



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS DA TRADUÇÃO

IAKOB LOURENÇO MOTA

**O PROCESSAMENTO DE HOMÓGRAFOS INTERLINGUÍSTICOS DO INGLÊS
PARA O PORTUGUÊS BRASILEIRO DURANTE A TRADUÇÃO**

FORTALEZA
2024

IAKOB LOURENÇO MOTA

O PROCESSAMENTO DE HOMÓGRAFOS INTERLINGUÍSTICOS DO INGLÊS PARA O
PORTUGUÊS BRASILEIRO DURANTE A TRADUÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Estudos da Tradução (POET) da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Estudos da Tradução. Área de concentração: linguagem, cognição e recursos tecnológicos.

Orientador: Prof^a. Dr^a. Pâmela Freitas Pereira Toassi.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

M871p Mota, Iakob Lourenço.
O processamento de homógrafos interlinguísticos do inglês para o português brasileiro durante a tradução/ Iakob Lourenço Mota. – 2024.
124 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Humanidades,
Programa de Pós-Graduação em Estudos da Tradução, Fortaleza, 2024.
Orientação: Profa. Dra. Pâmela Freitas Pereira Toassi.

1. estudos da tradução. 2. psicolinguística. 3. cognição. 4. bilinguismo. 5. homógrafos interlinguísticos. I. Título.

CDD 418.02

IAKOB LOURENÇO MOTA

O PROCESSAMENTO DE HOMÓGRAFOS INTERLINGUÍSTICOS DO INGLÊS PARA O
PORTUGUÊS BRASILEIRO DURANTE A TRADUÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Estudos da Tradução (POET) da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Estudos da Tradução. Área de concentração: linguagem, cognição e recursos tecnológicos.

Aprovada em: 26/02/2024.

BANCA EXAMINADORA

Profª. Dra. Pâmela Freitas Pereira Toassi (Orientadora)
Universidade Federal do Ceará (UFC)

Profª. Dra. Diana Costa Fortier Silva
Universidade Federal do Ceará (UFC)

Prof. Dr. Rafael Ferreira da Silva
Universidade Federal do Ceará (UFC)

Profª. Dra. Mailce Borges mota
Universidade Federal de Santa Catarina (UFSC)

AGRADECIMENTOS

Gostaria de agradecer à minha família pelo apoio incondicional nos estudos. Mesmo eu trabalhando na docência, uma área que ainda falta o devido prestígio, eles nunca deixaram de me apoiar em minhas escolhas.

Não posso deixar de agradecer também à minha orientadora Profa. Dra. Pâmela Freitas Pereira Toassi que me ajudou a concluir mais essa etapa da minha vida acadêmica. Agradeço também aos membros da banca examinadora, professores doutores Diana Costa Fortier Silva, Rafael Ferreira da Silva e Mailce Borges Mota que participaram ativamente com indagações e observações deveras frutíferas para a construção dessa dissertação.

Gostaria de agradecer também aos meus colegas do Laboratório de Fonética e Multilinguismo (LabFoM) que, seja por meio de conversas despreziosas ou em reuniões mais formais, contribuíram também para a construção final do texto.

RESUMO

O presente estudo explora os conceitos interligados de tradução, bilinguismo e métodos de acesso lexical bilíngue, esclarecendo sua evolução histórica e seu significado no mundo multilíngue de hoje. Começando com um vislumbre das origens da tradução, mergulhamos na rica história dessa prática fundamental, traçando seu desenvolvimento desde os tempos antigos até a era moderna. De textos religiosos a obras-primas literárias, a tradução desempenhou um papel vital na ponte entre as culturas e na facilitação da comunicação global. Adiante, a dissertação se aprofunda no âmbito do bilinguismo, destacando seu impacto na aquisição da linguagem e nos processos cognitivos. Ao estudar o contexto histórico e as pesquisas em andamento, obtemos informações valiosas sobre as complexidades do bilinguismo na sociedade cada vez mais globalizada e plural. Ademais, o estudo investiga o campo dos métodos de acesso lexical bilíngue, explorando como os indivíduos bilíngues acessam e processam seu conhecimento linguístico. Discutimos várias estruturas teóricas e modelos computacionais que foram desenvolvidos para entender as complexidades do processamento lexical bilíngue (Kroll; Stewart, 1994; Kroll; De Groot, 1997; Dijkstra; Van Heuven; Grainger, 1998; Dijkstra; Van Heuven, 2002; Dijkstra *et al.*, 2018). Esses métodos abrangem diferentes mecanismos cognitivos envolvidos na recuperação lexical e na seleção da linguagem, esclarecendo a natureza dinâmica do processamento do sujeito bilíngue. Este estudo tem como objetivo analisar o tempo de reação e a acurácia na tradução de homógrafos interlinguísticos do par linguístico inglês – português brasileiro com o propósito de verificar quais efeitos essa classe de palavras tem na tradução. Para fazer a coleta de dados foi utilizado o *software* Psytoolkit (Stoet, 2010, 2017) que opera de maneira *online* captando o tempo de reação das respostas. O corpus do estudo foi alimentado no *software* e, após a coleta de dados, a análise foi feita no RStudio, um programa dedicado ao processamento e análise de dados. Após a análise, foi possível observar um efeito de interferência dos homógrafos interlinguísticos para a acurácia e tempo de resposta. Os resultados obtidos acrescentam à discussão acerca da utilização da tradução como instrumento de pesquisa sobre o acesso lexical.

Palavras-chave: estudos da tradução; psicolinguística; cognição; bilinguismo; homógrafos interlinguísticos.

ABSTRACT

This study explores the intertwined concepts of translation, bilingualism and bilingual lexical access methods, shedding light on their historical evolution and their significance in today's multilingual world. Starting with a glimpse into the origins of translation, we delve into the rich history of this fundamental practice, tracing its development from ancient times to the modern era. From religious texts to literary masterpieces, translation has played a vital role in bridging cultures and facilitating global communication. Further on, the dissertation deepens into the scope of bilingualism, highlighting its impact on language acquisition and cognitive processes. By studying the historical context and ongoing research, we gain valuable insights into the complexities of bilingualism in an increasingly globalized and pluralistic society. Furthermore, the study investigates the field of bilingual lexical access methods, exploring how bilingual individuals access and process their linguistic knowledge. We discuss several theoretical frameworks and computational models that have been developed to understand the complexities of bilingual lexical processing (Kroll; Stewart, 1994; Kroll; De Groot, 1997; Dijkstra; Van Heuven; Grainger, 1998; Dijkstra; Van Heuven, 2002; Dijkstra *et al.*, 2018). These methods encompass different cognitive mechanisms involved in lexical retrieval and language selection, clarifying the dynamic nature of the bilingual subject's processing. This study aims to analyze the reaction time and accuracy in the translation of interlingual homographs from the English – Brazilian Portuguese language pair with the purpose of verifying what effects this class of words has on the translation. To collect the data, Psytoolkit software (Stoet, 2010, 2017) was used, which operates online, capturing the reaction time of responses. The study corpus was fed into the software and, after data collection, the analysis was carried out in RStudio, an application dedicated to data processing and analysis. After the analysis, it was possible to observe an interference effect of interlingual homographs for accuracy and response time. The results obtained add to the discussion about the use of translation as a research instrument on lexical access.

Keywords: translation studies; psycholinguistics; cognition; bilingualism; interlingual homographs

LISTA DE FIGURAS

Figura 1	– Pedra de Roseta	21
Figura 2	– Operação do computador IBM 701	24
Figura 3	– Modelo do Processo Tradutório	32
Figura 4	– (V)Wandinha Addams em 4 fases: Cartum (1938), Lisa Loring (1964), Christina Ricci (1991) e Jenna Ortega (2022)	36
Figura 5	– Modelo de Associação de Palavras	45
Figura 6	– Modelo de Mediação de Conceito	46
Figura 7	– Modelo Hierárquico Revisado	47
Figura 8	– Modelo de Traços Distribuídos	48
Figura 9	– Demonstração das sobreposições entre os lemas <i>elephant</i> e <i>elefante</i>	50
Figura 10	– Demonstração das sobreposições entre os lemas <i>sorrow</i> e <i>tristeza</i>	51
Figura 11	– Modelo de Ativação Interativa Bilíngue (BIA)	53
Figura 12	– Modelo de fonte ortográfica e traços de características visuais do BIA	55
Figura 13	– Modelo de Ativação Interativa Bilíngue Revisado (BIA+)	57
Figura 14	– Modelo Multilink	59
Figura 15	– Identificação do participante	82
Figura 16	– Nível estimado de língua inglesa nas quatro habilidades	83
Figura 17	– Forma(s) de aprendizado da língua inglesa	84
Figura 18	– Questionário de conhecimento de línguas estrangeiras	84
Figura 19	– Início da tarefa de tradução	86
Figura 20	– Identificação do participante	87
Figura 21	– Texto de introdução à tarefa de tradução (treinamento)	87
Figura 22	– Ponto de fixação	88
Figura 23	– Estímulo (treinamento)	89

Figura 24	– Estímulo (treinamento) com a caixa de texto para resposta	89
Figura 25	– Texto de início da tarefa de tradução	90
Figura 26	– Tela de aviso de metade da tarefa concluída	91
Figura 27	– Término da tarefa de tradução	91

LISTA DE GRÁFICOS

Gráfico 1 – Acurácia por condição	97
Gráfico 2 – Histograma do tempo de reação da tarefa de tradução	100
Gráfico 3 – Diagrama de nuvem de chuva do tempo de reação (TR) da tarefa de tradução por cada condição	101
Gráfico 4 – Histograma do tempo de reação da tarefa de tradução (todas as respostas)	107
Gráfico 5 – Diagrama de nuvem de chuva do tempo de reação (TR) da tarefa de tradução por cada condição	108

LISTA DE TABELAS

Tabela 1 – A escala Zipf de frequência de palavras	79
Tabela 2 – Porcentagem da acurácia por condição	98
Tabela 3 – Modelo generalizado de efeitos mistos (acurácia por condição)	99
Tabela 4 – Tempo de reação por condição da tarefa de tradução	102
Tabela 5 – Modelo linear de efeitos mistos (TR por condição)	103
Tabela 6 – Tempo de reação por condição da tarefa de tradução (geral)	108
Tabela 7 – Modelo linear de efeitos mistos (TR por condição - geral)	110

LISTA DE QUADROS

Quadro 1	– Comparativo entre a tradução automática do Google em 2012 (pré-neural) e 2022 (neural) do texto <i>Superstitie</i> de Marin Sorescu	25
Quadro 2	– Comparativo entre a tradução automática do Google em 2012 (pré-neural) e 2022 (neural) de parte do texto <i>Shakespeare</i> de Marin Sorescu	27
Quadro 3	– Resumo dos critérios de dimensões do bilinguismo	41
Quadro 4	– Homófonos e homógrafos do par inglês/português brasileiro	61
Quadro 5	– Características gerais dos Participantes	67
Quadro 6	– Nível autodeclarado e aprendizagem de línguas	68
Quadro 7	– Lista dos estímulos da tarefa de tradução (grupo experimental [EXP])	71
Quadro 8	– Lista dos estímulos da tarefa de tradução (grupo controle inglês [CI])	73
Quadro 9	– Lista dos estímulos da tarefa de tradução (grupo controle português [CP])	74
Quadro 10	– Lista dos homógrafos interlinguísticos do treinamento	75
Quadro 11	– Lista das palavras controle do treinamento	75
Quadro 12	– Representação da escala Zipf versus Frequência por milhão	79

SUMÁRIO

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	20
2.1	Breve histórico sobre a tradução no mundo	20
2.1.1	<i>Panorama da tradução automática</i>	23
2.2	A tradução como processo	30
2.2.1	<i>O tradutor como centro da tradução</i>	33
2.3	A psicolinguística e o processo de tradução bilíngue	40
2.3.1	<i>Caracterizando o bilinguismo</i>	41
2.4	Estado da arte dos modelos de acesso lexical bilíngue	44
2.4.1	<i>O Modelo Hierárquico Revisado</i>	45
2.4.2	<i>O Modelo de Traços Distribuídos</i>	48
2.4.3	<i>O Modelo de Ativação Interativa Bilíngue</i>	52
2.4.4	<i>O Modelo De Ativação Interativa Bilíngue Plus (BIA+)</i>	56
2.4.5	<i>O Modelo Multilink</i>	59
2.5	Homônimos e o processamento de línguas	61
3	METODOLOGIA	64
3.1	Objetivos	64
3.1.1	<i>Objetivo geral</i>	64
3.1.2	<i>Objetivos específicos</i>	64
3.2	Perguntas da pesquisa	64
3.3	Hipóteses	65
3.4	Procedimentos da pesquisa	65
3.4.1	<i>Desenho da pesquisa</i>	65
3.5	Participantes	67
3.6	Aparato tecnológico	70
3.7	Criação dos estímulos da tarefa de tradução	71
3.7.1	<i>A criação de um novo corpus e medição da frequência dos estímulos</i>	76
3.8	Procedimentos da pesquisa experimental	80
3.8.1	<i>Identificação</i>	81
3.8.2	<i>Questionário biográfico e linguístico</i>	82
3.8.3	<i>Questionário de tecnologia utilizada</i>	85

3.9	Tarefa de tradução.....	86
3.10	Teste de produção de vocabulário em língua inglesa.....	92
3.11	Tratamento dos dados.....	93
4	RESULTADOS	96
4.1	Análise estatística descritiva: Acurácia	97
4.1.2	<i>Análise estatística inferencial: Acurácia</i>	98
4.2	Análise estatística descritiva: Tempo de reação	100
4.2.1	<i>Análise estatística inferencial: Tempo de reação</i>	102
4.3	Discussão dos resultados	104
4.3.1	<i>A acurácia na tarefa de tradução</i>	105
4.3.2	<i>O tempo de reação na tarefa de tradução</i>	105
4.3.3	<i>Análise do tempo de reação com respostas incorretas</i>	106
5	CONSIDERAÇÕES FINAIS	112
	REFERÊNCIAS	115

1 INTRODUÇÃO

A escolha de uma área temática dentro do curso de Letras-Inglês para me aprofundar a nível de mestrado foi uma tarefa relativamente fácil, pois desde cedo fui um entusiasta em traduzir (de jogos digitais a letras de músicas) e foi essa paixão de criança que me levou a trilhar o caminho em que estou agora. Posso dizer também que a própria vivência acadêmica contribuiu para que meu caminho culminasse neste trabalho, já que a primeira vez que tive contato com o curso de letras, em meados de 2009, eu tive um fascínio pelo campo da psicolinguística pois aquilo era tudo novo e funcionava como uma motivação extra para eu continuar a aprimorar os meus conhecimentos. Por desventuras da vida não pude concluir o curso de Letras na primeira vez em que ingressei, porém, assim como no velho ditado, “há males que vem para o bem”, eu ingressei na Universidade Federal do Ceará (UFC) no ano de 2016 e pouco tempo depois consegui uma bolsa pela Pró-Reitoria de Relações Internacionais e Desenvolvimento Institucional (PROINTER) para fazer parte do projeto do Laboratório de Edição, Tradução e Revisão de Textos Acadêmicos (LETRARE). O LETRARE, sob a coordenação da professora Dr^a. Diana Fortier, atua como uma plataforma para a qual os docentes da UFC enviam artigos (em inglês) e os alunos-bolsistas fazem a revisão destes. Para além da parte prática do projeto, há também uma parte teórica em que os alunos-bolsistas têm contato com diversas teorias e teóricos da tradução, e foi nessa essencial parte em que eu pude assim me aperfeiçoar e me apaixonar ainda mais pelo campo da tradução.

Em se tratando da área da psicolinguística, eu tive meu primeiro contato ainda quando estudante da Universidade Estadual do Ceará (UECE), onde tive o prazer de ver a base desse campo de estudo e algumas pesquisas que tentavam explicar a aquisição da linguagem e a evolução da linguagem durante os períodos iniciais da vida humana. Ao ingressar na UFC, o meu escopo sobre a atuação da psicolinguística foi ampliado e comecei a ver estudos além da aquisição de linguagem, como também a compreensão da linguagem e a aquisição de segunda língua.

O conceito para esta dissertação veio a partir de um projeto guarda-chuva da professora Dr^a Pâmela Toassi que busca estender o nosso conhecimento empírico sobre tais processos cognitivos que acontecem durante a tradução, em especial a leitura. Por ser um tema abrangente, eu fiquei encarregado de analisar este aspecto nas palavras homógrafas do par inglês/português brasileiro que é apenas um minúsculo ponto no meio do oceano que é a compreensão da nossa capacidade de tradução. Como dito por Chesterman (2020, p. 25, tradução nossa) “Mas a gente pode descobrir algo, e este algo pode eventualmente ‘evoluir’ em

um conhecimento mais amplo, melhores explicações e melhor entendimento”¹. O objetivo deste estudo versa sobre um campo de pesquisa que está atrelado à tradução, mas não ao fato de “traduzir e ver se está correto”, e sim às etapas da cognição que podem facilitar ou dificultar o processo tradutório, sendo mais específico o desencadear da leitura e tradução de homógrafos interlinguísticos do Inglês para o Português. Além da motivação pessoal como aluno e entusiasta em pesquisas, é de grande valia pesquisas na área da psicolinguística da tradução, uma vez que o comprometimento deste campo de estudo foca no processo ou no “como” aquele texto foi traduzido e não na tradução como produto final, conseguindo assim elaborar uma ponte de conhecimento entre produto e processo de tradução.

Aproveito o ensejo, o tom mais leve desta introdução e a temática da curiosidade para também comentar sobre um aspecto que me chamou a atenção no início da vida acadêmica e continua chamando atenção até hoje que são os nomes dos autores que influenciam nas pesquisas: o psicólogo Jacob Kantor, que fora a primeira pessoa a usar o termo “psicolinguística” em 1936², há também Roman Jakobson, um pioneiro e importantíssimo linguista na área da análise estrutural da linguagem, e para complementar há outro pesquisador, Arnt Lykke Jakobsen, que juntamente com Fabio Alves contribuem de maneira intensa para esta mudança na visão epistemológica dos estudos da tradução. Apesar das variações nos nomes, Jacob, Jakob(son) e Jakob(sen) todos derivam do nome hebraico Ya’akov que significa “calcanhar” ou “aquele que usurpa” assim como acontece na história bíblica de Esaú e Jacó. Aqui neste estudo eu, Iakob, “usurpo” do conhecimento que estes grandes estudiosos deixaram para o legado da linguística e empenho-me a utilizá-los para assim também ajudar a construir um caminho de estudos cada vez maior.

A perspectiva mais ortodoxa sobre os estudos da tradução se importava apenas com o produto final, a tradução *per se*, e pouco ou nada se importava com os processos envolvidos até chegar àquele resultado. Hoje essa visão já não é mais unânime e podemos entender o porquê ao analisar alguns dos resultados de traduções automáticas, sendo tais traduções um tópico de assunto muito comum entre entusiastas e profissionais da tradução tecer comentários acerca das “pérolas” que tais tradutores automáticos cometem quando algum texto lhes é submetido. Caso o processo de tradução fosse irrelevante, já teríamos um ponto final para a longa e morosa história da tradução, utilizaríamos Google Tradutor, DeepL ou qualquer outro tradutor automático e tudo estaria resolvido, o que seria particularmente maravilhoso. Porém, uma

¹ “But we might get to know something, and this something might eventually ‘evolve’ into greater knowledge, better descriptions, better explanations and better understanding.” (CHESTERMAN, 2020)

² An Objective Psychology of Grammar. (KANTOR, 1936)

pessoa indistinta mesmo que não seja profissional da área, e sem ter domínio avançado das duas línguas que estão sendo traduzidas, consegue notar um ruído ou algo estranho em determinadas traduções automáticas. Por a tradução ser um processo em que o objeto que está sendo processado é a linguagem, a conceituação de que a língua é viva deve também ser transposta para a tradução. Sendo assim a tradução é, portanto, viva e evolui conforme a língua.

