: <https://doi.org/10.3758/s13428-018-01193-y>

PEIRCE, Jonathan W. PsychoPy—Psychophysics software in Python. **Journal of Neuroscience Methods**, [S. l.], v. 162, n. 1–2, p. 8–13, 2007. Disponível em: <https://doi.org/10.1016/j.jneumeth.2006.11.017>

POST DA SILVEIRA, A. Retrieving L2 word stress from orthography: Evidence from word naming and cross-modal priming. **Ilha do Desterro A Journal of English Language, Literatures in English and Cultural Studies**, v. 73, n. 1, p. 409–442, 31 jan. 2020.

POST DA OLIVEIRA, Amanda; LEUSSEN, Jan-Willem. **GENERATING A BILINGUAL LEXICAL CORPUS USING INTERLANGUAGE NORMALIZED LEVENSHTAIN DISTANCES**. [S. l.: s. n.]. *E-book*.

POTTER, Mary *et al.* Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal ...* **Journal of Verbal Learning and Verbal Behaviour**, [S. l.], v. 38, n. February 1984, p. 23–38, 1984. Disponível em: [https://doi.org/10.1016/S0022-5371\(84\)90489-4](https://doi.org/10.1016/S0022-5371(84)90489-4)

PREUSS, Elena Ortiz; ARÊAS DA LUZ FONTES, Ana Beatriz; FINGER, Ingrid. Efeito cognato no processo de lexicalização bilíngue. **Letras de Hoje**, [S. l.], v. 50, n. 1, p. 73, 2015. Disponível em: <https://doi.org/10.15448/1984-7726.2015.1.18375>

R CORE TEAM. **R: A language and environment for statistical computing**. Vienna/Austria: [s. n.], 2014. *E-book*.

ROSENTHAL, Robert. **Meta-analytic procedures for social research, Rev. ed.** Thousand Oaks, CA, US: Sage Publications, Inc, 1991. (Applied social research methods series, Vol. 6.) *E-book*. Disponível em: <https://doi.org/10.4135/9781412984997>

SÁNCHEZ-CASAS, R.; GARCÍA-ALBEA, José. The representation of cognate and noncognate words in bilingual memory: Can cognate status be characterized as a special kind of morphological relation? **Handbook of Bilingualism: Psycholinguistic Approaches**, [S. l.], p. 226–250, 2005.

SCHEPENS, Job; DIJKSTRA, Ton; GROOTJEN, Franc. Distributions of cognates in Europe as based on Levenshtein distance. **Bilingualism**, [S. l.], v. 15, n. 1, p. 157–166, 2012 a. Disponível em: <https://doi.org/10.1017/S1366728910000623>

SCHEPENS, Job; DIJKSTRA, Ton; GROOTJEN, Franc. Distributions of cognates in Europe as based on Levenshtein distance. **Bilingualism**, [S. l.], v. 15, n. 1, p. 157–166, 2012 b. Disponível em: <https://doi.org/10.1017/S1366728910000623>

SCHOONBAERT, Sofie *et al.* Semantic and translation priming from a first language to a second and back: Making sense of the findings. **Memory and Cognition**, [S. l.], v. 37, n. 5, p. 569–586, 2009. Disponível em: <https://doi.org/10.3758/MC.37.5.569>

THORNDIKE, E. L.; LORGE, I. **The teacher's word book of 30,000 words**. Oxford, England: Bureau of Publications, Teachers Co, 1944. *E-book*.

TOASSI, Pâmela Freitas Pereira; MOTA, Mailce Borges; TEIXEIRA, Elisângela Nogueira. O efeito de palavras cognatas no acesso lexical do inglês como terceira língua. **Cadernos de Tradução**, [S. l.], v. 40, n. esp2, p. 74–96, 2020. Disponível em: <https://doi.org/10.5007/2175-7968.2020v40nesp2p74>

TOKOWICZ, Natasha. **Lexical Processing and Second Language Acquisition**. 1. ed. [S. l.]: Routledge, 2014. *E-book*.