O ato de traduzir está presente na vida de todos nós desde os primórdios da humanidade. O mundo inteiro depende diretamente ou indiretamente de algum tipo de tradução, seja para fins recreativos (filmes, livros, revistas, quadrinhos, etc.), para trabalho (artigo científico, leis, tradução simultânea [interpretação], etc.) ou uso geral como em bulas de remédios e manuais de instruções. Ao considerarmos o simples ato de falar como a passagem de ideias contidas na mente para o mundo real, falar por si só já constitui um ato tradutório. Um conceito bem explanatório é representado por Octavio Paz (1971) “Aprender a falar é aprender a traduzir: quando a criança pergunta a sua mãe o significado desta ou daquela palavra, o que ela realmente quer é que traduza para sua linguagem o termo desconhecido.”³. Mas afinal de contas, o que é tradução? Vejamos a seguir alguns dos verbetes presentes no dicionário *Michaellis*. Podemos encontrar seis ocorrências:

1 - Ação ou efeito de traduzir;

2 - Transposição ou versão de uma língua para outra; técnica que consiste em traduzir palavra, enunciado, texto, obra etc. falado ou escrito, de uma língua para outra, possibilitando sua compreensão por alguém que não conhece ou não domina a língua em que originalmente o enunciado foi emitido: *Ela se especializou em tradução de textos do inglês para o português*;

3 - Obra traduzida: “*Em março de 1963 a versão original da Anatomia foi lançada no Brasil. E quando no ano seguinte Martim Francisco recebeu o primeiro exemplar da tradução americana, o governo de João Goulart havia sido derrubado e os militares estavam no poder*” (EV);

4 - [LINGUAGEM FIGURADA] Aquilo que indiretamente expressa ou reflete; imagem, reflexo, repercussão: “*A candidez com que Estêvão disse isto era a fiel tradução de seu espírito, e a razão de tais palavras, não a procure o leitor em outra parte mais que não seja aquele sorriso de há pouco [...]*” (MA1);

5 - Ação de tornar claro o significado de algo; explicação, interpretação; e

³ Traduzido por Doralice Alves de Queiroz (FALE/UFMG)

6 - [GENÉTICA] Etapa do processo de síntese de proteínas na qual o RNA (ou ARN) mensageiro opera a síntese da proteína pelo ribossomo; translação.

É notório o quão simplista e não explicativo é a primeira acepção do dicionário. Já a segunda entrada nos mostra uma definição mais bem elaborada com um breve exemplo para esclarecer ainda mais a definição. O terceiro significado fala sobre a tradução em si de obras (literárias ou não), já na quarta possibilidade temos a palavra sendo utilizada com o seu sentido figurado, a quinta entrada nos remete a um dos clássicos tipos de tradução (interlingual) que será abordada em breve, e o sexto e último conceito deste dicionário nos traz um significado inesperadamente complicado pois a palavra “tradução” foi utilizada no campo da Biologia, fazendo assim com que ela ganhasse uma nova realidade para leigos completamente diferente dos demais.

Para facilitar a compreensão de um determinado texto nós podemos reformulá-lo até que fique inteligível o suficiente para o público alvo, e é aí que o conceito-chave dos tipos tradução entra em cena, segundo Jakobson (1959) existem três tipos de tradução:

- a) a tradução intralingual (reformulação), que faz a interpretação dos signos verbais dentro da mesma língua. Neste caso uma tradução intralingual poderia nos ajudar a compreender melhor uma determinada passagem ou até um texto inteiro. Podemos observar como exemplo de tradução intralingual o belíssimo e bem-humorado trabalho dos estudantes Marcelo Henrique “Audino Vilão” e Thiago Torres “Chavoso da USP” que traduzem conceitos filosóficos e políticos para a linguagem da periferia. Os vídeos são disponibilizados na plataforma Youtube;
- b) a tradução interlingual (propriamente dita), que é a interpretação dos signos verbais entre duas línguas distintas. Para ilustrar este tipo de tradução, temos o a obra literária *O pequeno príncipe* (1943) que já foi traduzida para trezentos e oitenta e dois idiomas, detendo assim o recorde de livro não religioso mais traduzido no mundo;⁴ e
- c) a tradução intersemiótica (transmutação), que consiste na interpretação de um sistema de signos para outro. Podemos citar como exemplo dessa tradução as famosas adaptações filmicas que em sua maioria passam de uma semiose (processo de significação) de um livro para filme cinematográfico, porém os exemplos são vastos, pois temos também peças teatrais que viram filmes ou

⁴ <https://www.guinnessworldrecords.com/world-records/most-translated-author-same-book>

livros, desenho animado que vira peça teatral, e entrando em caminhos mais modernos observa-se também várias adaptações de filmes para jogos digitais e vice-versa.

Em se tratando de modernidade, em especial no que tange aos jogos digitais e programas de computadores, nós podemos observar também uma nova modalidade de tradução chamada de “localização”. De forma bastante resumida, a localização nada mais é do que uma tradução com adaptações para se inserir na cultura do local onde a marca ou produto está sendo inserido(a). A localização de uma tradução, apesar de ser um conceito recente (surgiu por volta da década de oitenta), tem um expressivo crescimento e é cada vez mais apreciada e requerida por quem adquire tais produtos ou serviços.

Ademais, assim como ocorre a evolução da língua, a evolução da tradução é uma realidade que acrescenta um novo tipo para a tríade inicial dos modelos de tradução de Jakobson. Hoje temos também a chamada *tradução intermodal*, que faz o papel de traduzir entre línguas de diferentes modalidades, a exemplo desta podemos ter uma tradução intermodal de Português para a Língua Brasileira de Sinais (LIBRAS).

Uma visão mais renovada e contempladora sobre o conceito de tradução é retratada por Guerini e Costa (2006, p. 2) em que diz:

“Atualmente, seu leque de significados é muito amplo e além do original ‘transferir’ quer dizer, entre outras coisas, também ‘transportar, trasladar de uma língua para outra’, ‘revelar, explicar, manifestar, explanar’, ‘representar, simbolizar’. *Traduzir* no sentido de ‘passar de uma língua a outra’ é uma metáfora do ato físico de transferir. (2005, p. 2)”

A tradução é então algo bem mais abrangente do que a passagem de uma língua para outra, a tradução é metalinguística à medida que usa da própria linguagem para ser explicada/feita, o que faz com que o seu processo seja também metacognitivo pois o próprio indivíduo (tradutor) faz o controle de todos os processos cognitivos envolvidos durante a tradução, como a solução e adequação de problemas e a autoavaliação do resultado final (MELLINGER, 2019).

Todo este emaranhado de ideias acerca da tradução “traduz” o quão complexo e meticuloso é este processo. Portanto se torna grave negligenciar a parte da cognição que acontece dentro da mente do tradutor e manter o foco apenas no produto final (texto de chegada), e é neste viés que este trabalho é embasado, nos processos do componente humano.

Para buscar o objetivo, esta dissertação está dividida em cinco capítulos. O capítulo 1 traz a introdução para o tema. No capítulo 2, mostraremos as principais referências teóricas

acerca da tradução e da psicolinguística, e também adotaremos uma visão de bilinguismo como base para o estudo. No capítulo 3, discorreremos acerca da metodologia aplicada durante a execução da pesquisa e a organização dos participantes e dados. No quarto capítulo, apresentamos os resultados para a pesquisa em questão com os devidos comentários. Reservamos o quinto e último capítulo para abordar a conclusão do estudo com as implicações dos achados na educação e tradução, suas limitações e por fim propusemos propostas para futuras pesquisas na área.

2 REFERENCIAL TEÓRICO

Neste capítulo, dialogaremos sobre toda a fundamentação teórica que servirá como base para o estudo em questão. Por ser um estudo que envolve uma ligação entre duas grandes áreas de estudo (tradução e psicolinguística), e que cada uma tem seus diversos pesquisadores com ideais e propostas ímpares, esta seção será dividida nos seguintes subtópicos: A história das mais antigas traduções; a tradução com foco no ensino de língua estrangeira (LE); as categorias de tradução; a tradução como processo; o acesso lexical e o papel da psicolinguística na tradução.

2.1 Breve histórico sobre a tradução no mundo

Muito utilizada no meio acadêmico, em empreendimentos grandiosos e na cultura do lazer, a tradução é um processo indispensável nas nossas vidas, existem até mesmo legiões de fãs, de amadores a profissionais que interagem e discutem sobre tradução “x” ou “y” de determinada obra, e temos logicamente o viés acadêmico que discorre também sobre tais traduções, mas então como se deu o início de tudo isso? Quais foram as primeiras traduções documentadas na nossa história? Vejamos então um compilado que trata de maneira breve e sucinta acerca das traduções mais antigas e impactantes do nosso mundo.

Falar sobre tradução é antes de mais nada, falar sobre interação. Não se sabe ao certo quando e qual foi a primeira tradução que aconteceu no mundo, há relatos de que o *Épico de Gilgamés*, escrito em aproximadamente 2100 a.C., fora o primeiro texto a ser traduzido, sendo este “transportado” do sumério (acadiano) para outras línguas asiáticas. É indiscutível, porém, que um dos achados mais significantes para o campo da tradução seja a Pedra de Roseta, um fragmento de um monólito feito de granodiorito (uma rocha ígnea semelhante ao granito), material este utilizado também em grandes construções como a Vila Adriana e as colunas do Panteão, ambos na Itália. A Pedra de Roseta é datada do ano 196 a.C., mas só foi descoberta em 1799 pelo oficial francês Pierre-François Bouchard durante uma campanha napoleônica no Egito. Após a Capitulação da Alexandria, em 1801, os britânicos saíram vitoriosos contra a

França e levaram a Pedra para o Museu Britânico; o item é inclusive o objeto mais visitado do museu.⁵

Figura 1 – Pedra de Roseta



Fonte:

https://upload.wikimedia.org/wikipedia/commons/2/23/Rosetta_Stone.JPG

A Pedra de Roseta tem cravado em sua superfície um decreto ordenado pelo rei Ptolemeu V Epifânio (210 a.C. – 180 a.C.). O texto fala sobre o pagamento de impostos e taxas e enaltece a generosidade do então rei. O mais interessante é que tal decreto fora escrito em três manuscritos diferentes e duas línguas diferentes, a parte superior da Pedra é escrita em egípcio antigo usando hieróglifos, já a parte do meio é escrita também em egípcio antigo porém utilizando a escrita demótica, isto é, a escrita propriamente dita; a parte inferior nos traz a tradução do decreto escrito em língua grega antiga (clássica).

A descoberta da Pedra teve um papel fundamental para a possibilidade de tradução dos hieróglifos, pois como ela continha o mesmo texto com o seu teor tanto em demótico como em grego antigo, o filólogo e estudioso da cultura oriental Jean-François Champollion conseguiu fazer associações (transliteração) com os caracteres pictográficos até então “desconhecidos” e assim decifraram muitos aspectos da escrita por hieróglifos, aspectos esses

⁵ <https://www.britishmuseum.org/blog/14-things-not-miss-british-museum>
<https://www.britishmuseum.org/visit/object-trails/one-hour-museum#:~:text=The%20Rosetta%20Stone%20is%20the,without%20seeing%20it%20for%20yourself.>

tanto gramaticais quanto fonéticos. A Pedra de Roseta ganhou uma fama tão imensa que ao buscar o nome em dicionários de inglês podemos encontrar geralmente duas definições, vejamos a seguir as definições contidas nos dicionários online *Merriam-Webster*, *Collins* e *Oxford Learners's Dictionary* (tradução nossa):

1 – Dicionário *Merriam-Webster*:

- a) Uma pedra de basalto negro encontrada em 1799 que contém uma inscrição em hieróglifos, caracteres demóticos e grego e é celebrada por ter dado a primeira pista para a decifração dos hieróglifos egípcios;
- b) Aquele que dá dica para o entendimento.⁶

2 – Dicionário *Collins*:

- a) Uma tábua de pedra encontrada em 1799 perto de Rosetta que contém inscrições paralelas em grego, hieróglifo egípcio, e caracteres demóticos, tornando possível a decifração dos hieróglifos egípcios antigos.
- b) Uma dica, um avanço ou descoberta que proporciona um conhecimento crucial para a solução de um problema ou quebra-cabeça.⁷

3 – *Oxford Learner's Dictionary*:

- a) Algo, especialmente uma descoberta, que ajude as pessoas a entender ou achar uma explicação para um mistério ou uma área de conhecimento que pouco se sabia sobre.⁸

Todos os três dicionários consultados apresentam definições similares para o mesmo vocábulo, e o mais interessante foi constatar que o *Oxford Learner's Dictionary* nem sequer chegou a apresentar a definição mais canônica de o que seria a Pedra de Rosetta e foi diretamente mostrando o novo significado que a palavra ganhou conforme toda a sua história.

6

<https://www.merriam-webster.com/dictionary/Rosetta%20stone#:~:text=Rosetta%20stone%20%E2%80%A2%20%5Croh%2DZET,gives%20a%20clue%20to%20understanding>

a) a black basalt stone found in 1799 that bears an inscription in hieroglyphics, demotic characters, and Greek and is celebrated for having given the first clue to the decipherment of Egyptian hieroglyphics

b) one that gives a clue to understanding

⁷ <https://www.collinsdictionary.com/dictionary/english/rosetta-stone>

a) a stone slab, found in 1799 near Rosetta, bearing parallel inscriptions in Greek, Egyptian hieroglyphic, and demotic characters, making possible the decipherment of ancient Egyptian hieroglyphics

b) a clue, breakthrough, or discovery that provides crucial knowledge for the solving of a puzzle or problem

⁸ <https://www.oxfordlearnersdictionaries.com/definition/english/rosetta-stone?q=rosetta+stone>

something, especially a discovery, that helps people to understand or find an explanation for a mystery or an area of knowledge that not much was known about

Para além da Pedra de Rosetta, nós podemos traçar uma trajetória ainda mais categórica e escavar, literalmente, mais a fundo na história e falar sobre o achado das Tábuas de Ebla. Assim como a Pedra, as Tábuas de Ebla só foram descobertas muito tempo depois de sua criação, para ser mais preciso elas foram descobertas entre 1974 e 1975 em Ebla, Síria, pela equipe de escavação do arqueólogo e professor de História da Arte, Paolo Matthiae. As Tábuas de Ebla têm sua criação datada entre 2500 e 2300 a.C., sendo assim ainda mais antiga até que o *Épico de Gilgamés*; ao total as tábuas contabilizam um conjunto de 1800 tabletes de argila com inscrições em escrita cuneiforme em duas línguas, a língua suméria e a língua eblaíta.

Diferentemente do teor aristocrático encontrado da Pedra de Rosetta ou dos lances literários do *Épico de Gilgamés*, as Tábuas de Ebla possuem em seus inscritos um conteúdo mais voltado para o registro propriamente dito, tanto é que as Tábuas continham descrições de registros econômicos de importação e exportação, além do registro de relações comerciais e políticas com outras cidades; o texto traz também descrições de que a cidade-império de Ebla realizava comércio de gado, trigo, ovelha, vinho e cevada, e também era produtora de cervejas, uma delas intitulada de “Ebla” em homenagem à cidade.

Como pode ser observado, a interação que a tradução busca entre os interlocutores é algo que já acontece há mais de quatro mil anos, tanto a Pedra de Rosetta quanto as Tábuas de Ebla são documentos históricos importantes que nos contam um pouco sobre esse processo (ou seria arte?) e como pode afetar de maneira positiva as relações entre culturas de povos diferentes, já o *Épico de Gilgamés* deve ser tratado de uma maneira diferenciada, atente para o uso da palavra “diferenciada”, pois dentre os três achados históricos ele é o único que se trata de uma obra literária, e sendo assim difere-se dos demais.

Na próxima seção abordaremos como a tradução evoluiu de seu contexto exclusivamente manual e passou a ser também um processo automatizado feito por intermédio de computadores.

2.1.1 Panorama da tradução automática.

A tradução automática da maneira que temos à nossa disposição nem sempre foi como conhecemos: as máquinas evoluíram bastante com o passar do tempo e hoje conseguem produzir traduções a níveis que rivalizam com a tradução humana (Muftah, 2022).

Nos primórdios da tradução automática, em meados da década de cinquenta, os computadores eram volumosos e tinham uma capacidade muito limitada em comparação às atuais máquinas e por isso a tradução era restrita apenas a traduzir um pequeno compilado de

palavras preestabelecidas. Em sete de janeiro de 1954 o experimento conhecido como “Experimento de Georgetown-IBM” mostrou ao público a capacidade do computador IBM 701 traduzir algumas frases do par linguístico Russo - Inglês.

Figura 2 – Operação do computador IBM 701



Fonte:

http://www.cs.umd.edu/class/fall2018/cmssc470/slides/slides_18.pdf

A tradução do Experimento de Georgetown-IBM, apesar de rudimentar e previamente doutrinada, atraiu a atenção pública e com isso a tradução automática foi alvo de fomento e conseguiu se estabelecer como um novo campo de estudo com um futuro muito promissor pois todos que viram aquilo desejavam essa facilidade de comunicação. Quase uma década havia passado e a tradução automática não viu avanços em comparação àquela demonstração em 1954, pois o experimento teria sido mais uma simulação do que uma comprovação prática, pois o computador já estava pré-configurado para fazer tal tradução (Hutchins, 1998; 2004).

Com os computadores um pouco mais desenvolvidos foi possível criar a “Tradução Automática Baseada em Regras” (Rule-Based Machine Translation) na década de 70. Esta nova etapa marcava uma era em que os textos começavam a ficar legíveis e precisavam de menos intervenção humana na parte da pós-edição. Foi somente no final da década de 80 e início dos anos 90, quando as máquinas deram um salto na tecnologia, que nós pudemos observar traduções automáticas com mais precisão e autonomia. A partir daqui a tradução automática se torna mais sofisticada e temos o que é chamado de “Tradução Automática Estatística” (Statistical Machine Translation) que tinha a vantagem de associar robustos corpora em suas traduções. Ainda na década de 90 surge o que é chamado de “Tradução Automática Híbrida” (Hybrid Machine Translation) que, como sugere o nome, faz a junção das operações da tradução baseada em regras com a tradução por estatística.

Contudo, nenhum desses métodos de tradução supracitados funcionam tão bem quanto o modelo que nós temos hoje em dia chamado de “Tradução Automática Neural” (Neural Machine Translation). Este modo de tradução foi primeiramente proposto por Cho *et al.* (2014) e apesar de não ter atraído muita atenção à época da publicação, grandes empresas como a Google utilizaram a ideia da rede neural para implementar seus tradutores; foi então em 2016 que a gigante da tecnologia aplicou em seu produto “Google Tradutor” esse novo modelo de tradução automática.

A tradução automática neural acontece de maneira análoga ao próprio pensamento humano, muito diferente das versões antigas de traduções automáticas em que geralmente o criador do sistema teria que alimentar um banco de dados com regras e corpus para que este funcionasse. Na modalidade neural, o sistema tradutor “aprende” semântica e sintaxe; em outras palavras, é como se o próprio tradutor automático possuísse uma inteligência artificial que aprende a cada dia e toda vez que nós o utilizamos.

Para termos uma ideia de como a tradução automática neural representa um salto na qualidade geral das traduções automáticas, vejamos a seguir dois textos utilizados por Precup-Stiegelbauer (2013) para comparar traduções automáticas do romeno para o inglês.

Quadro 1 – Comparativo entre a tradução automática do Google em 2012 (pré-neural) e 2022 (neural) do texto *Superstitie* de Marin Sorescu.

Texto fonte	Google Tradutor (2012)	Google Tradutor (2022)
Superstitie	Superstition	Superstition
"Pisica mea se spală Cu laba stângă, Iar o să avem un război. Fiindcă, am observat, De câte ori se spală Crește considerabil încordarea Internațională.	"My cat washed With the left paw. And have war I noticed how many times I washed With the left paw Considerable crests the strain International.	My cat is bathing With the left paw, And we will have a war. Because, I noticed, How many times it is washed With the left paw, The tension increases considerably International.
Cum vede ea Cele cinci continente? S-a mutat cumva. În pupilele sale Cea care știe să-ți spună dinainte	How he sees she. The one five continents? He moved somehow. In his pupils. He who say The history.	As she sees The five continents? He moved somehow In his pupils

Toată istoria Fără punctuație?	Charm. Comes me weep. When think I Prompt it which mixed I bound it. Rearwards. Depending, in last of the of a caprices cats! Absconds from grasps? Don't the maul world. Character of the devil. Of stench!"	Pythia, the one who knows how to tell you in advance All history No punctuation? I want to cry When I think that I And the sky with souls that I tied to me Behind It ultimately depends on By the sounds of a cat! Run away from catching mice, Stop unleashing world wars, Damn you It stinks!
-----------------------------------	--	--

Fonte: Adaptado de Precup-Stiegelbauer, L-R (2013)

Nota: Para compor o quadro utilizamos a tradução destacada em 2013 no artigo *Automatic translations versus human translations in nowadays world* de Precup-Stiegelbauer, L-R que utilizamos a versão antiga de motor de tradução do google e criamos uma terceira coluna com a tradução do google utilizando o novo modelo de tradução neural.

Existe uma clara diferença entre as traduções de 2012 e 2022, mesmo uma pessoa sem o menor conhecimento entre as línguas romenas e inglesa conseguiria notar elementos não diretamente relacionados à tradução como a própria extensão do texto em si. Na tradução pré-neural (2012) o texto alvo fica muito “similar” ao texto fonte, na medida em que a tradução se parece mais com um jogo de palavras em que fora traduzido uma por uma sem muito esmero com o resultado final; e para um conhecedor da língua em questão é de chamar atenção também a ininteligibilidade que fica presente no texto traduzido (2012), partes do texto contém inclusive desvios de padrões simples como no verso “*How he sees she*” em que os pronomes em destaques têm função sintática de sujeito, quando na verdade o último pronome deveria ter a função de objeto, e este não é o único desvio padrão, apesar de o texto a ser traduzido tratar-se de literatura, a “liberdade poética” fica demasiadamente exagerada no verso traduzido “in the last of the of a” que em tradução livre para o português ficaria algo como “no último do o de um”. A tradução feita pelo Google em 2012 é de fato um texto que aparenta ser truncado, é como se a tradução não estivesse completa e permanecesse na margem do possível.

Já no texto traduzido pelo mesmo tradutor, porém em 2022, com a tecnologia neural, podemos notar um “acréscimo” de palavras ou de estruturas sintáticas em relação ao texto fonte. É interessante notar também que nesta versão há um trecho no qual pode-se notar a presença de assonância do som do ditongo vocálico (glide) “[aɪ]” em: “**I** [aɪ] want to **CRY** [kraɪ] when **I** [aɪ] think that **I** [aɪ] And the **SKY** [skaɪ] with souls that **I** [aɪ] **TIED** [taɪd] to me **BEHIND** [bɪ'hamd]”, creio que esta figura de linguagem foi uma feliz coincidência na tradução. Ademais, o texto é perfeitamente inteligível. Uma pessoa que domina a língua inglesa pode até estranhar a alternância de pronomes e a agramaticalidade em alguns pontos do texto, porém fica perceptível que a ideia geral está presente e pode ser facilmente captada.

O segundo texto utilizado por Precup-Stiegelbauer (2013) é também da autoria de Marin Sorescu. O texto intitulado de *Shakespeare* retrata de forma poética a trajetória do Bardo de Avon, fazendo um comparativo entre as temáticas e obras de Shakespeare com o livro bíblico de Genesis e a criação do mundo por Deus.

Quadro 2 – Comparativo entre a tradução automática do Google em 2012 (pré-neural) e 2022 (neural) de parte do texto *Shakespeare* de Marin Sorescu.

Texto fonte	Google Tradutor (2012)	Google Tradutor (2022)
Shakespeare	Shakespeare	Shakespeare
"Shakespeare a creat lumea în șapte zile. În prima zi a făcut cerul, pământul și prăpăstiile sufletești.	"Shakespeare created the world in able-bodied days. In first charmed cut prompt sky, earth and deep souls.	"Shakespeare created the world in seven days. On the first day he made the sky, the earth and the spiritual chasms.
În a doua zi a făcut râurile, mările, oceanele	In second charmed cut the rivers, seas, the oceans	On the second day he made the rivers, the seas, the oceans
Și celelalte sentimente Și le-a dat lui Hamlet, lui Iulius Caesar, lui Antoniu, Cleopatrei și Ofeliei,	And other feelings. He gave them his Hamlet, Iulius his Caesar Antoniu, Cleopatra his Ofelia Othello	And the other feelings And he gave them to Hamlet, Julius Caesar, Antony, Cleopatra, and Ophelia,
Lui Othello și alora,	and other.	To Othello and others,

<p>Să le stăpânească, ei și urmașii lor,</p> <p>În vecii vecilor.</p> <p>În ziua a treia a strâns toți oamenii</p> <p>Și i-a învățat gusturile:</p> <p>Gustul fericirii, al iubirii, al deznădejdiei,</p> <p>Gustul geloziei, al gloriei și așa mai departe, până s-au terminat toate gusturile."</p>	<p>They follow them in eternities.</p> <p>In diurnally gathered the peoples</p> <p>Teach the tastes:</p> <p>The taste of happiness, of the love,</p> <p>The taste of the jealousy, of the glory, of all the tastes".</p>	<p>Let them and their descendants rule them,</p> <p>Forever and ever.</p> <p>On the third day he gathered all the people</p> <p>And learned his tastes:</p> <p>The taste of happiness, of love, of despair,</p> <p>The taste of jealousy, the taste of glory, and so on, until all tastes are finished."</p>
---	--	--

Fonte: Adaptado de Precup-Stiegelbauer, L-R (2013)

Nota: Para compor o quadro utilizamos a tradução destacada em 2013 no artigo *Automatic translations versus human translations in nowadays world* de Precup-Stiegelbauer, L-R que utilizada a versão antiga de motor de tradução do google e criamos uma terceira coluna com a tradução do google utilizando o novo modelo de tradução neural.

As mesmas observações acerca do texto *Superstitie* podem ser feitas para *Shakespeare*: a tradução de 2012 é truncada, ininteligível e contém inúmeros erros sintáticos e até semânticos se levarmos em consideração a tradução de 2022.

A autora do artigo transborda em sua conclusão o quão imprecisa é a tradução automática dos textos escolhidos. Vemos assim em Precup-Stiegelbauer (2013, p. 1776, traduzido por Google):

“Lendo as traduções acima, pode-se facilmente perceber que todas elas estão longe de serem precisas. As palavras de ligação não foram traduzidas, enquanto as terminações das palavras não foram levadas em consideração. Além disso, há palavras que nem foram traduzidas, porque o computador não conseguiu reconhecê-las. Em suma, pode-se apenas entender o que o original diz, pois é óbvio que as traduções foram feitas mecanicamente e precisam de muito aperfeiçoamento para serem completamente compreensíveis.

No entanto, a tecnologia MT está melhorando o tempo todo. Muitos programas estão sendo executados em todo o mundo agora e constituem uma área interessante de pesquisa de tradução, especialmente quando combinada com o toque humano. É

provável que, ao longo do tempo, esta pesquisa vá ampliando gradativamente os limites dentro dos quais o MT pode operar ou dentro de sua evolução.”⁹

Porém a autora reconhece que o campo da tradução automática estava evoluindo e tinha espaço para muitas melhorias, fato que pode ser observado com a tradução feita nesta dissertação com novo modelo de tradução neural do Google do par inglês/português do excerto do texto da própria autora (QUADRO 1 e QUADRO 2). Com o avanço da tecnologia, a tradução por rede neural efetuada pela máquina está cada vez mais próxima da forma de pensar dos seres humanos, isto é, a tradução automática não está mais limitada a dicionários e regras pré-estabelecidas; agora a máquina consegue processar e racionalizar de uma melhor forma as etapas da tradução; a cognição computacional é uma realidade que permitiu tamanho avanço nesta área que ficou por muito tempo delimitada por questões tecnológicas.

Há um debate incessável acerca da temática entre a qualidade da tradução humana versus a qualidade da tradução automática: vários autores (Ahrenberg, 2017; Hasibuan, 2020; Cadwell; O’Brien; Teixeira, 2018) alegam que a tradução automática não está em um nível de aceitabilidade de acordo com a tradução humana e por isso não deve ser utilizada, ou ainda, caso necessitem utilizá-la, que mantenham cautela durante o uso. Podemos ver também casos mais extremos como em Boukreris e Labeled (2016) que afirmam ser uma “falha” trocar um humano por uma tradução totalmente automática por conta desta última não ser capaz de interpretação. No entanto, existe também uma série de estudos (Muftah, 2022; Takakusagi *et al.*, 2021; Turner *et al.*, 2014) que advogam a favor da tradução automática, não no sentido de ela ser melhor ou poder trocar completamente a parte humana, mas sim que em determinadas situações a sua utilização se mostra tão válida quanto uma tradução humana; há um caso no estudo de Fischer e Läubli (2020) em que a tradução automática do par linguístico alemão—italiano mostrou resultados melhores do que a tradução humana por ter mostrado mais segmentos que não necessitaram de pós edição e um número ligeiramente menor de segmentos que necessitaram de correções na pós edição do texto traduzido.

⁹ “Reading the translations above, one can easily notice that all of them are far from being accurate. Linking words have not been translated, while word' endings have not been taken into account. Moreover, there are words that have not even been translated, because the computer could not recognize them. All in all, one can only get a grasp of what the original says, as it is obvious that translations have been made mechanically and they need much improvement in order to be completely understandable.

However, MT technology is improving all the time. Many programs are running around the world right now and it constitutes an exciting area of translation research, especially when combined with the human touch. It is likely that, over time, this research will gradually extend the boundaries within which the MT can operate or within its evolution.”

2.2 A tradução como processo

Nesta seção abordaremos as perspectivas de como os teóricos do campo dos Estudos da Tradução a veem como um processo. Vários conceitos discutidos aqui se entrelaçam ou complementam-se para formar o arcabouço teórico mais palpável possível acerca do escopo deste estudo.

Segundo Holmes (1972), os Estudos da Tradução podem ser divididos em “puro” e “aplicados”. Por sua vez, o estudo puro é subdividido em “teórico” e “descritivo”. A parte descritiva engloba ainda três divisões: estudo descritivo de tradução orientado ao produto (ênfase no resultado da tradução), estudo descritivo de tradução orientado ao processo (ênfase nos processos cognitivos da tradução) e estudo descritivo de tradução orientado à função (ênfase na tentativa de descrever a tradução no seu contexto sociocultural inserido). Das três (sub)divisões de Holmes, a que mais nos interessa é a do estudo descritivo da tradução orientado ao processo, pois é aí onde habita a peça principal do nosso objeto de estudo.

A tradução orientada ao processo tenta entender exatamente o que se passa dentro da “caixa preta” que é a mente do tradutor à medida em que ele cria o novo texto em uma outra língua (Holmes, 1988). O termo caixa preta (ou seria cinza?) é louvavelmente apropriado para o objeto em questão pois o que se passa dentro da cabeça de qualquer um que faça uma tradução ainda permanece um profundo mistério. As outras divisões comportam muito bem o que antigamente era demasiadamente enfatizado como sendo o propósito dos estudos da tradução (resultado), e também se envolvem com a proposta de Even-Zohar (1990; 1997) ao reconhecer o status mais vivo/humano da tradução (contexto sociocultural). Seria leigo, porém, invalidar ou negligenciar as outras pontas desta divisão e focar com visão em túnel apenas no processo, até mesmo porque Holmes admite que não é uma divisão totalmente segregadora entre os três princípios, na verdade é o contrário, um não pode ser compreendido sem que o outro seja estudado.

Ao dissertar sobre a tradução de palavras, Dijkstra *et al.* (2018, p. 17, tradução nossa) discorrem que “[...] apesar de todo bilíngue, desde uma jovem criança a um intérprete profissional, conseguir até certo ponto realizar essa tarefa, ela é uma das atividades linguísticas mais complexas que um falante humano pode se enveredar a fazer.”¹⁰. Esta passagem explícita e resume de uma maneira ímpar o quão difícil é a atividade de tradução, e por ter essa

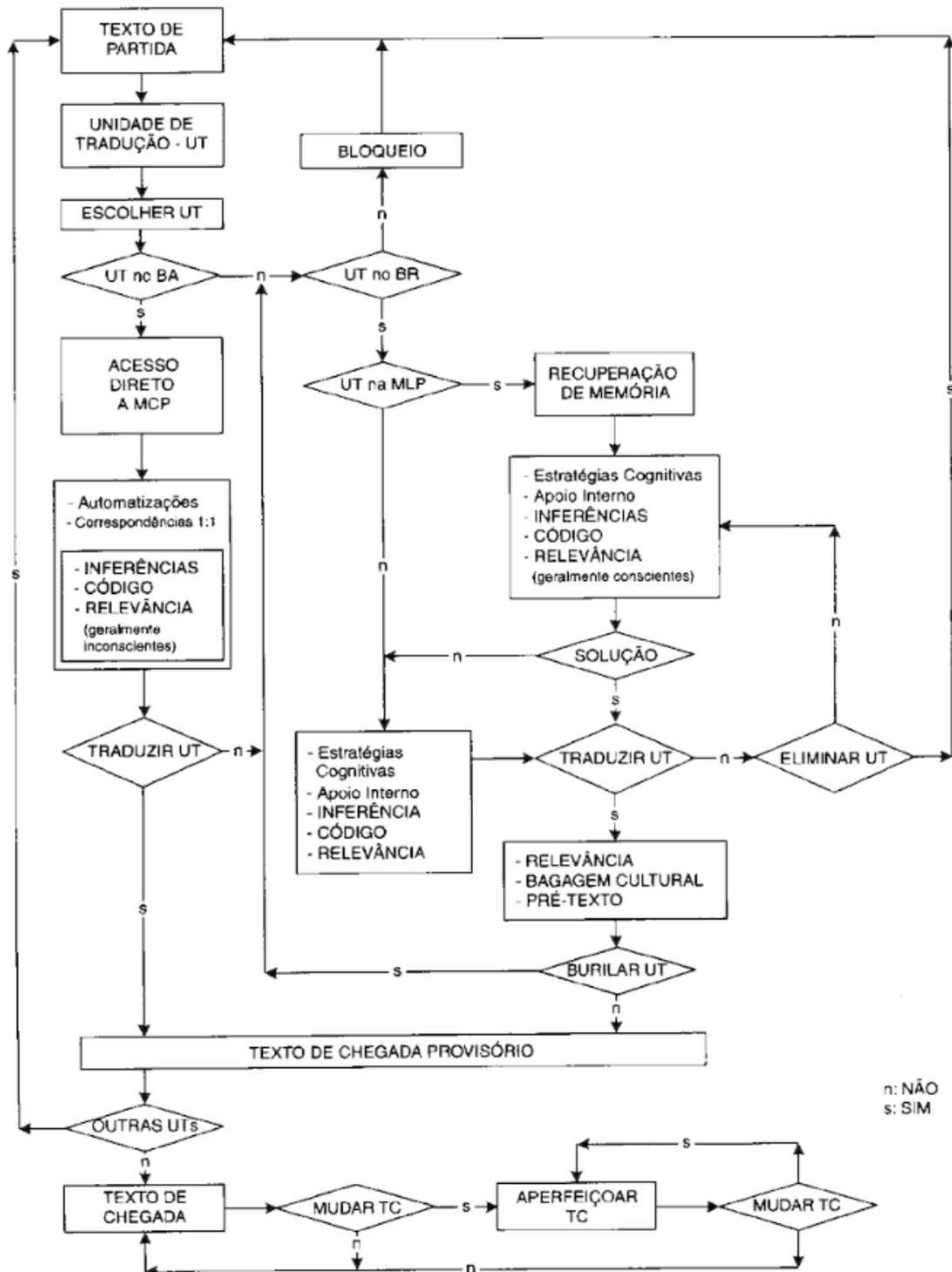
¹⁰ “although every bilingual from a young child to a professional interpreter can to some extent perform this task, it is one of the most complex language activities that a human speaker can engage in”

dificuldade intrínseca à tarefa, a sua análise também não poderia deixar de ser menos complicada.

Nos primórdios dos estudos da tradução como processo, os pesquisadores utilizavam os protocolos verbais, mais conhecidos como TAP (think-aloud protocol) para se ter uma melhor compreensão do processo tradutório (cf. Krings, 1986; Jääskeläinen, 1987; Löscher, 1991; Kussmaul, 1995; Kiraly, 1995; Seguinot, 1989; Dancette, 1995). Os protocolos verbais consistem em o participante do experimento (neste caso a tradução) falar literalmente o que passa em sua mente durante toda a atividade de tradução; apesar de não ser um método intrusivo, vimos que a tradução *per se* já constitui uma tarefa que demanda um alto esforço cognitivo (Alves, 2001; Shahrokhi, 2016), e, adicionando a esta tarefa uma sobrecarga cognitiva por uma atividade subsequente, incrementa um peso inesperado à pesquisa. Bernardini (2002) já se preocupava com a forma em que os protocolos verbais eram adotados em pesquisas de tradução orientada ao processo, e ela atacou a falta de maturidade com que o método era utilizado, sendo as vezes os seus resultados representados “em forma de anedota” e sem critérios de categoria, transformando-se assim num “vazio teórico”. Não serei aqui um obstinado defensor do uso de protocolos verbais, nem tampouco um *hater* de quem faz uso dele. Vejo que na pesquisa da tradução há, sim, espaço para a utilização protocolos verbais, todavia, o pesquisador que optar por usar tal método deve estar ciente de todos os contras e tudo mais que isso implica nos resultados apurados na exploração de dados, e deverá tratar e explicitar todo e qualquer ruído advindo deste método para que outros pesquisadores possam vir a replicar um determinado experimento.

Alves (2015) discorre que muito ainda tem de ser feito em pesquisas do processo tradutório, pois a grande maioria dos estudos que envolvem a tradução como processo ainda baseia-se ou se utiliza bastante dos arsenais de ideias de outros campos de estudo. O autor cita uma série de publicações que trouxeram à tona toda a complexidade pertinente ao processo tradutório como as obras de (Göpferich; Jakobsen; Mees (eds.), 2009; Mees; Alves; Göpferich (eds.), 2009; Göpferich; Alves; Mees (eds.), 2010; Shreve; Angelone (eds.), 2010; Alvstad; Hild; Tiselius (eds.), 2011). Para se ter uma dimensão da complexidade da tarefa que é traduzir, vejamos o esquema elaborado por Fábio Alves (1997) como um possível modelo de adentrar na mente do tradutor durante o processo tradutório.

Figura 3 – Modelo do Processo Tradutório



Fonte: (ALVES, 1997:28).

Este esquema, em forma de fluxograma, é uma tentativa ilustrada de demonstrar o passo a passo de cada etapa da tradução, com os losangos fazendo papel dos processos de decisão (sim ou não) e os retângulos os processos “compulsórios” (nem todos os processos são terminantemente obrigatórios, tudo dependerá das escolhas do tradutor).

Um bom exercício para transformar esse fluxograma em prosa é a leitura do trecho escrito por Bruno Osimo no sítio Logos (traduzido por Mauro Rubens da Silva e Nadia Fossa):

“O tradutor que lê o texto que vai traduzir o faz projetando os possíveis metatextos em um espaço virtual, no qual o novo texto começa a tomar forma; primeiro, como material mental (processamento do material tal como percebe o tradutor) e, a seguir, como a inserção concreta de tal material em uma estrutura rígida e convencional: o código do futuro metatexto (a linguagem do texto traduzido). A mente humana examina, de maneira veloz e nem sempre totalmente consciente, as diferentes possibilidades de projeção do prototexto na língua do metatexto e, mediante um processo de eleição, que guarda muito em comum com a teoria dos jogos, **decide-se pela solução que considera ótima entre as previstas.**”¹¹(Grifo nosso).