VAN HEUVEN, Walter J. B. *et al.* SUBTLEX-UK: A new and improved word frequency database for British English. **Quarterly Journal of Experimental Psychology**, [S. l.], 2014. Disponível em: <https://doi.org/10.1080/17470218.2013.850521>

VAN HEUVEN, Walter J. B.; DIJKSTRA, Ton; GRAINGER, Jonathan. Orthographic Neighborhood Effects in Bilingual Word Recognition. **Journal of Memory and Language**, [S. l.], v. 39, n. 3, p. 458–483, 1998. Disponível em: <https://doi.org/10.1006/jmla.1998.2584>

VOGA, Madeleine; GRAINGER, Jonathan. Cognate status and cross-script translation priming. **Memory and Cognition**, [S. l.], 2007. Disponível em:

<https://doi.org/10.3758/BF03193467>

WANG, Xin. Language dominance in translation priming: Evidence from balanced and unbalanced Chinese-English bilinguals. **Quarterly Journal of Experimental Psychology**, [S. l.], v. 66, n. 4, p. 727–743, 2013. Disponível em: <https://doi.org/10.1080/17470218.2012.716072>

WEN, Yun; VAN HEUVEN, Walter J. B. Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. **Psychonomic Bulletin and Review**, [S. l.], v. 24, n. 3, p. 879–886, 2017. Disponível em: <https://doi.org/10.3758/s13423-016-1151-1>

WILLIAMS, John N. The relationship between word meanings in the first and second language: Evidence for a common, but restricted, semantic code. **European Journal of Cognitive Psychology**, [S. l.], v. 6, n. 2, p. 195–220, 1994. Disponível em: <https://doi.org/10.1080/09541449408520143>

## APÊNDICES

## APÊNDICE A

Estão descritos nas tabelas abaixo as listas utilizadas no experimento de *priming* encoberto realizado nesta dissertação. As condições são listadas conforme variáveis independentes estabelecidas no capítulo 4 desta dissertação: Tipo de palavra e relação de *priming*. Ainda, foi adicionado um número de identificação (id) para cada uma das condições experimentais e suas respectivas palavras *prime* e alvo para todas as listas.

### EXPERIMENTO 1A, LISTA 1

id	Palavra	Relação	prime	Alvo
1	co	re	value	valor
2	co	nr	author	vitória
3	co	re	production	produção
4	co	nr	model	público
5	co	re	night	noite
6	co	nr	bank	abril
7	co	re	region	região
8	co	nr	election	papel
9	co	re	problem	problema
10	co	nr	industry	dezembro
11	co	re	film	filme
12	co	nr	service	música
13	co	re	police	polícia
14	co	nr	control	título
15	co	re	decision	decisão
16	co	nr	success	outubro
17	co	re	sector	setor
18	co	nr	objective	qualidade
19	co	re	moment	momento
20	co	nr	victory	dólar
21	co	re	point	ponto
22	co	nr	article	população
23	co	re	period	período
24	co	nr	image	sociedade
25	co	re	car	carro
26	co	nr	salary	setembro
27	co	re	economy	economia
28	co	nr	december	pessoa
29	co	re	congress	congresso
30	co	nr	creation	senador
31	co	re	fault	falta
32	co	nr	culture	eleição
33	co	re	cause	causa
34	co	nr	public	autor
35	co	re	situation	situação

36	co	nr	april	banco
37	co	re	family	família
38	co	nr	person	modelo
39	co	re	justice	justiça
40	co	nr	leader	indústria
41	co	re	june	junho
42	co	nr	october	imagem
43	co	re	direction	direção
44	co	nr	title	controle
45	co	re	question	questão
46	co	nr	september	serviço
47	co	re	january	janeiro
48	co	nr	music	líder
49	co	re	defense	defesa
50	co	nr	dollar	cultura
51	co	re	center	centro
52	co	nr	population	objetivo
53	co	re	august	agosto
54	co	nr	society	sucesso
55	co	re	movement	movimento
56	co	nr	quality	salário
57	co	re	security	segurança
58	co	nr	senator	artigo
59	co	re	text	texto
60	co	nr	paper	criação
61	nc	re	deadline	prazo
62	nc	nr	lawyer	sábado
63	nc	re	start	início
64	nc	nr	bus	proposta
65	nc	re	increase	aumento
66	nc	nr	farm	século
67	nc	re	businessman	empresário
68	nc	nr	review	maneira
69	nc	re	book	livro
70	nc	nr	travel	corpo
71	nc	re	majority	maioria
72	nc	nr	teacher	manhã
73	nc	re	woman	mulher
74	nc	nr	purchase	tarde
75	nc	re	research	pesquisa
76	nc	nr	change	filho
77	nc	re	thing	coisa
78	nc	nr	leaf	força
79	nc	re	soccer	futebol
80	nc	nr	answer	renda