É, portanto, interessante observar que a tradução passa por uma maiêutica socrática (autorreflexão, sucessão de respostas até encontrar a mais correta) quando atinge o que poderia ser o ponto final (texto de chegada), com uma “pequena” diferença que o tradutor deve conversar consigo próprio para obter as suas respostas. É de se esperar que ao findar a tradução sempre exista a possibilidade do questionamento “Ficou bom ou precisa melhorar?” e saltando de Sócrates para Shakespeare caímos na eterna pergunta “Melhorar ou não melhorar? Eis a questão!”; a resposta para a pergunta habita única e exclusivamente no tradutor e no tempo que ele tem disponível para realizar o trabalho.

2.2.1 O tradutor como centro da tradução

Sendo assim vemos que a tradução não está atrelada somente ao fator linguístico; o fator humano, a cognição, a expertise e a experiência de vida do tradutor são fatores que se somam e mudam a perspectiva do que antes era visto, não como simples, mas, como um estudo que deixava de lado tantas características cruciais para o seu entendimento

¹¹ “Il traduttore, che legge il testo che dovrà tradurre, lo legge proiettando i possibili metatesti in uno spazio virtuale, nel quale comincia a prendere forma il testo nuovo, prima in termini di materiale mentale (elaborazione del materiale percepito dal traduttore), poi in termini di concreto inserimento di tale materiale in una struttura rigida e convenzionale che è il codice del futuro metatesto (la lingua della traduzione). La mente umana prende in esame - in modo veloce e non sempre del tutto consapevole - le varie possibilità potenziali di proiezione del prototesto nella lingua del metatesto e, con un procedimento di scelta che ha molto a che vedere con la teoria dei giochi, opta per la soluzione considerata ottimale tra quelle prefigurate.”

(Alves; Magalhães; Pagano, 2000). Toda essa ciência envolta da tradução como processo e jogo de decisões tem ganhado força desde os primeiros ensaios do tema, é tanto que podemos ver em Levý (1967, traduzido por Gustavo Althoff e Cristiane Vidal) a mesma associação da tradução com um jogo:

“[...] Do ponto de vista do trabalho do tradutor, a qualquer momento desse trabalho (isto é, do ponto de vista pragmático), traduzir é um PROCESSO DE TOMADA DE DECISÃO: uma série de um certo número de situações consecutivas – movimentos, como em um jogo – que impõem ao tradutor a necessidade de escolher dentre um certo número (muitas vezes exatamente definível) de alternativas.”¹²

De um ponto de vista mais analítico, são justamente essas alternativas da tradução que podem dificultar o trabalho do tradutor, tanto a tradução de literatura quanto uma tradução livre ou técnica sofrem muito com esta tomada de decisões pois o tradutor tenta ser o mais invisível possível para assim cumprir com o papel de uma tradução “ideal” (Shapiro, 1986)¹³, porém na maioria das vezes o tradutor se vê em uma escolha em cascata iniciando no paradigma semântico da palavra (ou de um construto semântico mais complexo) no texto de partida (fonte) em que suas várias interpretações devem ser levadas em consideração e logo após no paradigma de palavras (construtor verbais) da língua alvo (chegada) que leva em consideração a escolha feita no paradigma anterior (Levý, 1968, p. 1117).

A invisibilidade do tradutor é um assunto que vira tópico central para discussões entre os estudiosos do tema; uma das frases mais comentadas sobre essa questão é o antigo ditado italiano “traduttore, traditore” que quer dizer algo como “tradutor, traidor” pois os defensores da famosa tradução “perfeita, ideal, imaculada, *et cetera*” alegam que toda e qualquer tradução perde a essência, o seu caráter “original” e por isso toda tradução é consequentemente uma traição do ideal imaginário do autor original da obra. Bem, é uma maneira de pensar sobre tradução, um pouco exagerada se pararmos para pensar, é como se a palavra em português “lança de arremesso” não fosse a tradução do inglês “javelin”, ou o contrário, como se o clássico “saudade” não pudesse ser traduzido para “miss” pois perde a essência da palavra; tais alegações constituem uma aberração do ponto de vista lógico, analogamente é como dizer que a palavra inglesa “miss” não é ter saudade de alguém, ou até mesmo simplesmente manifestar que o sentimento da palavra “miss” é menor do que “saudade”

¹² “From the point of view of the working situation of the translator at any moment of his work (that is from the pragmatic point of view), translating is a DECISION PROCESS: a series of a certain number of consecutive situations—moves, as in a game—situations imposing on the translator the necessity of choosing among a certain (and very often exactly definable) number of alternatives.”

¹³ “I see translation as an attempt to produce a text so transparent that it does not seem to be a translation. A good translation is like a pane of glass. You only notice that it's there when there are little imperfections -- scratches, bubbles. Ideally, there shouldn't be any. It should never call attention to itself.”

é algo que foge do racional. Qual seria então o critério para ser uma tradução “perfeita”? É aquela que o leitor concordar que ficou boa? Qual leitor? Existe um precipício de questionamentos a serem feitos sobre este assunto.

Apenas para fazer um curto jogo prático sobre a tomada de decisão(ões) na tradução, peguemos por exemplo a vigésima terceira letra do alfabeto português, a letra “w”. Conhecemos esta letra pelo seu nome “dáblio” [d’ablju], que tem esse nome/som por conta da letra se chamar “double u” (duplo u) na língua inglesa, porém a forma da letra se assemelha mais com um “duplo v” ao invés de “u” e a sua fonética na língua portuguesa também carrega essa característica dupla, ou sem “identidade cultural”; as palavras de origem alemã que tem o “w” em sua grafia são foneticamente pronunciadas com som de “v” (e.g. Wagner [vag’neʁ]) enquanto as palavras de origem inglesa teriam o som de “u” (e.g. Washington [uaʃɪŋ’tõ]), porém por ser algo sem grande relevância para a sociedade como um todo, a letra “w” acaba sendo ignorada, e nomes como o britânico Darwin [’da:rwɪn] acabam sendo pronunciado com som de “v” e em contrapartida o compositor austríaco Wolfgang [’vɔlfgaŋ] Mozart tem por muitas vezes seu nome pronunciado com o som de “u”.

Um caso recente que envolve a manipulação fonética da letra “w” é no grande sucesso do final de 2022 da Netflix, *Wandinha* (2022). Na série, a personagem principal se chama “Wednesday” (quarta-feira) no original em inglês, por conta disso tem seu nome abreviado para “Wendy”, e aqui no Brasil de modo a soar parecido com o nome original, adotou-se o nome “Wandinha” com o “w” sendo pronunciado como um “v” fugindo assim da pronúncia original em que esta mesma letra tem o som de “u”. Nem sempre (V)Wandinha teve esse nome: durante a sua primeira aparição em terras brasileiras ela era chamada de Benenilda, porém este nome não conseguiu se manter. Uma outra versão desse nome não-escolhido fora Venenilda (que é até interessante por parecer com a palavra “veneno”, encaixando-se assim com as características da personagem), partindo também da proximidade entre as letras “V” e “B” que sofreram um metaplasmo (fenômeno de mudança de grafia) chamado de Betacismo, em que atualmente a letra “V” é mais comumente utilizada; este fenômeno é notado por exemplo ao ver pessoas mais antigas pronunciarem palavras como “bassoura” ou “assobio”. Em outras nacionalidades a ideia de “quarta-feira” é mantida no nome, na Espanha ela é chamada de “Miércoles” e nos países falantes de espanhol da América Latina é conhecida por “Merlina”, na França ela se chama “Mercredi” e seu nome é “Mercoledì” em italiano; em vários outros países seu nome não muda, uma provável explicação de o porquê se chamar (V)Wandinha em português pode estar atrelada à similaridade não só fonética mas também de tamanho entre os nomes Wednesday/Wandinha fazendo assim com que algum aspecto do nome original

transborde para a tradução. Outro aspecto que deixa pistas sobre a escolha da troca do “V” pelo “W” no nome da protagonista pode estar relacionado à sua vestimenta característica possuir a gola em formato de “w” como visto da figura 4.

Figura 4 – (V)Wandinha Addams em 4 fases: Cartum (1938), Lisa Loring (1964), Christina Ricci (1991) e Jenna Ortega (2022)



Fonte: Montagem elaborada pelo autor.

Todo esse jogo de escolhas culminou em uma cena ímpar no primeiro capítulo da série *Wandinha* (2022). Em um diálogo entre a protagonista e a diretora da escola, o assunto sobre o nome peculiar é mostrado ao espectador, e a conversa dublada em português ocorre da seguinte maneira “(Diretora Weems) – Wandinha, que nome mais singular! Imagino que foi porque nasceu toda fofinha.” (Wandinha Addams) – Eu nasci numa sexta-feira 13”, já no áudio original em inglês temos a seguinte conversa “(Principal Weems) – Wednesday such a unique name! I’m guessing it was the day of the week you were born!?” (Wednesday Addams) – “I was born on a Friday the 13th”, toda a conversa revolve sobre o tópico do nome Wednesday (quarta-feira) e a dublagem para o português sofreu uma perda de sentido por conta do nome adaptado “Wandinha” não fazer referência ou ter relação alguma com o dia da semana; já a legenda em português tomou um rumo inusitado e tentou abranger a referência ao dia da semana de alguma maneira, nela podemos ver a seguinte transcrição quanto ao motivo do nome “(Diretora Weems) – Homenagem a uma pessoa ou *data*?”. Mostrado desta forma pode-se entender que a dublagem para o português desvirtua em algum aspecto, porém o espectador que se depara com a cena pode notar o tom de ironia quando a diretora fala “fofinha” que traz justamente a mensagem irônica de antítese entre o que é falado e o que é visto, portanto é ultrajante afirmar qual das escolhas funciona melhor, principalmente em uma mídia que acopla vários outros fatores como cenário, movimentação, tonalidade e não somente o texto escrito.

Vale salientar o nome que primeiramente vingou no Brasil foi “Vandinha” (com V) por questões de aceitação cultural¹⁴, mas não só a protagonista teve seu nome adaptado para a versão brasileira; outros personagens como o Tio Chico (Uncle Fester) também sofreram uma alteração com a expectativa de melhor se adequar ao público brasileiro. Uma coisa é fato: a série mostra uma família fora do comum com personagens de nomes bem únicos e, inclusive, o mordomo da família, Tropeço (em inglês “Lurch”), tem seu nome “traduzido” de forma literal e se fossemos pensar desta maneira, na tradução mais “ao pé da letra” dos nomes dos personagens faz sentido ao comparar com os outros países onde a série é exibida; ficaria até bem de acordo com a tonalidade mórbida da série se o personagem do Tio Chico fosse chamado de “Tio Chagas” para fazer referência a doença de chagas ou até mesmo ao substantivo “chaga (ferida, úlcera)”. No mesmo ritmo de adaptação/tradução de nomes, há também no emblemático *Harry Potter* um nome que causou um certo estranhamento e desconforto entre os fãs mais ávidos da série. O pai do bruxo é chamado de James Potter na história original em inglês e aqui em terras brasileiras o nome foi adaptado/traduzido para Tiago Potter. Questionada sobre como foi lidar com a tradução de vários dos nomes próprios encontrados na obra de J.K Rowling, a tradutora brasileira Lia Wyler respondeu o seguinte:

“Recriando-os, procurando remontar às mesmas fontes que a autora, consultando-a. Aliás, à minha primeira consulta, a Sra. Rowling - que fala o português - me respondeu pedindo que lhe mandasse uma lista das traduções a que eu chegara. Ela as aprovou sem ressalvas - os livros ainda não tinham virado série, não eram sucesso mundial, nem assunto para tablóides e revistas domingueiras.” (Lia Wyler, 2001).

Ao levarmos em consideração o mundo de feitiçaria da obra fictícia como em *Harry Potter*, em que há outros personagens já com seus nomes na versão brasileira tais como - Cedrico, Hermione, Voldemort, Neville e Dumbledore -, é intrigante imaginar o porquê do nome “James” ter que virar “Tiago”, porém a série virou um fenômeno mundial e por aqui não foi diferente, todas as decisões tomadas pela tradutora parecem ter sido acertadas, pois o público brasileiro ficou cativado pelo mundo (re)criado em nossa língua.

É de fato uma situação caótica tentar equilibrar a tradução e manter coerência durante toda uma história que foi pensada em outra língua. Estes exemplos com nomes próprios são mais alguns dos casos em que a (im)possibilidade de tradução ocorre. Todas essas escolhas são pensadas pelo tradutor na hora do processo tradutório, o marketing, a recepção pelo público,

¹⁴ <https://www.metropoles.com/entretenimento/televisao/diretor-explica-por-que-wednesday-vingou-na-versao-brasileira>

as similaridades com o original: tudo entra no árduo jogo de escolhas e cabe ao profissional tentar percorrer o caminho mais assertivo possível.

Um modo de pensar que dialoga melhor com a (im)possibilidade de tradução é demonstrado no livro *Escola de Tradutores* de Paulo Rónai. Lá podemos observar o esclarecimento do autor acerca da incapacidade de produzir uma tradução literária, porém ele reconhece que mesmo com esse empecilho quase proibitivo, todos os tradutores ainda assim fazem a tradução. Rónai faz uma comparação entre a tradução e a arte: ora, se um texto literário é uma obra de arte, e por que a sua tradução também não deveria ser? Rónai afirma que “O objetivo de toda arte não é algo impossível? O poeta exprime (ou quer exprimir) o inexprimível, o pintor reproduz o irreproduzível, o estatuário fixa o infixável. Não é surpreendente, pois, que o tradutor se empenhe em traduzir o intraduzível.” (Rónai, 1987, p. 14). Assim conseguimos combinar o melhor dos dois mundos; continua-se a ideia de que a língua original é intraduzível, mas ao mesmo tempo sabe-se que é o dever do tradutor fazer com que essa impossibilidade se concretize em novas palavras com “novos” significados, ao invés de embasar-se no dito italiano aqui podemos levar em consideração o trocadilho húngaro “fordítás, ferdítés” que quer dizer algo como “tradução é distorção”. Para esclarecer mais o assunto, Rónai ainda exemplifica o porquê do pensamento de que todo tradutor é traidor não faz muito sentido:

“Assim, por exemplo, a própria opinião de que o tradutor trai necessariamente a idéia do autor talvez seja devida, antes de tudo, à possibilidade no italiano, do trocadilho traduttori traditori; em qualquer outra língua, em que as duas palavras não têm forma semelhante, a idéia nasceria mais dificilmente e não teria a mesma oportunidade de generalização [...]” (Rónai, 1987, p.15).

“Tudo isso mostra, apenas, que as dificuldades do tradutor não podem ser circunscritas e também que não há problema de tradução definitivamente resolvido. **Cada palavra se apresenta, de cada vez, num contexto diferente, que a embebe de sua atmosfera e lhe altera o sentido, às vezes quase imperceivelmente.**” (Rónai, 1987, p.19, grifo nosso).

Em consonância com Rónai, podemos ver também a posição de Henrik (1993) no seu livro *As (in)fidelidades da tradução: servidões e autonomia do tradutor*:

“Parece evidente que não se pode exigir uma fidelidade àquilo que é por definição inacessível: no caso em pauta, a mensagem pretendida do emissor original. Mesmo a mensagem virtual não é diretamente acessível, mas apenas pela intermediação parcial do processo de decodificação. Assim, a matriz primária da fidelidade há de ser, por imposição dos fatos, a mensagem efetiva que o tradutor apreendeu enquanto um entre vários receptores do texto original, experiência individual e única, não-reproduzível por inteiro nem mesmo pelo próprio receptor-tradutor, em outro momento ou sob outras condições de recepção.” (Henrik, 1993, p.75).

Com tantas decisões a serem tomadas é uma árdua tarefa a de se manter invisível perante uma tradução, é provável dizer que agora com toda a tecnologia e conhecimento teórico,

a tradução de qualquer texto é de fato possível, o que se torna ligeiramente impossível é apagar o tradutor e a criação deste novo produto em relação à obra original.

Em determinadas situações o tradutor pode até tentar ficar escondido entre as linhas do texto, porém existe um tipo de tradução em que essa opção é completamente inviável - a tradução intermodal. Com a pandemia da covid-19, o País viu um recente e crescente aumento na procura de curso de Libras¹⁵; pois como um dos resultados da pandemia foram vários empregos, palestras e aulas sendo transportadas do modelo presencial para o modelo remoto. O profissional intérprete de Libras passou a ser mais requisitado para garantir a acessibilidade de todos¹⁶. É de se esperar que num país onde 5% da sua população é surda¹⁷, este idioma fosse ganhar destaque; outro elemento que fica em evidência é o tradutor e intérprete de Libras, uma vez que diferentemente do tradutor de textos escritos ou até mesmo orais, para a interpretação dos sinais o “anonimato” ou “invisibilidade” do tradutor intérprete é um fator inexistente. Por estes motivos, podemos evidenciar a importância do papel do tradutor, que agora deixa de ser esta presença ocultada pelo público e passa a ter um papel realmente central na tradução, é ele quem deve ocupar os holofotes e não o texto, as dúvidas, as decisões, as escolhas, os caminhos não tomados, tudo recai sobre o tradutor e ele tem o direito moral de reivindicar seu papel primordial na tradução.

Frequentemente é fácil encontrarmos pessoas discutindo em fóruns *online* que um texto “x” ou livro “y” está mal traduzido e falar que o tradutor não foi competente¹⁸, porém quando o contrário é verdade e um determinado texto/livro está muito bem traduzido, dificilmente você encontrará esse alguém para elogiar quem fez aquela tradução; podemos pegar clássicos da literatura internacional como *O Senhor dos Anéis*, *Harry Potter* e até o mais recente *Crepúsculo*. Quem lê estas obras dificilmente pensa que parte do sucesso é devido a uma boa tradução, e o tradutor que é a peça-chave destes “novos textos” fica “invisível”. Hoje há diversos tipos de mídia sendo consumidos mundo afora, sejam eles filmes, livros, jogos, revistas ou propagandas, tudo o que se pode imaginar é agora acessível para todos por conta da tradução, e como diz Arrojo (1992) “traduzir [...] implica [...], em primeiro lugar, reconhecer

¹⁵ <https://www.vidamaislivre.com.br/2022/09/30/libras-e-o-segundo-idioma-mais-buscado/>

¹⁶ <https://diversa.org.br/noticias/uso-de-libras-durante-a-pandemia-chamou-atencao-para-acessibilidade/#:~:text=A%20L%C3%ADngua%20Brasileira%20de%20Sinais,seja%20reconhecida%20oficialmente%20no%20Brasil.>

¹⁷ <https://g1.globo.com/rj/sul-do-rio-costa-verde/especial-publicitario/ubm/conhecimento-transforma/noticia/2020/02/12/populacao-brasileira-e-composta-por-mais-de-10-milhoes-de-pessoas-surdas.ghtml>

¹⁸ <http://dublanet.com.br/forum1/showthread.php?55540-Tradu%E7%F5es-ruins-problem%E1ticas/page3>

seu papel essencialmente ativo de produtor de significados e de representante e intérprete do autor e dos textos que traduz.”.

2.3 A psicolinguística e o processo de tradução bilingue

Apesar de o termo “psicolinguística” ter ganhado relevância em 1946 com a publicação do artigo *Language and psycholinguistics: A review* de Nicholas Pronko e começar a ser tratada como disciplina apenas em 1951 após o Seminário de Psicologia e Linguística na universidade de Cornell, há pouco menos de duas décadas de seu crescimento, o psicólogo Jacob Kantor cunhou o termo pela primeira vez em 1936 no seu livro *An Objective Psychology of Grammar*, uma das primeiras publicações da universidade de Indiana, tendo em vista que esta inaugurou seu conteúdo da “série científica” em 1935.

Kantor foi um psicólogo behaviorista que viu em seus estudos a necessidade de aproximar o campo de estudo da Psicologia com a Linguística. Kantor (1936) afirmava que a gramática era em parte um conjunto psicológico, e com isso queria estudos mais aprofundados da gramática com os métodos de pesquisas científicas utilizados na psicologia, algo que tornasse a compreensão de aquisição, aprendizado e produção da língua em algo mais concreto. Em seu livro, Kantor tentou estudar os fenômenos gramaticais do ponto de vista psicológico utilizando as atuais descobertas e avanços do campo da psicologia; é interessante notar que o autor atribui em sua obra que o estudo da “gramática psicológica” é de serventia para solucionar problemas linguísticos e gramaticais, o termo psicolinguística veio então para “criar uma abordagem unificada e nova para a comunicação da linguagem humana”¹⁹.

Por lidar tanto com a língua escrita como a língua falada, a psicolinguística não fica refém da gramática normativa tradicional. Os estudos psicolinguísticos “não têm consideração” para a padronização da língua, tendo em vista que este não é seu escopo. Dentre os diversos tópicos de pesquisa neste ramo de estudo, a seletividade ou não-seletividade da língua no acesso lexical bilingue, é um tópico que gera debate e vamos aprofundar um pouco mais por ser inerente a esta pesquisa.

¹⁹ “to create a new, unified approach to human linguistic communication.” Levelt (2013).

2.3.1 Caracterizando o bilinguismo

Quando se fala em “bilíngues”, notavelmente um pensamento leigo acerca deste termo já será de “alguém que fala duas línguas”, definição esta muito simplória porém não errada. Contudo o bilinguismo não se trata de uma constante ou uma definição binária entre monolíngue ou bilíngue, ele pode variar seu grau e características de acordo com a proficiência ou o modo em que a língua foi aprendida. Citamos anteriormente Dijkstra (2018) para enfatizar a dificuldade que é o processo tradutório, mesmo sendo capaz de ser realizado “até um certo ponto” tanto por *crianças* quanto *adultos*. Nasce dessa dicotomia criança-adulto uma categorização de bilinguismo proposta por Weinreich (1953) que o divide em 3 grupos:

- a) Bilinguismo simultâneo – quando a L1 e a L2 são aprendidas ao mesmo tempo, isto é, a criança aprende as duas línguas logo nos primeiros anos de vida.
- b) Bilinguismo sequencial – quando a L1 é aprendida pela criança nos seus primeiros anos de vida e a L2 é aprendida ainda na sua fase infantil.
- c) Bilinguismo tardio – quando a L2 só é aprendida bem mais tarde, na adolescência ou após.²⁰

Seguindo a tendência de expansão das ideias e no intuito de classificar os falantes bilíngues em classes mais diferenciadas, Hamers e Blanc (1989) estruturaram um esquema um pouco mais completo em que é levado em consideração não apenas a idade do sujeito no tocante ao aprendizado da L2 mas também outros fatores como: competência relativa, organização cognitiva, idade de aquisição, exogeneidade, status sociocultural e identidade cultural. Vejamos a seguir o quadro explicativo de cada um dos elementos abordados pelos autores.

Quadro 3 – Resumo dos critérios de dimensões do bilinguismo

DIMENSÕES	TIPO	DEFINIÇÃO
De acordo com a competência em ambas as línguas	(a) Bilinguismo balanceado	Competência na L1=L2
	(b) Bilinguismo dominante	Competência na L1 menor ou maior do que na L2

²⁰ Em concordância com Rod Ellis (1994), L1 indica a língua primeiramente aprendida pelo sujeito (língua materna), e L2 a língua aprendida após a primeira língua. Apesar de não ser uma definição proveitosa por existirem pessoas que aprendem as duas línguas de maneira simultânea (bilinguismo simultâneo) e existir ainda aquelas que trocam a L1 pela L2 no decorrer da vida, seja por motivo de trabalho ou outra mudança qualquer. Esta será a convenção utilizada para designar estes termos (L1/L2) nesse estudo.

De acordo com a organização cognitiva	(a) Bilinguismo composto		Unidade na L1 equivale à L2, isto é, apenas uma unidade conceitual
	(b) bilinguismo coordenado		Um conceito para a unidade na L1 e outro conceito para o equivalente na L2
De acordo com a idade de aquisição	(a) Bilinguismo infantil	(i) Simultâneo	L1 e L2 Adquiridas ao mesmo tempo e antes de 10 anos
		(ii) Consecutivo	L2 adquirida após a L1, porém também antes de 10 anos
	(b) Bilinguismo adolescente		L2 adquirida entre os 11 e 17 anos
	(c) Bilinguismo adulto		L2 adquirida após os 17 anos
De acordo com a presença da L2 na comunidade	(a) Bilinguismo endógeno		Presença da L2 na comunidade
	(b) Bilinguismo exógeno		Ausência da L2 na comunidade
De acordo com o status relativo das duas línguas	(a) Bilinguismo aditivo		L1 e L2 são socialmente valorizadas
	(b) Bilinguismo subtrativo		L2 é mais valorizada do que a L1
De acordo com a identidade cultural	(a) Bilinguismo bicultural		Identidade cultural positiva para a L1 e L2
	(b) Bilinguismo monocultural		Identidade cultural voltada para a L1
	(c) Bilinguismo acultural		Identidade cultural voltada para a L2
	(d) Bilinguismo descultural		Sem identidade cultural

Fonte: Adaptado de Hamers e Blanc (1989).

A esquema mostrado por Hamers e Blanc já se propõe a abordar um conjunto de definições mais focadas em tentar categorizar o bilinguismo perante uma série de fatores que não somente a idade de aquisição da L2. Vejamos em Kenji e Hakuta (2006, tradução nossa) uma designação do que pode (ou não) ser o bilinguismo:

“Os autores aqui presentes definem bilíngues como indivíduos ou grupos de pessoas que adquirem habilidades comunicativas em variados graus de proficiência na forma oral e/ou escrita para conseguir interagir com falantes de uma ou mais línguas em uma determinada sociedade. Sendo assim, bilinguismo pode ser definido como estado

psicológico e social de indivíduos ou grupos de pessoas que resulta de interações através da língua em que dois ou mais códigos linguísticos (incluindo dialetos) são usados para comunicação.” (Kenji e Hakuta, 2006, p. 115).²¹

Há também uma tabela elaborada para definir todos os graus e competências do bilinguismo (Kenji; Hakuta, 2006, p. 116-117) que aborda de maneira ainda mais exploratória as categorizações de bilinguismo. Nela os autores abordam a definição (tipologia) do bilinguismo, a dimensão, as características do aprendizado da L2, os possíveis efeitos das relações L1 e L2 e ainda os problemas relacionados e implicações educacionais; os autores abordam também uma rápida revisão sobre o que é bilinguismo sob a ótica de vários outros estudiosos (Kenji; Hakuta, 2006, p. 115-120). Os autores salientam que o bilinguismo possui um “comportamento linguístico multidimensional e muito complexo”, e afirmam também a característica transitória do bilinguismo, não sendo estático; o indivíduo, por ser alguém inserido na sociedade, pode por algum evento passar a ter que usar a sua L2 em detrimento da L1, como por exemplo uma pessoa que vá morar ou trabalhar em um local onde a L2 é mais falada que a L1. É provável que as habilidades orais na L2 com o passar do tempo superem as habilidades da L1.

Uma das perguntas que serve como força motriz para as pesquisas na área do bilinguismo é o acesso lexical de uma determinada língua ser seletivo ou não: isto é, ao encontrar um estímulo (geralmente uma palavra) a nossa mente seleciona uma língua a priori para só depois ativar a outra (acesso lexical seletivo) ou na chegada do estímulo nossa mente já ativa as duas línguas de maneira paralela (acesso lexical não seletivo)?

O presente estudo abordará o bilinguismo como a capacidade individual do sujeito em ser capaz de produzir e receber insumo em duas línguas, neste caso em particular as línguas inglesa e portuguesa. Diferentemente do proposto por Kenji e Hakuta (2006), não levaremos em consideração a capacidade operacional social da expertise nas línguas, e aqui consideraremos também a visão já bem consolidada (Grainger, 1993; Guttentag *et al.*, 1984; Hermans *et al.*, 1998; van Heuven *et al.*, 1998) da não-seletividade da língua.

Para adentrar neste jogo do bilinguismo os pesquisadores do tema elaboraram vários modelos teóricos de acesso lexical. Uma gama destes estudos é dedicada a revisar, testar, criticar ou dar suporte a tais modelos, veremos de maneira sucinta o que cada modelo propõe

²¹ The present authors define bilinguals as individuals or groups of people who obtain communicative skills, with various degrees of proficiency, in oral and/or written forms, in order to interact with speakers of one or more languages in a given society. Accordingly, bilingualism can be defined as psychological and social states of individuals or groups of people that result from interactions via language in which two or more linguistic codes (including dialects) are used for communication.

como base para suas teorias e alguns exemplos elaborados para exemplificar com a nossa língua portuguesa.

2.4 Estado da arte dos modelos de acesso lexical bilíngue

Para ter uma melhor compreensão e avançar cada vez mais nos estudos, os pesquisadores do campo do bilinguismo tentam propor modelos que simulam o processo que ocorre na mente do bilingue durante a tradução. Esta é talvez a mais árdua tarefa que enfrentamos ao realizar pesquisas com este foco, pois adentrar na mente humana e captar os microscópicos detalhes acerca de qualquer assunto é no mínimo um desafio de décadas e décadas de estudos e testes empíricos para aferir a eficácia de tais modelos.

Começaremos por citar mais uma vez Weinreich (1953) quando na ávida tentativa de classificar duma maneira mais elaborada os variados tipos de bilíngues, fez a diferença entre bilíngues coordenados, compostos e subordinados. Os bilíngues coordenados são aqueles que separam o significante e o significado para cada língua que dominam (i.e., [book - /buk/] e [kniga - /kn'iga/]) o termo “livro” é separado para cada uma das línguas e assim o falante junta (coordena/associa) cada termo independentemente entre as línguas; os bilíngues compostos fazem a diferenciação dos dois significantes porém utilizam um significado apenas (i.e., book = kniga e /buk/ ≠ /kn'iga/), o termo “livro” é compreendido como tal independente da língua em questão, a diferença está apenas na forma física e não no conceito; e os bilíngues subordinados são aqueles que aprendem uma L2 com a ajuda de outra língua previamente aprendida (i.e., book → /buk/ → /kn'iga/), fazendo assim com que o termo “livro” na L2 esteja subordinado ao significante e significado na L1. Com esse pensamento de associações entre as línguas, Weinreich plantou este desafio que paira até hoje quando o assunto é acesso lexical. Para entendermos melhor alguns dos modelos de acesso lexical veremos adiante os seguintes modelos: Modelo Hierárquico Revisado, Modelo de Traços Distribuídos, Modelo de Ativação Interativa Bilíngue, Modelo de Ativação Interativa Bilíngue Plus, e Multilink. Para fazer a seleção de modelos lexicais a serem revisados, deve-se levar em conta o propósito do estudo pois como visto em Toassi e Mota (2015) os diferentes tipos de modelos servem a estudos distintos e variam também em seus respectivos resultados:

“A literatura não oferece resposta única e conclusiva para explicar a organização e o processamento do léxico mental de bilíngues. Existem diferentes modelos e hipóteses, sob diferentes perspectivas, com o objetivo de explicar o léxico mental de bilíngues.

Além disso, estudos com foco no acesso lexical de bilíngues, na compreensão e na produção da língua, apresentam diferentes resultados.” (Toassi e Mota, 2015, p.393).

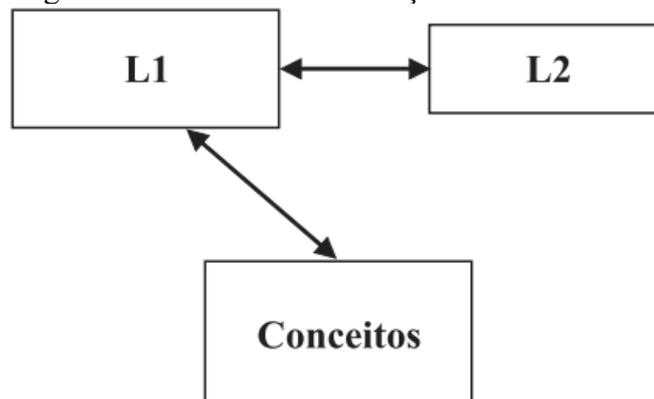
Ao levar em consideração tal afirmação e entender que este estudo pretende adentrar na seara da tradução de palavras feita pelo falante bilíngue, os modelos de acesso lexicais aqui escolhidos foram baseados numa segmentação lógica de “evolução” em que todos têm traços marcantes que se desenvolvem até chegar ao último modelo analisado, Multilink, pois é o modelo que mais se assemelha a uma tradução humana entre duas línguas distintas que possam conter características gráficas, fonéticas, e fonológicas discrepantes entre si.

2.4.1 O Modelo Hierárquico Revisado

O Modelo Hierárquico Revisado foi proposto por Kroll e Stewart (1994) como uma alternativa ou resposta para os conceitos do modelo hierárquico em Potter *et al.* (1984). O modelo previamente apresentado por Potter *et al.* tinha como uma das hipóteses a suposição de que o significado de um termo da L2 está sempre relacionado ao seu correspondente na L1. Já na segunda hipótese, os autores do modelo assumiam que os termos tanto da L1 quanto da L2 estavam ligados aos seus respectivos conceitos e não necessitavam assim de interferência entre as línguas para ter acesso às multiplicidades de representações das palavras.

Para a primeira hipótese de Potter *et al.*, o modelo previa esta dependência nata da L2 para a L1. É como se absolutamente todas as palavras que o sujeito tivesse aprendido na L2 fossem obrigatoriamente associadas a um léxico na L1, sendo assim impossível o acesso à caixa de conceitos da L2 se não por meio da tradução. Somente a L1 tem o acesso imediato à caixa dos conceitos. Observe a figura 5 para uma representação gráfica do modelo; pode-se dizer que esta hipótese faz analogia ao bilinguismo subordinado (proficiência da L2 menor que da L1).

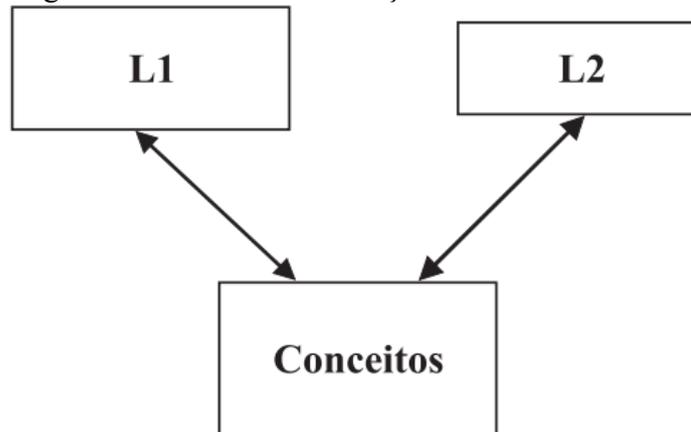
Figura 5 – Modelo de Associação de Palavras



Fonte: Adaptado de Potter *et al.* (1984).

Já a segunda hipótese dos autores contemplava o bilinguismo composto, em que o sujeito já tem um grau maior de proficiência da L2 e consegue acessar a caixa dos conceitos sem necessidade de intermediação ou tradução na L1.

Figura 6 – Modelo de Mediação de Conceito

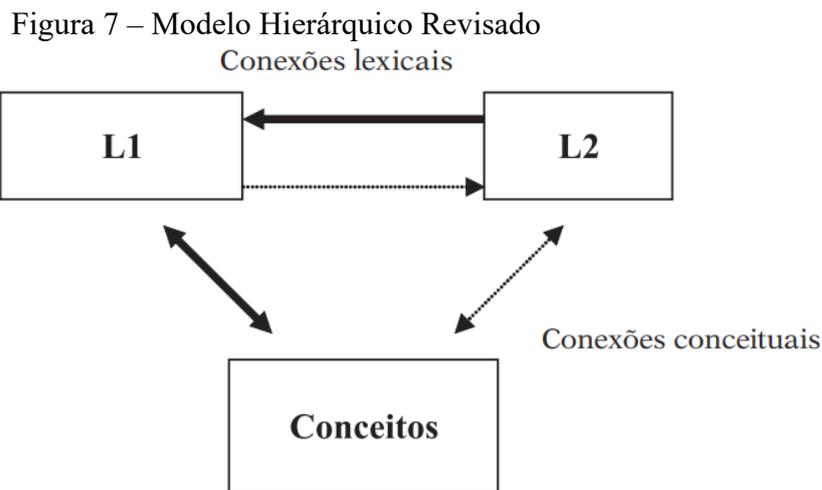


Fonte: Adaptado de Potter *et al.* (1984).

O modelo da segunda hipótese assumia, porém, um desaparecimento entre as mediações da L1 com a L2 e vice-versa, o que na prática não aconteceria. A depender da palavra utilizada como estímulo, as ligações entre as duas línguas ainda eram presentes (e.g. palavras cognatas, palavras não cognatas, substantivos concretos e substantivos abstratos, familiaridade da palavra, frequência da palavra); o modelo, então, não contemplava a diversidade do universo do vocabulário das línguas.

Contudo, Kroll e Curley (1988) e Chen e Leung (1989) primavam que os dois modelos seriam aplicados para diferentes etapas de aprendizado de uma língua. O Modelo de Associação de Palavras seria mais adequado para um aprendiz iniciante que deve ter o trabalho de traduzir toda e qualquer palavra da L2, e o Modelo de Mediação de Conceitos leva em consideração o bilíngue que já é proficiente na L2, não necessitando assim de fazer traduções do léxico da L2 para a L1; a tradução ou acesso ao léxico do bilíngue acontecia de forma assimétrica, geralmente uma das línguas “pesava” mais durante o processo tradutório (tradução direta ou busca na caixa de conceitos). Essa assimetria foi aferida em pesquisas com tarefas de tradução (De Groot; Dannenburg; Van Hell, 1994; Kroll *et al.* 1998b; Kroll *et al.* 2002; Sánchez-Casas; Davis; García-Albea, 1992; Sholl; Sankaranarayanan; Kroll, 1995), *priming* (Fox, 1996; Gollan; Forster; Frost, 1997; Jiang, 1999; Keatley; Spinks; De Gelder, 1994) e tarefas de nomeação (Kroll *et al.*, 1998b, 2002).

Foi então que, para contabilizar essas assimetrias (tradução da L2 para a L1 acontecia com mais rapidez) ou lacunas deixadas pelos modelos de Potter *et al.*, Kroll e Stewart (1994) produziram um novo modelo de acesso lexical. O Modelo Hierárquico Revisado leva em consideração a proficiência do bilingue para explicar a força das relações entre os léxicos L1-L2 (Kroll; Sholl, 1992; Kroll; Stewart, 1994) e, por conseguinte, apresenta ligações que caminham tanto da L2 para a L1 como da L1 para a L2.



Fonte: Adaptado de Kroll e Stewart (1994).