81	nc	re	man	homem
82	nc	nr	exit	terra
83	nc	re	support	apoio
84	nc	nr	recipe	entrevista
85	nc	re	sell	venda
86	nc	nr	path	prefeito
87	nc	re	player	jogador
88	nc	nr	magazine	cabeça
89	nc	re	fall	queda
90	nc	nr	army	consumo
91	nc	re	Sunday	domingo
92	nc	nr	century	advogado
93	nc	re	return	volta
94	nc	nr	manner	ônibus
95	nc	re	departure	partida
96	nc	nr	Saturday	fazenda
97	nc	re	measure	medida
98	nc	nr	proposal	revisão
99	nc	re	sense	sentido
100	nc	nr	body	revista
101	nc	re	people	gente
102	nc	nr	strength	professor
103	nc	re	truth	verdade
104	nc	nr	afternoon	mudança
105	nc	re	bill	conta
106	nc	nr	interview	compra
107	nc	re	field	campo
108	nc	nr	morning	folha
109	nc	re	health	saúde
110	nc	nr	earth	resposta
111	nc	re	brand	marca
112	nc	nr	income	saída
113	nc	re	right	direito
114	nc	nr	son	receita
115	nc	re	meeting	reunião
116	nc	nr	head	caminho
117	nc	re	war	guerra
118	nc	nr	mayor	viagem
119	nc	re	currency	moeda
120	nc	nr	consumption	exército
121	ps	re	crumedue	crumedue
122	ps	nr	rutiambo	varbaspe
123	ps	re	tidonfol	tidonfol
124	ps	nr	stapeixo	rutiambo
125	ps	re	obarlavo	obarlavo

126	ps	nr	cadedus	stapeixo
127	ps	re	hialhes	hialhes
128	ps	nr	tiparta	cadedus
129	ps	re	ubessus	ubessus
130	ps	nr	osuriva	tiparta
131	ps	re	mevisia	mevisia
132	ps	nr	montagu	osuriva
133	ps	re	uedação	uedação
134	ps	nr	sorrodi	montagu
135	ps	re	polunta	polunta
136	ps	nr	aismier	sorrodi
137	ps	re	togadai	togadai
138	ps	nr	epefase	aismier
139	ps	re	gasmili	gasmili
140	ps	nr	cefifre	epefase
141	ps	re	culerdu	culerdu
142	ps	nr	vardodo	cefifre
143	ps	re	cobodie	cobodie
144	ps	nr	clafiel	vardodo
145	ps	re	piopiru	piopiru
146	ps	nr	elizaie	clafiel
147	ps	re	veapeas	veapeas
148	ps	nr	iposurm	elizaie
149	ps	re	hiveafa	hiveafa
150	ps	nr	figibam	iposurm
151	ps	re	becizal	becizal
152	ps	nr	patuaço	figibam
153	ps	re	chilaro	chilaro
154	ps	nr	videvom	patuaço
155	ps	re	averton	averton
156	ps	nr	nunhese	videvom
157	ps	re	aultuse	aultuse
158	ps	nr	hucafim	nunhese
159	ps	re	fundure	fundure
160	ps	nr	sechas	hucafim
161	ps	re	impras	impras
162	ps	nr	icação	sechas
163	ps	re	hichos	hichos
164	ps	nr	omibis	icação
165	ps	re	tirora	tirora
166	ps	nr	enifis	omibis
167	ps	re	mestie	mestie
168	ps	nr	adolos	enifis
169	ps	re	neidus	neidus
170	ps	nr	ifiana	adolos