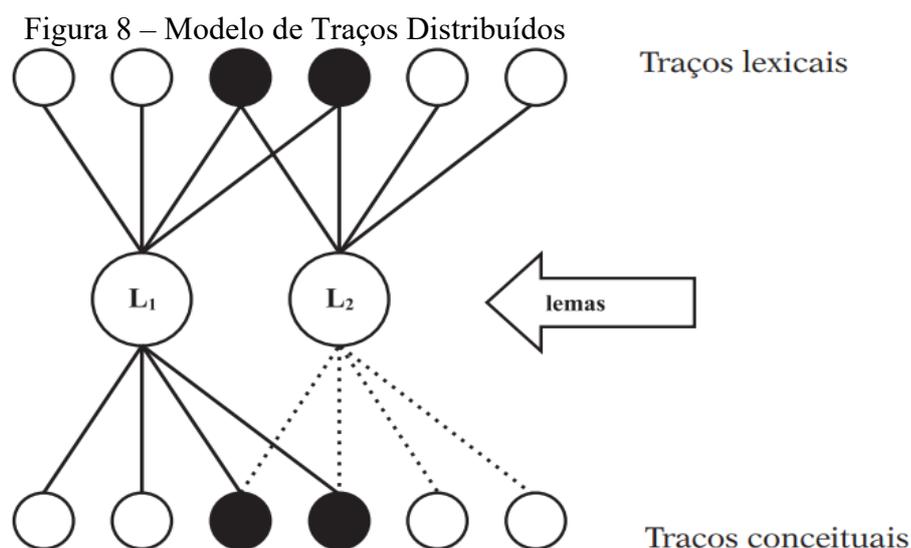
O modelo apresenta seta (ligações) mais forte que parte da L2 para a L1 por compreender que a tradução desta condição acontece de maneira mais rápida por ser o caminho mais frequente para o sujeito que está aprendendo uma L2, já a conexão da L1 para a L2 é mais fraca pois este tipo de tradução é bem menos comum durante o aprendizado da L2. O modelo apresenta também uma ligação mais forte entre os *conceitos* e a L1, explicado por esta ter sido aprendida durante a fase da infância, fortificando assim a assimilação; por levar em consideração a proficiência bilingue do sujeito, a ligação entre os *conceitos* e a L2 pode com o tempo ser fortificada e passar a se igualar ao nível da ligação *conceitos*-L1. Todo esse processo de considerar o nível de bilinguismo faz com que a aparente assimetria do modelo possa se tornar mais equilibrada, principalmente ao ter apoio contextual semântico durante a tarefa a ser executada (La Heij; Kerling; Van Der Velden, 1996).

Tanto o Modelo Hierárquico quanto o Modelo Hierárquico Revisado receberam diversas críticas a respeito da incapacidade de conseguir controlar os fatores que ordenavam as estruturas do acesso lexical do bilingue, fazendo com que os resultados obtidos fossem muito interpretativos; os fatores que acarretavam nessa quase aleatoriedade dos resultados são aspectos já conhecidos como a proficiência do sujeito, a abstração da palavra (substantivo) e

quanto a natureza cognata das palavras (Heredia, 1996; Van Hell, 1998; Fraga; Teijido; Alameda, 2002); outros pesquisadores, em um período de tempo já mais recente em relação à publicação original de Kroll e Stewart, criticaram ainda a visão seletiva de língua adotada pelo método em questão (Brysbaert; Duyck, 2010), crítica esta rebatida pelos próprios autores (Kroll *et al.*, 2010) ao considerarem que os estudos da época evidenciavam a favor pela seletividade da língua (Gerard; Scarborough, 1989) e que o ponto focal do seu modelo não era sobre a seletividade ou não-seletividade da língua e nem sobre o reconhecimento de palavras, mas sim tinha seu foco voltado para a produção de palavras e a proposição para contabilizar o desempenho na produção de traduções²². Os autores ainda pontuam diversos estudos que corroboram para a não-seletividade da língua em reconhecimento visual de palavras (Dijkstra, 2005), reconhecimento da palavra falada (Marian; Spivey, 2003) e na produção da palavra falada (Kroll; Bobb; Wodniecka, 2006) e finda citando diversos outros artigos que contribuíram para a elucidação de um novo pensamento a respeito da seletividade da língua no acesso lexical (Duyck *et al.*, 2007; Libben; Titone, 2009; Schwartz; Kroll, 2006; Van Hell; De Groot, 2008).

2.4.2 O Modelo de Traços Distribuídos

Em meio às críticas em relação ao modelo, Kroll e De Groot (1997) elaboram um novo arquétipo para explicar de maneira mais precisa as divergências que ocorrem na tradução a nível da palavra. Este modelo ficou conhecido como Modelo de Traços Distribuídos.



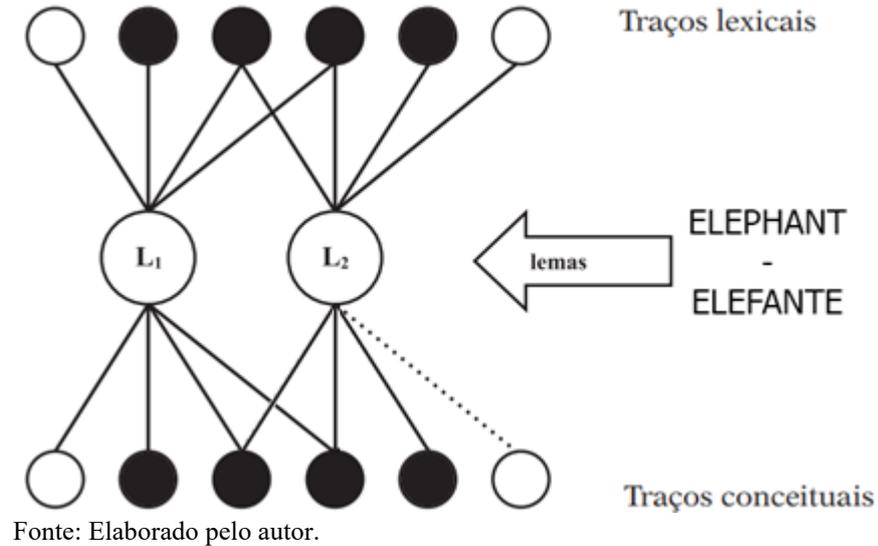
Fonte: Adaptado de Kroll e de Groot (1997).

²² the RHM is fundamentally a model of word production, proposed to account for performance in translation production. (Kroll *et al.*, 2010, p. 3)

O novo modelo trabalhava levando em consideração os pormenores de cada lema em cada língua, e assim conseguia explicar uma ativação mais rápida de substantivos concretos e palavras cognatas frente a substantivos abstratos e palavras não cognatas. A representação gráfica do modelo mostra traços contínuos para quando uma determinada característica daquele lema é facilmente percebido/ativado em uma dada língua e apresenta também traços pontilhados, indicando que tal característica não é facilmente recuperada de um outro lema. Os círculos do modelo representam as características, e os círculos preenchidos são as sobreposições que podem ocorrer entre os lemas em línguas diferentes. Estas sobreposições ajudariam um rápido acesso às informações daquela palavra; o modelo prevê que um dado lema pode partilhar tanto características lexicais quanto traços de conceito, quanto mais características são compartilhadas, mais imediata é a acessibilidade da palavra, tendo em vista que o modelo opera com base nas sobreposições entre as línguas, ele ainda tem a vantagem de contabilizar a independência ou interdependência entre as línguas em questão, tornando a visão da não-seletividade das línguas um ponto mais concreto no estudo. (Kroll; De Groot, 1997).

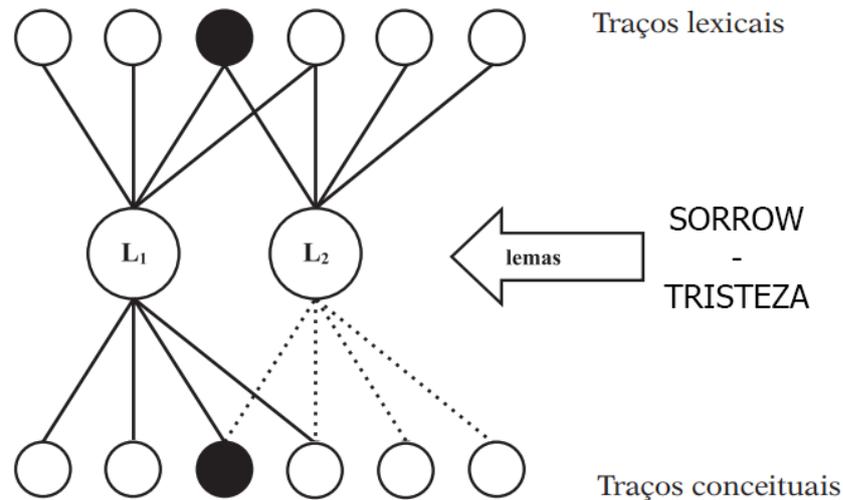
Para representar as implicações do modelo podemos pegar as palavras *elephant* e *elefante*. Por ser um substantivo concreto e ainda ter a mais a vantagem de ser uma palavra cognata entre as línguas inglesa e portuguesa, este lema teria seus círculos de características lexicais e conceituais preenchidos por representarem exatamente o mesmo objeto em ambas as línguas. No exemplo demonstrado a seguir é como se as palavras preenchessem quatro critérios de similaridade nos traços lexicais. Podemos dizer que ambas começam com *ele*, as duas possuem oito letras, têm o grupo de letras *ant* mais deslocado para o final da palavra e possuem o fonema [f] no meio da palavra, a única diferença está na grafia deste fonema, enquanto em inglês usa-se o dígrafo *PH*, em português usa-se apenas a letra *F*. Já nos traços conceituais, por se tratar de um substantivo concreto e de amplo conhecimento geral, decidimos sobrepor também quatro características para critério de comparação, ambas as palavras vão fazer referência a um animal, de grande porte, com tromba, acinzentado e inteligente.

Figura 9 – Demonstração das sobreposições entre os lemas *elephant* e *elefante*.



Se para o mesmo modelo utilizarmos as palavras *sorrow* e *tristeza* (FIGURA 10), o nível de sobreposição entre elas já decai bastante pois trata-se de um substantivo abstrato que, além do mais, não é um cognato entre as línguas, sendo assim o acesso a esse lema seria mais dificultoso e demorado. O exemplo para esta situação já se torna bem mais complicado, primeiramente que para associar traços lexicais entre as duas palavras é uma tarefa quase impossível, elas não possuem o mesmo número de letras, nem o mesmo número de vogais, a parte fonológica das duas são completamente distintas, podemos no máximo falar que, visualmente, as duas têm um tamanho parecido. Para os traços conceituais também temos dificuldade em ligar as duas palavras, um ponto em comum e bem óbvio é que ambas fazem referência a um sentimento ruim, porém por se tratar de um sentimento, cada indivíduo tem suas particularidades sobre o que é tristeza e como senti-la.

Figura 10 – Demonstração das sobreposições entre os lemas *sorrow* e *tristeza*.



Fonte: Elaborado pelo autor.

Os exemplos citados anteriormente coincidem de traços lexicais e conceituais serem ou muito próximos ou muito afastados, mas se pegarmos por exemplo o lema *nordeste* e compararmos com *northeast*, ou até mesmo *dezembro* e *december*? Ambos os lemas têm semelhança nos traços lexicais, porém na parte conceitual um tradutor ao encontrar essa palavra em um determinado contexto pode não compreender o seu significado. Vejamos um trecho do livro *Como aprendi o português e outras aventuras*, de Rónai, em que ele explica a dificuldade conceitual das palavras.

O que, porém, me atrapalhava sobretudo eram as palavras mais corriqueiras, mais simples. Os sábios glotologistas do meu café, embora com relutância, tiveram de concordar comigo quando lhes mostrei que uma das palavras brasileiras mais difíceis de se traduzir e encaixar num verso húngaro era dezembro. O nosso december, etimologicamente idêntico, mas que evocava noções de gelo, neve e miséria, não poderia sugerir a nenhum leitor húngaro a imagem de um Natal carioca, tórrido e abafado. Ou então, que significava a palavra Nordeste? Foi necessário uma longa carta de Ribeiro Couto (então cônsul na Holanda) para dar-me uma idéia aproximativa do complexo sentido geográfico, antropológico, sociológico e, sobretudo, poético, dessa denominação. (Rónai, 1956, p. 15)

É importante elucidar que todos os exemplos utilizados nesta seção foram pensados com o propósito de facilitar a explanação dos modelos. É certo salientar também que as ligações dos traços lexicais e conceituais não se limitam apenas aos seis círculos presentes nas figuras representativas do modelo, elas estão ali como um apoio visual para melhor compreensão; é possível que determinados lemas tenham 2, 5 ou 20 traços conceituais em comum, e daí teríamos que de maneira desnecessária representá-los com uma imagem demasiadamente grande.

2.4.3 O Modelo de Ativação Interativa Bilingue

Nas pesquisas mais antigas acerca da ativação das línguas de um bilingue, a concepção mais aceita era de que o sujeito não permanecia com ambas as línguas ativas ao mesmo tempo, enquanto a L1 estava em processamento, a L2 ficava em estado de dormência, esta hipótese foi confrontada em Brenders *et al.* (2010) que constatou a ativação de ambas as línguas tanto em bilíngues adultos quanto em crianças, mostrando assim que até mesmo o fator proficiência, que tem um grande peso para os modelos de acesso lexical, não foi capaz de impedir/inibir o acesso simultâneo às línguas. O modelo de Ativação Interativa Bilingue (Bilingual Interactive Access), também chamado de BIA, foi criado por Dijkstra, Van Heuven e Grainger (1998) e é essencialmente um modelo mais amadurecido por já assumir de princípio a não-seletividade da língua. Apesar de os autores assumirem a postura não-seletiva, é interessante notar que eles trazem à tona uma possibilidade do acesso seletivo, vejamos o excerto de Dijkstra, Van Heuven e Grainger (1998, p. 178, traduzido por Google, revisão e grifo nosso):

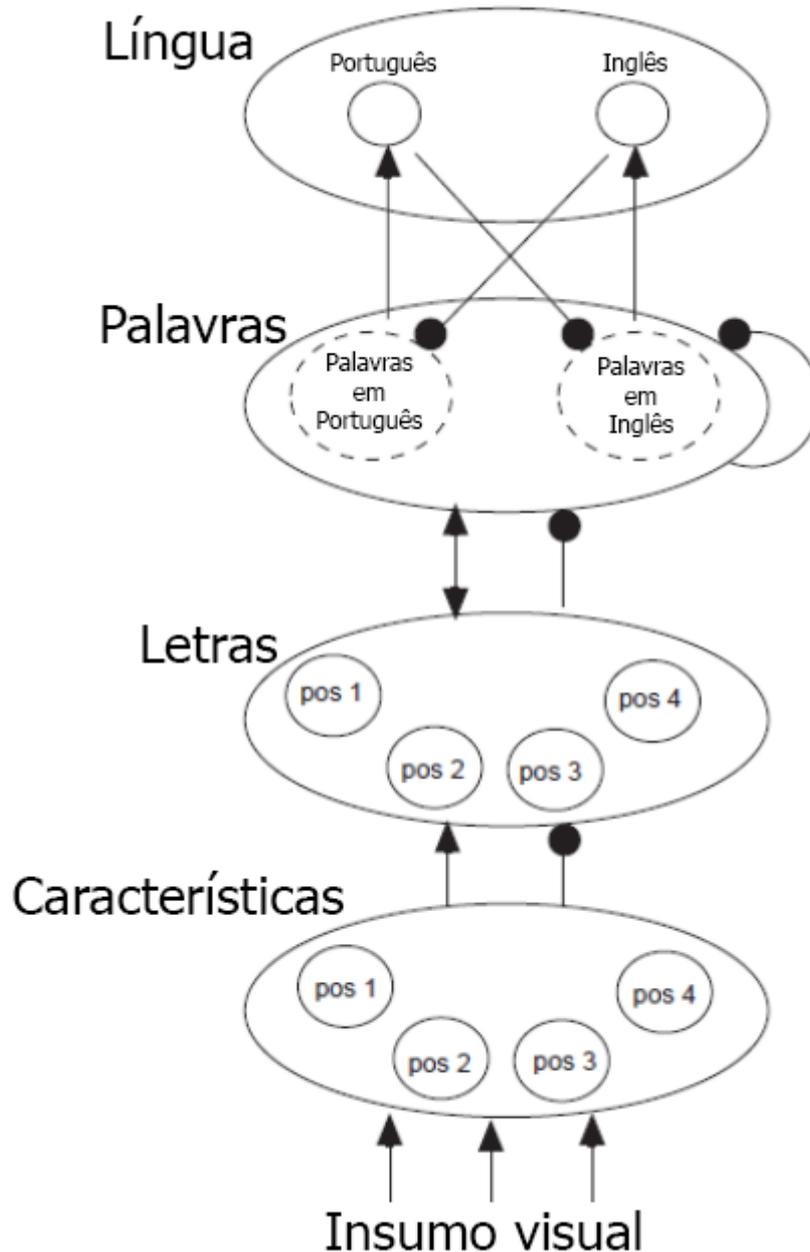
Os resultados aparentemente conflitantes na literatura experimental podem ser reconciliados assumindo que o reconhecimento de palavras bilíngues envolve basicamente um sistema de processamento não seletivo de linguagem que **pode funcionar mais ou menos seletivamente, dependendo das circunstâncias experimentais**. Em outras palavras, o estado de ativação das línguas de um bilingue não é rigidamente fixado em um ambiente monolíngue ou bilingue, mas flutua dependendo de (interações ainda desconhecidas) fatores "top-down" (demanda da tarefa, tipo de instrução e assim por diante) e fatores "bottom-up" (composição da lista de estímulos em termos de tipos e itens de linguagem, etc.)²³

Podemos notar, então, que o modelo por eles proposto leva em consideração um novo fator antes não existente. A tarefa proposta e suas instruções, além de colocar como um dos pedestais o nível de proficiência na L2 do sujeito. Uma característica marcante do BIA é que o estímulo acontece de forma visual e imediatamente quando o sujeito inicia a leitura de uma palavra. O foco do modelo está no reconhecimento da palavra pela representação ortográfica, assumindo uma ativação em paralelo entre as duas línguas. O léxico do bilingue no BIA é tratado de maneira integrada (i.e., todas as palavras estão competindo entre si e existem “ao mesmo tempo” nas duas línguas), porém o modelo possui *nós/nódulos* para diferenciar

²³ The seemingly conflicting results in the experimental literature can be reconciled by assuming that bilingual word recognition basically involves a language-nonselective processing system which may function more or less selectively depending on experimental circumstances. In other words, the state of activation of a bilingual's languages is not rigidly fixed in either monolingual or a bilingual setting, but fluctuates depending on (as yet unknown interactions between) "top-down" factors (task demand, type of instruction, and so on) and "bottom-up" factors (stimulus list composition in terms of types and language items, etcetera)

entre a L1 e L2. Outro ponto crucial do BIA é que ele é teoricamente uma adaptação do modelo de Ativação Interativa (IA) proposto por McClelland & Rumelhart (1981) com a adição de uma camada extra para as línguas por ser voltado ao indivíduo bilíngue em contrapartida ao modelo monolíngue adotado no modelo IA.

Figura 11 – Modelo de Ativação Interativa Bilíngue (BIA)



Fonte: Adaptado de Dijkstra, Van Heuven e Grainger (1998).

A execução do modelo se inicia no *input* (insuno) visual: os traços com seta na ponta indicam o caminho da ativação, já os traços com uma bola na ponta mostram uma inibição durante aquela determinada etapa do processo. Notadamente, o modelo BIA foi planejado em

sua origem para palavras em inglês e holandês de apenas quatro letras. Vamos respeitar então o modelo e exemplificar com palavras em português de mesmo tamanho. Tomando como exemplo a palavra “ALMA”, a palavra em si é o nosso insumo visual: ao visualizar os caracteres da palavra o sistema das características entra em operação. Esse sistema de caracterização visual leva em conta a disposição visual dos traços que formam os caracteres de cada letra: a letra “L” em “ALMA”, por exemplo, é formada por um traço reto na vertical que se estende por três posições (superior, mediana e inferior) “|” e um traço reto na horizontal que ocupa apenas a posição inferior “_”. O modelo BIA trabalha com enfoque na concepção de inibição, então sempre que uma característica de qualquer uma das camadas é ativada; outra característica que competiria com esta passa a ser inibida.

Após a ativação das características visuais da letra o nó passa para a etapa das letras propriamente ditas. Neste nó o que será ativado é a posição que cada letra ocupada na palavra, como o modelo fora alimentado apenas com palavras de quatro letras, então aqui as posições são apenas quatro; para o nosso exemplo “ALMA” teríamos a letra “A” ocupando duas posições, a primeira e a última (quarta) enquanto a segunda posição é ocupada pela letra “L” e a terceira está ocupada por um “M”. Aqui a inibição ocorre de maneira óbvia, é impossível duas letras ocuparem a mesma posição, caso contrário a palavra ficaria ilegível, uma vez que todas as letras foram observadas em suas posições, o nó dá mais um salto para a caixa das palavras.

Ao chegar na fase de palavras, o sistema opera em conjunto com o nó seguinte (Língua) com a decisão binária de escolha da língua, aferindo então se a palavra é pertencente ao vocabulário da L1 ou da L2. No exemplo citado o sistema escolheria a palavra “ALMA” como sendo uma palavra em português, o que já sobe para o nó da língua a ativação da língua Portuguesa, o que por consequência já parte para inibir as palavras em inglês. Há também uma conexão da caixa de palavras com a caixa das características, uma vez que a palavra pode ter vizinhos ortográficos que competem visualmente com a palavra alvo; no nosso exemplo temos competidores para a palavra “ALMA” as palavras “ALVA”, “ALTA”, “ALÇA”, “ALGA”, “ALFA”, “ALBA”, “ALNA”, “ALHA”, e a única diferença ortográfica entre elas está na letra da terceira posição, todas as palavras seguem o padrão “AL_A”, quanto mais semelhanças visuais os competidores (vizinhos) tiverem, mais eles tentarão competir com a palavra alvo, por outro lado, vizinhos um pouco mais distantes como “ALÉM” ou “ALGO” que compartilham apenas as duas primeiras posições, seriam menos ativados.

Uma crítica ou observação pessoal ao modelo BIA diz respeito à classificação dos aspectos visuais presentes no nó das características. O modelo leva em consideração um tipo de fonte ortográfica que é pouco usual e que mais se assemelha a um visor de calculadora, todas

as características presentes no modelo são representadas por uma combinação de 16 traços, horizontal, vertical, diagonal, superior, mediano ou inferior como na figura 12.

Figura 12 – Modelo de fonte ortográfica e traços de características visuais do BIA



Fonte: McClelland e Rumelhart (1981).

Analisando a fonte utilizada, fica pouco provável que o modelo funcione com determinadas palavras de outros idiomas, mesmo o idioma brasileiro poderia ser prejudicado caso o insumo visual fosse alguma palavra que contenha “Ç” como “CAÇA” pois o indicador da cedilha é um traço visual que ocupa um espaço mais abaixo do espaço dito como “inferior” pelo BIA, isso sem contar com os nossos outros sinais diacríticos (´ [agudo], ` [grave], ^ [circunflexo], ¨ [trema] e ~ [til]) que aparecem em posição mais acima da posição superior vista no modelo como na palavra “SOFÁ” ou uma bem marcada de diacríticos como “AÇAÍ”. Portanto para que o nó das características funcione como esperado ele deve ser remodelado para comportar todos estes aspectos visuais das línguas alvo em questão.

O modelo BIA conseguiu representar de uma maneira mais elaborada o acesso lexical do bilíngue, porém críticas ao modelo surgiram por conta de determinados aspectos como:

- a) representações fonológicas e características semânticas das palavras não apresentarem importância alguma para o modelo;
- b) palavras cognatas e homógrafos intralinguísticos também sofriam com falta de precisão em suas representações;
- c) os aspectos representacionais e funcionais em relação aos nós de linguagem são confusos;

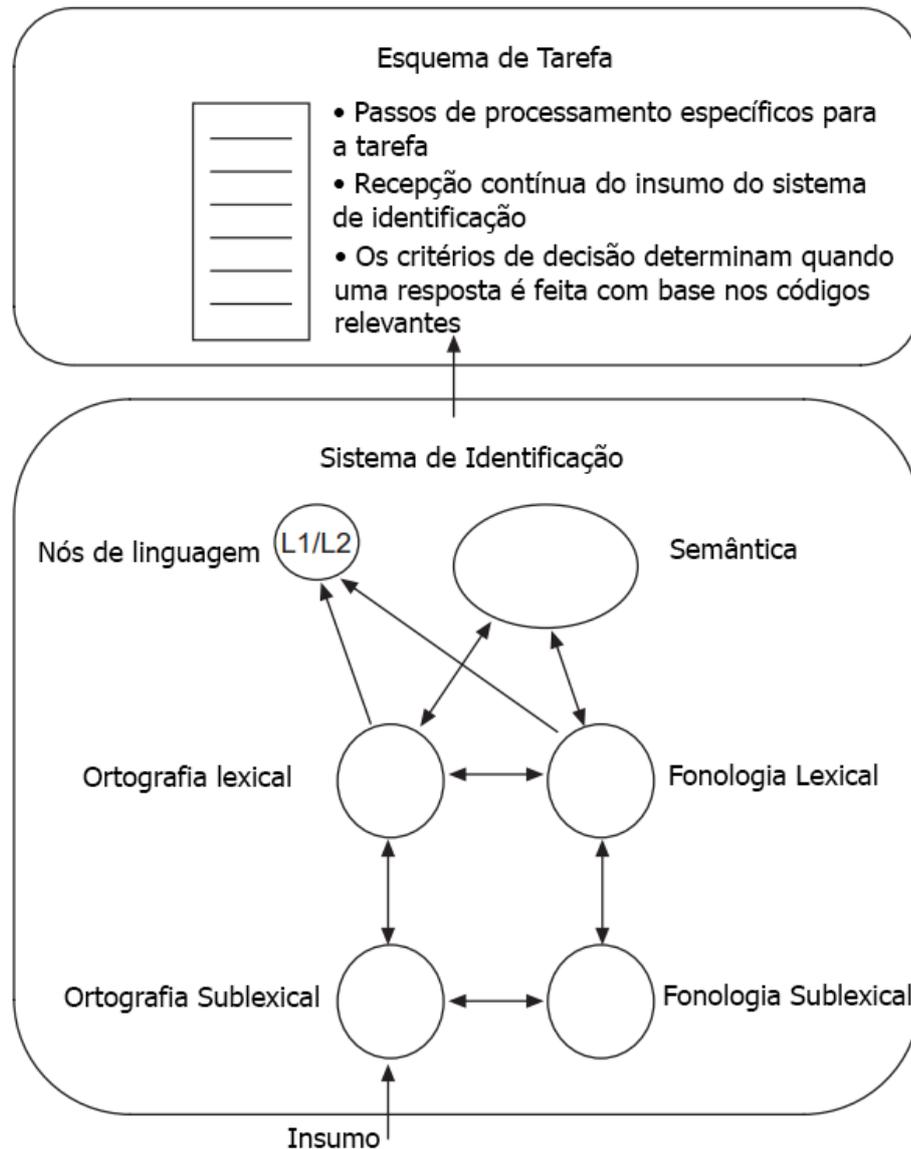
- d) há apenas relatos limitados de como os contextos não linguísticos e linguísticos afetam o reconhecimento bilíngue de palavras;
- e) não há descrição detalhada de como os participantes executam uma tarefa específica, por exemplo, decisão lexical;
- f) a relação entre identificação de palavras e as demandas da tarefa são pouco específicas.

Com base nas críticas, Dijkstra e Van Heuven (2002) criaram um novo modelo para tentar preencher os vazios deixados pelo BIA, surgiu então o *Bilingual Interactive Activation Plus* (BIA+).

2.4.4 O modelo de ativação interativa bilíngue plus (BIA+)

Dijkstra e Van Heuven trabalham em seu novo modelo comparando-o com o anterior, explicitando as similaridades e divergências existentes entre eles. O BIA+, assim como o BIA, leva em consideração a tarefa a ser executada, porém, um subconjunto novo e exclusivo é criado para comportar o esquema das tarefas, fazendo com que a identificação das palavras em si não seja afetada pelo componente das tarefas ou por fontes de informações não-linguísticas. Para discorrer sobre as características do BIA+, vejamos a seguir a sua representação visual.

Figura 13 – Modelo de Ativação Interativa Bilíngue Revisado (BIA+)



Fonte: Adaptado de Dijkstra e Van Heuven (2002).

A priori, podemos identificar a clara divisão entre o Esquema de Tarefa e o Sistema de Identificação, além disso, o modelo deixa embutido também um funcionamento puramente *Bottom-up*, isto é, “de baixo para cima”, pois a tarefa a ser realizada só poderá ser pensada após a identificação da(s) palavra(s), desta maneira o Esquema de Tarefa não tem influência alguma sobre o Sistema de Identificação (Van Heuven; Dijkstra, 2010). A outra adição tão bem aceita fora a adequação dos círculos que levam agora em consideração a fonologia e a semântica, onde anteriormente só era levado em consideração a ortografia. Ainda é notório também a exclusão do componente das “características” das letras, o que abre um leque maior de possibilidades para que o modelo possa ser utilizado em línguas variadas.

Outra característica pertinente ao BIA+ é o não retrocesso ao chegar no “nó da linguagem” onde processamento da escolha de língua da palavra é obtido de forma final, em contrapartida seus outros “nós” são todos intercambiáveis, mostrando assim seu aspecto mais bidirecional de processar as línguas.

Para se ter uma ideia melhor do processamento feito pelo BIA+ podemos pegar um conjunto de palavras que o BIA claudicava ao tê-las como insumo, vejamos por exemplo um homógrafo interlinguístico do par Inglês-Português como a palavra “TEAR”. Em inglês a palavra é polissêmica e tem sua pronúncia variada de acordo com o significado, ela pode ser pronunciada como [tɪər] ao se referir à “lágrima” ou como [tɛr] caso esteja fazendo referência para a forma verbal de “rasgar” enquanto que a palavra em português refere-se a um substantivo e tem como pronúncia [tɛ'ar], partindo do pressuposto de que o modelo BIA não leva em consideração a fonologia e nem mesmo a semântica, o modelo entraria em um *loop* infinito para tentar descobrir se a palavra TEAR está em inglês ou em português.

Já no modelo revisado (BIA+) a decisão poderia ser facilmente contornada pelos níveis fonológicos ou pelo nível semântico, é importante deixar claro que por se tratar de homógrafos interlinguísticos, qualquer uma das línguas escolhidas (inglês ou português) estaria correta caso a tarefa fosse apenas uma decisão arbitrária de idiomas em palavras descontextualizadas, porém ao contextualizar ou inseri-las em uma frase, o modelo conseguiria suprir suas carências e responder com uma melhor precisão a qual língua aquela palavra pertence. Os componentes de Ortografia Sublexical (primeira ativação após o insumo) e Fonologia Sublexical (segunda ativação em conjunto com a Ortografia lexical) inseridos aqui já trabalham de modo a caracterizar língua X ou Y, tal como se uma palavra tem muitas consoantes a probabilidade de ser inglês ao invés de português aumenta, da mesma maneira que um som mais nasalado terá uma grande chance de ser português e não inglês; a ortografia sublexical pode ser pensada de maneira análoga ao velho conhecimento de sempre falar “letra ‘A’” ao jogar o famoso “jogo da força” é algo que o sujeito adquire e internaliza em cada língua. Deste modo, o Sistema de Identificação opera de maneira autônoma em relação ao Esquema de Tarefa, mas a tarefa e o contexto podem sim influenciar o resultado final do processamento das palavras em se tratando de performance do sujeito.

O novo modelo é, portanto, endossado por diversos outros estudos como uma alternativa mais abrangente e capaz de lidar com pormenores que antes escapavam das análises, podemos observar um estudo recente que utilizou o BIA+ no processamento de palavras tanto em bilíngues unimodais (e.g., Português-Inglês) quanto em bilíngues bimodais (e.g. Libras-

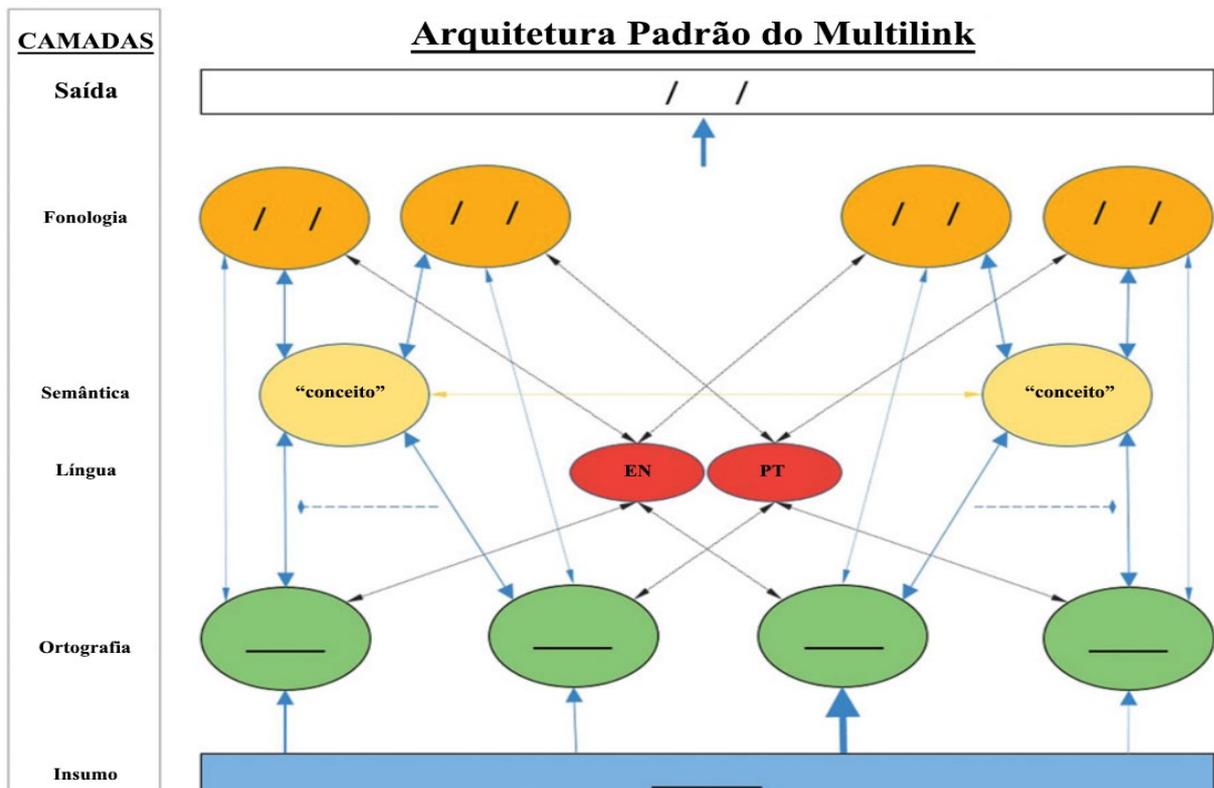
Português) onde o BIA+ demonstrou de maneira favorável os efeitos da influência interlinguística e da alternância de código linguístico *code-switching* (Wang *et al.*, 2022).

2.4.5 O Modelo Multilink

Dijkstra, novamente, cria um modelo computacional, desta vez com um nome totalmente novo pois o modelo em questão não é mais um “aperfeiçoamento” do BIA/BIA+ e sim uma estrutura completamente autêntica que captura a essência do Modelo Hierárquico Revisado (Kroll; Stewart, 1994) e também do próprio BIA+ (Dijkstra; Van Heuven, 2002) para dar vida a um arquétipo capaz de lidar tanto com o reconhecimento quanto a produção e tradução de palavras (Dijkstra *et al.*, 2018). O avanço do novo modelo pode ser percebido também na quebra do famoso limite de “quatro a cinco letras” para as palavras selecionadas como estímulo, podendo agora trabalhar com palavras de maior abrangência espacial com limites mínimos e máximo ampliados (entre três e oito letras), porém o alcance de línguas comportado pelo sistema continua o mesmo, o par Holandês e Inglês.

Ao ver a representação visual do Multilink (FIGURA 14), nós podemos ter uma melhor noção da complexidade adicionada pelo novo modelo e uma ideia de suas possibilidades.

Figura 14 – Modelo Multilink



Fonte: Adaptado de Dijkstra *et al.* (2018).

Nesta representação o insumo está figurado como o sublinhado azul mais abaixo, a representação ortográfica como o sublinhado verde e as representações fonológicas por barras “ / / ”. Na região mais central do modelo estão as línguas EN (inglês) e PT (português). A linha tracejada entre duas conexões da *ortografia* à *semântica* indica que a sua ativação é somada após tomar metade do insumo de ativação do segundo nó. A saída é dependente da tarefa e as barras indicam uma saída fonológica em línguas iguais ou diferentes (para nomeação de palavras ou tradução). Podemos notar também a gigantesca possibilidade de caminhos, depois que o insumo chega à representação ortográfica as setas passam a percorrer pelas mais diversas camadas, podendo ser ligadas diretamente à língua, conceito (semântica) ou fonologia, a etapa final do Multilink passa a não ser mais a “língua” em si pois como dito anteriormente, a tarefa é quem vai ditar as regras da “saída” que é nó final do modelo.

Como os autores descrevem (Dijkstra *et al.*, 2018, p. 5-6, traduzido por Google, revisão nossa), o Multilink ampliou o escopo das pesquisas pois ele pode:

Primeiro, o Multilink permite a simulação de fala monolíngue e processamento bilíngue de palavras que variam em frequência de uso, extensão e semelhança entre idiomas. [...] Em segundo lugar, por incluir um sistema de tarefa/decisão, o modelo permite simular processamento de texto em tarefas psicolinguísticas, como decisão lexical, *priming* ortográfico e semântico, nomeação de palavras e de tradução de palavras. [...] Terceiro, por ter a capacidade de ajuste perfeito para a proficiência na L2 do seu léxico e parâmetros de configurações, o Multilink pode simular a performance de bilíngues com muita e pouca proficiência nestas tarefas.²⁴

Por fim, os autores performam uma série de testes para validar o novo modelo mediante os modelos prévios, e, ao final, constatam que o Multilink pode ser a base para um novo paradigma neste campo experimental, principalmente por comportar múltiplos (multi) fatores de interação/ligações (link) entre os mais variados *nós* (camadas).

Tão qual a progressão do BIA para o BIA+, o Multilink é o passo natural na cadeia evolutiva de crescimento dos modelos teóricos/empíricos de processamento do léxico bilíngue. Neste capítulo abordamos alguns dos modelos de acesso lexical, alguns outros modelos incluem o Modelo para Compreensão da Fala Bilíngue ou BLINCS (Bilingual Language Interaction Network for Comprehension of Speech), o Modelo Interativo Bilíngue de Acesso Lexical ou BIMOLA (Bilingual Interactive Model of Lexical Access) e inclusive uma outra variante do

²⁴ First, Multilink allows the simulation of monolingual and bilingual processing of words that vary in frequency of usage, length, and cross-linguistic similarity. [...] Second, because the model includes a task/decision system, it allows simulating word processing in psycholinguistic tasks such as lexical decision, orthographic and semantic priming, word naming, and word translation production. [...] Third, because Multilink’s lexicon and parameter settings can be fine-tuned to L2-proficiency, it can simulate performance of both high and low L2-proficiency bilinguals in these tasks.

BIA, o modelo de Ativação Interativa Bilingue Desenvolvidor ou BIA-d (Developmental Bilingual Interactive-Activation).

A ênfase na demonstração dos tipos de acesso lexical ficou para escola de modelos criados por Ton Dijkstra e colaboradores, como o presente estudo abordará uma tarefa de tradução do par linguístico Inglês para Português, logo era mais sensato escrutinar o autor e sua linha de raciocínio até a chegada do seu então último modelo (Multilink), pois é este que nos traz uma proposta mais adequada para a tarefa de tradução proposta por não conter tantos limites linguísticos/espaciais quanto os demais modelos de acesso lexical bilíngue.