171	ps	re	chente	chente
172	ps	nr	uionso	ifiana
173	ps	re	ameipo	ameipo
174	ps	nr	gubiro	uionso
175	ps	re	ininir	ininir
176	ps	nr	itasem	gubiro
177	ps	re	olhere	olhere
178	ps	nr	obrobo	itasem
179	ps	re	sentur	sentur
180	ps	nr	miafas	obrobo
181	ps	re	anabal	anabal
182	ps	nr	iecode	miafas
183	ps	re	colopo	colopo
184	ps	nr	pepaco	iecode
185	ps	re	didati	didati
186	ps	nr	asmova	pepaco
187	ps	re	babivo	babivo
188	ps	nr	siscem	asmova
189	ps	re	redabi	redabi
190	ps	nr	elempe	siscem
191	ps	re	izanie	izanie
192	ps	nr	chatol	elempe
193	ps	re	caizau	caizau
194	ps	nr	douema	chatol
195	ps	re	aulatu	aulatu
196	ps	nr	cuevim	douema
197	ps	re	feclar	feclar
198	ps	nr	vegiro	cuevim
199	ps	re	lercue	lercue
200	ps	nr	ênizar	vegiro
201	ps	re	mibece	mibece
202	ps	nr	vedegu	ênizar
203	ps	re	henego	henego
204	ps	nr	bigaci	vedegu
205	ps	re	fedeva	fedeva
206	ps	nr	ruious	bigaci
207	ps	re	epovim	epovim
208	ps	nr	tusmom	ruious
209	ps	re	iemavo	iemavo
210	ps	nr	frulur	tusmom
211	ps	re	obiodr	obiodr
212	ps	nr	vaupru	frulur
213	ps	re	lhotua	lhotua
214	ps	nr	hiemus	vaupru
215	ps	re	fiefom	fiefom

216	ps	nr	burnos	hiemus
217	ps	re	maisão	maisão
218	ps	nr	potosa	burnos
219	ps	re	sofite	sofite
220	ps	nr	ecorda	potosa
221	ps	re	eviada	eviada
222	ps	nr	tefece	ecorda
223	ps	re	hirofe	hirofe
224	ps	nr	japors	tefece
225	ps	re	escupa	escupa
226	ps	nr	nambue	japors
227	ps	re	lagrol	lagrol
228	ps	nr	glospe	nambue
229	ps	re	stecus	stecus
230	ps	nr	breia	glospe
231	ps	re	abeso	abeso
232	ps	nr	rapea	breia
233	ps	re	açano	açano
234	ps	nr	basua	rapea
235	ps	re	omica	omica
236	ps	nr	voror	basua
237	ps	re	onopo	onopo
238	ps	nr	elaba	voror
239	ps	re	vasme	vasme
240	ps	nr	varbaspe	elaba

### EXPERIMENTO 1A, LISTA B

id	cognate	relation	prime	target
241	co	nr	text	valor
242	co	re	victory	vitória
243	co	nr	region	produção
244	co	re	public	público
245	co	nr	congress	noite
246	co	re	april	abril
247	co	nr	problem	região
248	co	re	paper	papel
249	co	nr	night	problema
250	co	re	december	dezembro
251	co	nr	sector	filme
252	co	re	music	música
253	co	nr	fault	polícia
254	co	re	title	título
255	co	nr	film	decisão
256	co	re	october	outubro
257	co	nr	decision	setor

258	co	re	quality	qualidade
259	co	nr	point	momento
260	co	re	dollar	dólar
261	co	nr	moment	ponto
262	co	re	population	população
263	co	nr	car	período
264	co	re	society	sociedade
265	co	nr	period	carro
266	co	re	september	setembro
267	co	nr	family	economia
268	co	re	person	pessoa
269	co	nr	june	congresso
270	co	re	senator	senador
271	co	nr	police	falta
272	co	re	election	eleição
273	co	nr	situation	causa
274	co	re	author	autor
275	co	nr	cause	situação
276	co	re	bank	banco
277	co	nr	center	família
278	co	re	model	modelo
279	co	nr	production	justiça
280	co	re	industry	indústria
281	co	nr	economy	junho
282	co	re	image	imagem
283	co	nr	defense	direção
284	co	re	control	controle
285	co	nr	justice	questão
286	co	re	service	serviço
287	co	nr	direction	janeiro
288	co	re	leader	líder
289	co	nr	january	defesa
290	co	re	culture	cultura
291	co	nr	question	centro
292	co	re	objective	objetivo
293	co	nr	security	agosto
294	co	re	success	sucesso
295	co	nr	august	movimento
296	co	re	salary	salário
297	co	nr	movement	segurança
298	co	re	article	artigo
299	co	nr	value	texto
300	co	re	creation	criação
301	nc	nr	sense	prazo
302	nc	re	Saturday	sábado

303	nc	nr	heath	início
304	nc	re	proposal	proposta
305	nc	nr	departure	aumento
306	nc	re	century	século
307	nc	nr	measure	empresário
308	nc	re	manner	maneira
309	nc	nr	return	livro
310	nc	re	body	corpo
311	nc	nr	people	maioria
312	nc	re	morning	manhã
313	nc	nr	field	mulher
314	nc	re	afternoon	tarde
315	nc	nr	currency	pesquisa
316	nc	re	son	filho
317	nc	nr	truth	coisa
318	nc	re	strength	força
319	nc	nr	brand	futebol
320	nc	re	income	renda
321	nc	nr	Sunday	homem
322	nc	re	earth	terra
323	nc	nr	meeting	apoio
324	nc	re	interview	entrevista
325	nc	nr	right	venda
326	nc	re	mayor	prefeito
327	nc	nr	war	jogador
328	nc	re	head	cabeça
329	nc	nr	bill	queda
330	nc	re	consumption	consumo
331	nc	nr	increase	domingo
332	nc	re	lawyer	advogado
333	nc	nr	deadline	volta
334	nc	re	bus	ônibus
335	nc	nr	businessman	partida
336	nc	re	farm	fazenda
337	nc	nr	player	medida
338	nc	re	review	revisão
339	nc	nr	book	sentido
340	nc	re	magazine	revista
341	nc	nr	majority	gente
342	nc	re	teacher	professor
343	nc	nr	woman	verdade
344	nc	re	change	mudança
345	nc	nr	research	conta
346	nc	re	purchase	compra
347	nc	nr	thing	campo

348	nc	re	leaf	folha
349	nc	nr	start	saúde
350	nc	re	answer	resposta
351	nc	nr	man	marca
352	nc	re	exit	saída
353	nc	nr	support	direito
354	nc	re	recipe	receita
355	nc	nr	fall	reunião
356	nc	re	path	caminho
357	nc	nr	sell	guerra
358	nc	re	travel	viagem
359	nc	nr	soccer	moeda
360	nc	re	army	exército
361	ps	nr	vasme	crumedue
362	ps	re	varbaspe	varbaspe
363	ps	nr	crumedue	tidonfol
364	ps	re	rutiambo	rutiambo
365	ps	nr	tidonfol	obarlavo
366	ps	re	stapeixo	stapeixo
367	ps	nr	obarlavo	hialhes
368	ps	re	cadodus	cadodus
369	ps	nr	hialhes	ubessus
370	ps	re	tiparta	tiparta
371	ps	nr	ubessus	mevisia
372	ps	re	osuriva	osuriva
373	ps	nr	mevisia	uedação
374	ps	re	montagu	montagu
375	ps	nr	uedação	polunta
376	ps	re	sorrodi	sorrodi
377	ps	nr	polunta	togadai
378	ps	re	aismier	aismier
379	ps	nr	togadai	gasmili
380	ps	re	epefase	epefase
381	ps	nr	gasmili	culerdu
382	ps	re	cefifre	cefifre
383	ps	nr	culerdu	cobodie
384	ps	re	vardodo	vardodo
385	ps	nr	cobodie	piopiru
386	ps	re	clafiel	clafiel
387	ps	nr	piopiru	veapeas
388	ps	re	elizaie	elizaie
389	ps	nr	veapeas	hiveafa
390	ps	re	iposurm	iposurm
391	ps	nr	hiveafa	becizal
392	ps	re	figibam	figibam

393	ps	nr	becizal	chilaro
394	ps	re	patuaço	patuaço
395	ps	nr	chilaro	averton
396	ps	re	videvom	videvom
397	ps	nr	averton	aultuse
398	ps	re	nunhese	nunhese
399	ps	nr	aultuse	fundure
400	ps	re	hucafim	hucafim
401	ps	nr	fundure	impras
402	ps	re	sechas	sechas
403	ps	nr	impras	hichos
404	ps	re	icação	icação
405	ps	nr	hichos	tirora
406	ps	re	omibis	omibis
407	ps	nr	tirora	mestie
408	ps	re	enifis	enifis
409	ps	nr	mestie	neidus
410	ps	re	adolos	adolos
411	ps	nr	neidus	chente
412	ps	re	ifiana	ifiana
413	ps	nr	chente	ameipo
414	ps	re	uionso	uionso
415	ps	nr	ameipo	ininir
416	ps	re	gubiro	gubiro
417	ps	nr	ininir	olhere
418	ps	re	itasem	itasem
419	ps	nr	olhere	sentur
420	ps	re	obrobo	obrobo
421	ps	nr	sentur	anabal
422	ps	re	miafas	miafas
423	ps	nr	anabal	colopo
424	ps	re	iecode	iecode
425	ps	nr	colopo	didati
426	ps	re	pepaco	pepaco
427	ps	nr	didati	babivo
428	ps	re	asmova	asmova
429	ps	nr	babivo	redabi
430	ps	re	siscem	siscem
431	ps	nr	rebadi	izanie
432	ps	re	elempe	elempe
433	ps	nr	izanie	caizau
434	ps	re	chatol	chatol
435	ps	nr	caizau	aulatu
436	ps	re	douema	douema
437	ps	nr	aulatu	feclar

438	ps	re	cuevim	cuevim
439	ps	nr	feclar	lercue
440	ps	re	vegiro	vegiro
441	ps	nr	lercue	mibece
442	ps	re	ênizar	ênizar
443	ps	nr	mibece	henego
444	ps	re	vedegu	vedegu
445	ps	nr	henego	fedeva
446	ps	re	bigaci	bigaci
447	ps	nr	fedeva	epovim
448	ps	re	ruious	ruious
449	ps	nr	epovim	iemavo
450	ps	re	tusmom	tusmom
451	ps	nr	iemavo	obiodr
452	ps	re	frulur	frulur
453	ps	nr	obiodr	lhotua
454	ps	re	vaupru	vaupru
455	ps	nr	ilhotua	fiefom
456	ps	re	hiemus	hiemus
457	ps	nr	fiefom	maisão
458	ps	re	burnos	burnos
459	ps	nr	maisão	sofite
460	ps	re	potosa	potosa
461	ps	nr	sofite	eviada
462	ps	re	ecorda	ecorda
463	ps	nr	eviada	hirofe
464	ps	re	tefece	tefece
465	ps	nr	hirofe	escupa
466	ps	re	japors	japors
467	ps	nr	escupa	lagrol
468	ps	re	nambue	nambue
469	ps	nr	lagrol	stecus
470	ps	re	glospe	glospe
471	ps	nr	stecus	abeso
472	ps	re	breia	breia
473	ps	nr	abeso	açano
474	ps	re	rapea	rapea
475	ps	nr	açano	omica
476	ps	re	basua	basua
477	ps	nr	omica	onopo
478	ps	re	voror	voror
479	ps	nr	onopo	vasme
480	ps	re	elaba	elaba

## APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

### TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

#### Dados de identificação

**Título do Projeto:** Organização lexical em bilíngues: uma investigação sobre os efeitos na direcionalidade do priming encoberto de palavras cognatas e não-cognatas.

**Pesquisador Responsável:** Maylton da Silva Fernandes

**Instituição a que pertence o Pesquisa dor Responsável:** Universidade Federal da Paraíba (UFPB)

**Endereço do pesquisador:** Bacharel Wilson Flávio Moreira Coutinho, 76, Jardim Cidade Universitária, João Pessoa/PB, CEP 58052-5 10.

**Telefone para contato:** (83) 99339-2410

**E-mail:** maylton.fernandes@gmail.com

**Endereço do Comitê de Ética:** Universitário, S/N, Castelo Branco, João Pessoa/PB, CEP 5805 1900.

**Telefone para contato:** (83) 3216-7791

**E-mail:** comitedeetica@ccs.ufpb.br

O(A) senhor(a) está sendo(a) convidado(a) para participar da pesquisa “Organização lexical em bilíngues: uma investigação sobre os efeitos na direcionalidade do priming encoberto de palavras cognatas e não-cognatas”. Este estudo está sendo conduzido por Maylton Fernandes, aluno de Mestrado da Universidade Federal da Paraíba.

O objetivo deste estudo é investigar como o conhecimento de uma língua influencia no conhecimento linguístico de outras línguas.

Esta pesquisa está sendo conduzida com falantes nativos de Português e/ou de Inglês, que estejam cursando ou que já concluíram o ensino superior ou posterior.

Se você concordar em participar deste estudo, você será solicitado para completar um questionário online, bem como participar de uma tarefa de decisão lexical. Este questionário versará sobre o seu conhecimento linguístico prévio e sobre o seu uso atual do Português e do Inglês. A tarefa de decisão lexical será sobre o conhecimento lexical do participante em inglês e português, dividido em duas etapas, e seu tempo de leitura será gravado. As duas tarefas levarão aproximadamente 30 minutos (total) para completarem.

Nós acreditamos que os riscos decorrentes desta investigação são mínimos; no entanto, sempre há um risco de quebra de confidencialidade e, portanto, tomamos medidas para minimizar tais riscos.

Dentro do possível, suas respostas neste estudo serão confidenciais. Nós minimizaremos as possibilidades de vazamento de seus dados através de uma atribuição de código somente conhecido pelo pesquisador, e que não identificam pessoalmente sua pessoa.

Ao participar deste estudo, você não receberá pagamento.

Você não precisa participar deste estudo, caso não queira. Se você concordar em participar dele, mas posteriormente mudar de ideia, você poderá sair da amostra a qualquer momento. Não há penalidades ou consequências de quaisquer tipos se você decidir que não deseja mais participar.

Caso você tenha alguma pergunta sobre o projeto ou se você tiver algum problema derivado desta pesquisa, você poderá contactar o pesquisador Maylton Fernandes através do e-mail ou telefone dispostos no início deste termo.

Obs.: O sujeito da pesquisa ou seu representante responsável deverão rubricar todas as folhas do TCLE apondo suas assinaturas na última página do referido Termo. Assim, ao assinar abaixo, o(a) senhor(a) indica que tem pelo menos 18 anos, que leu este termo e concorda em participar da pesquisa.

De antemão, agradecemos sua contribuição.

\_\_\_\_\_, \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_

\_\_\_\_\_  
Assinatura do participante

\_\_\_\_\_  
Assinatura do pesquisador responsável

## APÊNDICE C – QUESTIONÁRIO SÓCIOLINGUÍSTICO NO GOOGLE FORMS

21/08/2021

Histórico Linguístico

### Histórico Linguístico

Você está sendo convidado(a) para participar de uma pesquisa. Este estudo está sendo conduzido por Maylton Fernandes, aluno de Mestrado da Universidade Federal da Paraíba.

O objetivo deste estudo é investigar como o conhecimento de uma língua influencia no conhecimento linguístico de outras línguas.

Esta pesquisa está sendo conduzida com falantes nativos de Português e/ou de Inglês, que estejam cursando ou que já concluíram o ensino superior ou posterior.

Se você concordar em participar deste estudo, você será solicitado para completar um questionário online. Este questionário versará sobre o seu conhecimento linguístico prévio e sobre o seu uso atual do Português e do Inglês.

Caso você tenha algum pergunta sobre o projeto ou se você tiver algum problema derivado desta pesquisa, você poderá contactar o pesquisador Maylton Fernandes através do e-mail [maylton.fernandes@gmail.com](mailto:maylton.fernandes@gmail.com).

De antemão, agradecemos sua ajuda.

**\*Obrigatório**

1. E-mail \*

---

2. Nome do participante \*

---

3. Número para contato

---

Histórico linguístico

4. Qual sua data de nascimento? \*

---

*Exemplo: 7 de janeiro de 2019*

21/08/2021

Histórico Linguístico

5. Onde você nasceu? (País, estado e cidade) \*

---

6. Gênero \*

*Marcar apenas uma oval.*

- Feminino
- Masculino
- Prefiro não responder
- Outro: \_\_\_\_\_

7. Qual o nível mais alto de educação que você completou? \*

*Marcar apenas uma oval.*

- Ensino Fundamental
- Ensino Médio
- Graduação
- Especialização
- Mestrado
- Doutorado

8. Além do Brasil, você já viveu em outro país por um longo período de tempo (mais de 6 meses)? Caso afirmativo, onde e por quanto tempo? \*

*Marcar apenas uma oval.*

- Não
- Outro: \_\_\_\_\_

Histórico linguístico

21/08/2021

Histórico Linguístico

9. Qual(is) língua(s) você adquiriu antes de começar a estudar? \*

---

10. Quantos anos você tinha quando começou a adquiri-la(s)?

---

11. Em qual(is) língua(s) você estudou a maior parte de sua educação antes do ensino superior? \*

---

12. Qual(is) língua(s) você aprendeu em um ambiente educacional formal? \*

---

13. Qual(s) língua(s) você aprendeu fora de um ambiente educacional formal? \*

---



---



---



---



---

14. Em geral, como você classificaria sua proficiência linguística em Inglês? \*

*Marcar apenas uma oval por linha.*

	Nenhuma	Péssima	Ruim	Suficiente	Boa	Excelente
Fala (Produção oral)	<input type="radio"/>					
Escuta (Compreensão oral)	<input type="radio"/>					
Proficiência em geral	<input type="radio"/>					

21/08/2021

Histórico Linguístico

15. Em geral, como você classificaria sua proficiência linguística em Português? \*

*Marcar apenas uma oval por linha.*

	Nenhuma	Péssima	Ruim	Suficiente	Boa	Excelente
Fala (Produção oral)	<input type="radio"/>					
Escuta (Produção oral)	<input type="radio"/>					
Proficiência em geral	<input type="radio"/>					

16. Com que frequência você fala Inglês? \*

*Marcar apenas uma oval.*

- Nunca
- Raramente
- Poucas vezes ao ano
- Mensalmente
- Semanalmente
- Diariamente

17. Você se sente mais confortável falando em Inglês ou Português? \*

*Marcar apenas uma oval.*

- Inglês
- Português
- Sem preferência

21/08/2021

Histórico Linguístico

18. Você poderia elaborar mais sua resposta sobre a pergunta anterior? Por que você se sente mais confortável com o Inglês/Português ou por que você não tem nenhuma preferência?

---

---

---

---

---

#### Histórico linguístico (pais)

19. Qual(is) língua(s) seu pai fala? \*

---

20. Qual(is) língua(s) você mais usa quando fala com seu pai? \*

---

21. Qual(is) língua(s) seu pai mais usa quando fala com você? \*

---

22. Qual(is) língua(s) sua mãe fala? \*

---

23. Qual(is) língua(s) você mais usa quando fala com sua mãe? \*

---

24. Qual(is) língua(s) sua mãe usa quando fala com você? \*

---

#### Uso linguístico

21/08/2021

Histórico Linguístico

26. Você poderia indicar na grade abaixo a frequência que usa Inglês nos contextos marcados? \*

*Marcar apenas uma oval por linha.*

	Nunca	Muito raramente	Raramente	Às vezes	Frequentemente	Todo o tempo
Com parentes	<input type="radio"/>					
Com amigos	<input type="radio"/>					
Com animais de estimação	<input type="radio"/>					
No trabalho	<input type="radio"/>					
Na escola	<input type="radio"/>					
Na igreja	<input type="radio"/>					
Em lojas/supermercados	<input type="radio"/>					
Em clubes/organizações	<input type="radio"/>					

27. Você escuta músicas em Inglês? \*

*Marcar apenas uma oval.*

- Sim  
 Não

28. Você assiste vídeos ou filmes em Inglês? \*

*Marcar apenas uma oval.*

- Sim  
 Não

21/08/2021

Histórico Linguístico

29. Você escuta programas de rádio ou podcasts em Inglês? \*

*Marcar apenas uma oval.*

Sim

Não

30. Você lê livros, jornais, revistas ou blogs em Inglês? \*

*Marcar apenas uma oval.*

Sim

Não

31. Você se sente desconfortável falando Inglês com um falante nativo de Inglês? \*

*Marcar apenas uma oval.*

Sim, sempre.

Sim, as vezes.

Não, nunca.

Final do questionário

32. Você chegou ao fim do questionário. Há algo que você gostaria de mencionar?  
Este comentário pode ser sobre qualquer assunto relacionado à pesquisa.

---

---

---

---

---

Este conteúdo não foi criado nem aprovado pelo Google.

Google Formulários