2.5 Homônimos e o processamento de línguas

Uma classe de palavras que despertou a atenção para este estudo foram os homônimos, que podem ser classificados em homófonos e homógrafos. Os homófonos são palavras que possuem a mesma fonética, porém têm diferenças ortográficas e semânticas, já os homógrafos são palavras que possuem a mesma ortografia, porém com diferenças fonéticas e semânticas, vide alguns pares nas línguas portuguesa e inglesa no quadro 4:

Quadro 4 – Homófonos e homógrafos do par inglês/português brasileiro

	PALAVRAS (ORTOGRAFIA)	TRANSCRIÇÃO (FONÉTICA)	SIGNIFICADO (SEMÂNTICA)
H O M Ó F O N O S	Sem	['sễ]	Ausência, nada.
	Cem	['sễ]	Centena, numeral.
	Acento	[a'sễtu]	Diacrítico, gramática
	Assento	[a'sễtu]	Cadeira, banco.
	Auto	['awtu]	Si próprio.
	Alto	['awtu]	Grande, tamanho.
	<i>Hear</i>	[hɪr]	Escutar, ouvir.
	<i>Here</i>	[hɪr]	Aqui, lugar.
	<i>Our</i>	['aʊər]	Nosso (pronome).
	<i>Hour</i>	['aʊər]	Hora, tempo.
	<i>Wait</i>	[weɪt]	Esperar, verbo.
	<i>Weight</i>	[weɪt]	Peso, substantivo.
	Boto	['botʊ]	Animal, golfinho.

H	Boto	['bɔtɔ]	Colocar, verbo.
O	Gosto	['gɔstɔ]	Sabor, substantivo.
M	Gosto	['gɔstɔ]	Apreciar, verbo.
Ó	Sede	['sedʒɪ]	Vontade de beber.
G	Sede	['sedʒɪ]	Matriz, substantivo.
R	<i>Lead</i>	[li:d]	Liderar, verbo.
A	<i>Lead</i>	[lɛd]	Chumbo (metal), substantivo
F	<i>Live</i>	[lɪv]	Viver, morar, verbo.
O	<i>Live</i>	[laɪv]	“Ao vivo”, adjetivo.
S	<i>Project</i>	['prɒdʒekt]	Projeto, substantivo.
	<i>Project</i>	[prə'dʒekt]	Projetar, arremessar, verbo.

Fonte: elaborada pelo autor.

Como pode ser observado, a semelhança entre os homógrafos e homófonos é que ambos distinguem na semântica, uma outra observação que pode ser feita acerca do quadro 4 é quanto à variante da língua utilizada para cada uma das línguas escolhidas, para a língua portuguesa fora escolhido o português brasileiro por motivo desta pesquisa ser conduzida no Brasil e para a língua inglesa o inglês americano por ser mais difundida no País do que a variante britânica. Ao usar homófonos é de suma importância explicitar a variante da língua a ser utilizada, pois como o objeto de estudo será focado na sonoridade, qualquer diferença poderá acarretar em ruído na observação dos dados.

Algumas das pesquisas em homônimos (Dijkstra; Van Jaarsveld; Ten Brinke, 1998; Dijkstra; Grainger; Van Heuven, 1999; Dijkstra et al., 2000; Schulpen et al., 2003; Kerkhofs et al., 2006; Smits et al., 2006; e Poort; Warren; Rodd, 2016) são partes ativas da evolução dos modelos de acesso lexical, pois as concepções dos modelos foram testadas de forma empírica e as considerações dos resultados acarretaram na elaboração de modelos cada vez mais avançados e que respondiam cada vez mais às lacunas deixadas pelos anteriores. Em suas pesquisas, Dijkstra chegou também a considerar a similaridade fonética entre os homógrafos interlinguísticos durante a realização de uma tarefa (Dijkstra; Grainger; Van Heuven, 1999), os resultados desta pesquisa concordam com a necessidade de considerar o acesso lexical não-seletivo e também julgam a similaridade fonética como um fator importante em uma tarefa de decisão lexical, uma vez que essa similaridade apresentou efeito inibitório.

Ao focar nos homógrafos interlinguísticos, podemos destacar resultados prejudiciais no estudo de Smits et al. (2006) em que um grupo de 34 participantes (condição 1)

e 24 participantes (condição 2) bilíngues em inglês e holandês tiveram que nomear várias palavras em inglês ou holandês, a depender da condição. Os resultados mostraram que tanto a acurácia como o tempo de resposta foram prejudicados quando a condição de homógrafos interlinguísticos era averiguada. Já em Dijkstra, Van Jaarsveld e Brinke (1998) os resultados achados em uma tarefa de decisão lexical mostraram respostas dependentes da condição e tipo da tarefa, podendo os homógrafos interlinguísticos apresentarem resultado nulo em relação às palavras controles, resultados inibitórios ou até mesmo resultados facilitadores, o que fora justificado pela lista e frequência das palavras utilizadas no experimento. É possível observar também resultados antagônicos em Elston-Güttler, Gunter e Kotz (2005) em que 48 participantes foram divididos em dois grupos de quantidades idênticas e o grupo que recebeu o priming na L2 (inglês) não apresentou distúrbios ou atrasos nos tempos de reação em tarefa de decisão lexical na L2, enquanto o grupo que teve o priming na L1 (alemão) apresentou maiores tempos de reação para a mesma tarefa, o estudo corroborou então com a teoria proposta pelo BIA+ e seu *schema* de tarefas que atua junto ao processamento das línguas.

Como apresentado neste capítulo, a tradução está presente na nossa sociedade há bastante tempo e é um processo essencial para o mundo contemporâneo. Muitas pesquisas são realizadas no intuito de compreender a mente daquele que traduz, seja de maneira receptiva ou produtora, então abordamos aqui também os vários modelos de acesso lexical bilíngue, percorrendo um matiz que nos leva à investigação do proposto pelo estudo: averiguar a eficiência na tradução de homógrafos interlinguísticos do par inglês/português. No capítulo a seguir discorrerei de maneira meticulosa toda metodologia da pesquisa e os processos para a sua realização.

3 METODOLOGIA

Neste capítulo, abordarei com a maior precisão possível de detalhes todas as etapas que concernem a pesquisa elaborada para o estudo, com o intuito de deixar o desenho da pesquisa de maneira facilmente replicável e discutível a fim de se tornar um estudo com possibilidades abrangentes para pesquisas futuras. Apresentarei a priori os objetivos geral e específicos do estudo e posteriormente a montagem da pesquisa, apresentando o corpus de palavras e como fora feito a sua seleção, os participantes da pesquisa e os demais dados relevantes para este estudo, finalizando então com a metodologia para o tratamento dos dados.

3.1 Objetivos

3.1.1 Objetivo geral

Este estudo tem como objetivo explorar os resultados do processamento dos dados (acurácia e tempo de resposta) aferidos durante uma tarefa de tradução de homógrafos interlinguísticos feitos por um grupo de participantes bilíngues (português como L1 e inglês como L2) a partir da ótica psicolinguística.

3.1.2 Objetivos específicos

- 1) Averiguar se os homógrafos interlinguísticos causam efeito de interferência na acurácia e tempo de reação da tradução destes do idioma inglês para o português brasileiro.
- 2) Averiguar se os homógrafos interlinguísticos causam um efeito facilitador na acurácia e tempo de reação da tradução destes do idioma inglês para o português brasileiro.

3.2 Perguntas da pesquisa

Para delinear bem o escopo da pesquisa podemos basear o estudo nas seguintes perguntas: “Os homógrafos interlinguísticos do par inglês/português brasileiro causam, na acurácia, efeito facilitador ou de interferência ao serem traduzidos?” e “O tempo de reação sofre algum efeito de interferência por conta dos homógrafos interlinguísticos?”.

3.3 Hipóteses

H1 – Os homógrafos interlinguísticos causam um efeito de interferência para a tradução das palavras. É esperado que os homógrafos interlinguísticos causem um efeito de interferência e a acurácia seja prejudicada, além do tempo de resposta. Assim como fora demonstrado em Macizo; Bajo e Cruz martín (2010), em que um estudo feito com as línguas inglesa e espanhola os participantes mostraram efeito inibitório para a compreensão de homógrafos interlinguísticos, levando então mais tempo para decidir os significados destas palavras (Macizo; Bajo; Cruz Martín, 2010).

H2 – Os homógrafos interlinguísticos causam um efeito facilitador para a tradução das palavras, apresentando uma maior acurácia e menor tempo de resposta. Ao fazer uma analogia com o “efeito de facilitação do cognato” que sustenta as palavras cognatas como sendo mais rapidamente processadas pelo falante bilíngue (Costa; Santesteban; Caño, 2005; Van Assche, et al., 2009; Gollan et al., 2007; Duñabeitia; Perea; Carreiras, 2010) em testes orais e escritos tanto de leitura quanto nomeação e tradução, esta segunda hipótese vai na contramão do esperado pois a ideia análoga é que os homógrafos interlinguísticos, diferentemente dos cognatos, causem um efeito inibitório.

3.4 Procedimentos da pesquisa

3.4.1 Desenho da pesquisa

Nesta subseção um breve detalhamento da pesquisa será comentado para servir de resumo aos tópicos seguintes que detalharão com todo o rigor necessário o passo a passo de todo o procedimento observado e elaborado durante cada fase do estudo, a escolha das palavras, desenho da tarefa, coleta de dados e processamento dos dados.

A escolha de homógrafos interlinguísticos é um tópico intrigante a ser escolhido, pois estudos mostram resultados conflituosos a respeito dos seus efeitos no processamento de palavras. Como visto anteriormente, alguns estudos observaram efeito de interferência (Smits *et al.*, 2006; Von studnitz; Green, 2002) enquanto outros mostram efeitos benéficos a depender da condição (Dijkstra; Van Jaarsveld; Ten brinke, 1998) além de resultados mistos (Elston-Güttler; Gunter; Kotz, 2005). Um ponto que chama a atenção nestes estudos é a possibilidade

dos resultados com homógrafos interlinguísticos terem a característica de variar dependendo da tarefa a ser feita ou da lista de estímulos (Dijkstra, 2009).

A pesquisa foi conduzida em ambiente *online* com o uso do *software* Psytoolkit (Stoet, 2010, 2017). O *software* citado permite a elaboração, aplicação e coleta de dados dos participantes, tudo isso feito de maneira integralmente remota, permitindo assim uma maior autonomia para que os indivíduos realizem a tarefa em momento oportuno. Para contextualizar e validar o *software* com outro já conhecido em pesquisas acadêmicas, E-prime 3.0, o Psytoolkit (Stoet, 2010, 2017) foi testado e avaliado em estudos comparativos e mostrou resultados com dados de tempo de reação e acurácia condizentes com os resultados do E-prime 3.0.

A pesquisa é dividida em duas etapas, na primeira os participantes respondem perguntas sobre seu perfil pessoal, o nível autodeclarado de língua inglesa e como ocorreu o aprendizado da língua por eles. Na segunda parte os participantes traduzem as palavras do estímulo (homógrafos interlinguísticos inglês-português brasileiro) contando com uma primeira fase de treinamento para o reconhecimento da atividade. Durante a fase de treinamento são apresentadas dez palavras, cinco do estímulo e cinco controles, as palavras-controle da fase de treinamento serão palavras comumente utilizadas na língua inglesa para caso o participante cometer falhas na tradução destas palavras a sua análise seja individualizada para evitar distorções estatísticas.

Foram escolhidas 3 variáveis preditoras para a escolha dos estímulos, a primeira sendo 40 homógrafos interlinguísticos do par Inglês/Português que foram selecionados como o estímulo experimental. A segunda variável (palavras controle) além de servir como artifício para evitar qualquer tipo de aprendizado durante a realização da tarefa de tradução, é composta por 40 palavras controle moderadas pela frequência do conjunto experimental (palavras em inglês que não são homógrafos interlinguísticos e nem cognatas) foram inseridas na tarefa de tradução. Além destes, um terceiro grupo fora inserido também com palavras controles, porém relacionados à frequência da tradução dos homógrafos em português do conjunto experimental.

Ao terminar a parte experimental da pesquisa os participantes deveriam também realizar um teste produtivo de vocabulário em língua inglesa²⁵. O teste consiste em completar uma palavra inserida em um contexto (frase) que contém apenas a sua inicial. Esta última etapa entra no conjunto da pesquisa com a finalidade de filtrar participantes que apresentassem valores atípicos.

²⁵ https://itt-leipzig.de/static/vltenglish_01.2p/index.html

Vejamos agora o detalhar de cada etapa ocorrida durante a elaboração e execução desta pesquisa.

3.5 Participantes

O estudo teve como público total o número de 26 participantes, falantes bilíngues de português brasileiro (L1) e inglês (L2), o convite para participar da pesquisa foi feito por meio de mensagem institucional (SIGAA-UFC) com o *link* para a realização das tarefas e uma breve mensagem falando sobre o estudo. No decorrer da pesquisa os dados de 3 participantes foram perdidos por motivos não esclarecidos, provavelmente por a tarefa ser realizada *online* algum tipo de interrupção na *internet*, ou até mesmo falhas entre o servidor da coleta de dados pode ter acarretado nessa perda, além destes, um outro participante teve que ser removido por apresentar respostas incompatíveis com o padrão dos demais, pois quase todas as respostas aproximavam-se do limite de tempo estabelecido e a margem de acertos foi de quase 100%, resultando assim em um número final de 22 participantes. Os participantes deveriam obrigatoriamente possuir idade de 18 anos ou mais e concordar com Termo de Consentimento Livre e Esclarecido (TCLE) para participar da pesquisa. Este estudo foi também aprovado pelo Comitê de Ética e Pesquisa (CEP) com o Certificado de Apresentação de Apreciação Ética número 73514223.8.0000.5054, cumprindo assim com todos os procedimentos legais para uma realização segura aos participantes envolvidos. Nenhuma forma de benefício monetário foi dada aos participantes, porém os mesmos ganharam certificado de participação para utilizar como horas complementares em seus devidos cursos. Vejamos a seguir os quadros com a distribuição das características dos participantes.

Quadro 5 – Características gerais dos Participantes

Tamanho da amostragem	22
Média de idade (DP), faixa	24,5 (4,2), 19-33
Sexo	12 (54,5%) masculino 10 (45,5%) feminino
Grau de escolaridade	21 (95,5%) ensino superior 1 (4,5%) mestrado
Curso	19 (86,5%) Letras – Inglês 1 (4,5%) Biblioteconomia 1 (4,5%) Publicidade e Propaganda

	1 (4,5%) Letras – Português e Inglês
Profissão	10 (45,5%) Estudantes 7 (32%) Professores (língua inglesa) 1 (4,5%) Tutor de reforço escolar 1 (4,5%) Supervisor 1 (4,5%) Ilustrador 1 (4,5%) Servidor público 1 (4,5%) Revisor ortográfico

Fonte: Elaborado pelo autor

Como pode ser observado pelo quadro 5, a população tem média de 24,5 (4,2) anos de idade e a grande maioria está cursando ou terminou o ensino superior (graduação), com a exceção de apenas um participante que indicou nível de mestrado. No quesito “curso de graduação”, a maioria é estudante do curso de Letras – Inglês e em relação à profissão dos participantes podemos destacar uma boa quantidade de estudantes e professores (de língua inglesa) que totalizam 75,5% da população observada. Além das características gerais (não diretamente relacionadas ao nível da L2, os participantes responderam às perguntas concernentes às habilidades e método de aprendizagem da língua inglesa de acordo com o quadro 6.

Quadro 6 – Nível autodeclarado e aprendizagem de línguas

Avaliação subjetiva da fala em L2	2 (9%) Ruim 7 (32%) Razoável 5 (22,7%) Bom 5 (22,7) Muito bom 3 (13,6) Excelente
Avaliação subjetiva da compreensão auditiva em L2	2 (9%) Ruim 2 (9%) Razoável 4 (18,4%) Bom 11 (50%) Muito bom 3 (13,6%) Excelente
Avaliação subjetiva da escrita em L2	10 (45,5%) Razoável 5 (22,7%) Bom 5 (22,7%) Muito bom

	2 (9,1%) Excelente
Avaliação subjetiva da leitura em L2	2 (9%) Razoável 7 (32%) Bom 6 (27%) Muito bom 7 (32%) Excelente
Forma(s) usada(s) para aprender a língua inglesa – Aqui a soma total de respondentes passa dos 22 pois cada participante poderia marcar quantas opções fossem verdadeiras	10 (20,4%) Com amigos 3 (6,1%) Em casa 13 (26,5%) Na escola 2 (4,1%) No trabalho 21 (42,9%) Autoaprendizagem
Idade que iniciou o aprendizado da língua inglesa	2 (9,1%) Entre 3 a 7 anos de idade 11 (50%) Entre 7 a 14 anos de idade 8 (36,4%) Entre 14 a 21 anos de idade 1 (4,5%) Após 21 anos de idade
Conhecimento geral autodeclarado de outras línguas	1 (4,5%) Francês e Alemão 1 (4,5%) Libras 6 (27%) Espanhol 14 (64%) Português e Inglês somente

Fonte: Elaborado pelo autor

Ao analisar a quadro de autodeclaração de conhecimento nas quatro habilidades é possível notar que nenhum dos participantes considera-se “ruim” nas habilidades de escrita e leitura em língua inglesa, fator relevante para a pesquisa em questão tendo em vista que a tarefa proposta envolve especificamente apenas estas duas habilidades e pouco ou nada pode ser observado da fala e compreensão auditiva. Em termos relativos, quase metade (42,9%) dos participantes responderam que uma das formas que eles utilizaram para aprender a língua inglesa foi a autoaprendizagem, do ponto de vista absoluto apenas um participante dos 22 não informou este modo de aprendizado. No momento da escolha das formas de aprendizagem havia ao lado de cada uma das opções um parêntese para exemplificar o significado de cada alternativa, e para a autoaprendizagem eles tinham como exemplo “assistir filmes ou séries, ler livros, jogos digitais entre outros”, enfim, todas as formas menos tradicionais de aprender outra língua. É interessante notar que nenhum dos participantes marcou a opção de aprendizado por imersão, que para este estudo foi definido como estar fisicamente situado em um lugar onde a

língua falada era a língua inglesa, logo o aprendizado da língua inglesa pelos participantes teve rumos bem semelhantes como comprovado pela análise do quadro 6.

Outro fato notório acerca do aprendizado de língua inglesa pelos participantes fora a idade de início deste aprendizado. Metade dos participantes (11) afirmaram ter iniciado o aprendizado no início da adolescência (entre 7 e 14 anos de idade) e uma outra parte considerável (8) iniciaram no período da adolescência para a idade adulta (entre 14 a 21 anos de idade). Dados sobre o início de aprendizado de uma segunda língua são bem relevantes pois muitos estudos pontuam um “período crítico” para a aprendizagem de uma L2, ainda não é consenso no que tange a uma idade específica para este período, porém as pesquisas convergem para um melhor aprendizado da segunda língua em idade pré-puberdade (Lightbown e Spada, 1998; Frazier e Brown, 2001; Krashen e Terrell, 1983; Patkowski, 1990).

Para finalizar a parte da pesquisa sobre o conhecimento de outras línguas os participantes teriam que indicar se eles detinham conhecimento de uma outra língua que não fosse do par português-ínglês. A maioria (14 [63,6%]) afirmou ter conhecimento apenas das duas línguas que servem para o objetivo do estudo, poucos outros (6 [27,4%]) declararam conhecimento na língua espanhola, uma única pessoa (1 [4,5%]) disse saber Libras e outro participante (1 [4,5%]) alegou conhecimento em Francês e Alemão.

3.6 Aparato tecnológico

Para realizar a atividade os participantes deveriam utilizar computadores (*desktop* ou *notebook*) e não era possível realizar a tarefa utilizando aparelho celular. A maior parte da tarefa foi realizada presencialmente no Laboratório de Fonética e Multilinguismo (LabFoM) da UFC, apenas 2 dos participantes válidos realizaram a tarefa em locais distintos. Foi disponibilizado aos participantes 1 computador de mesa (*desktop*) e 2 computadores portáteis (*notebook*), o computador de mesa estava ligado a um monitor de 17 polegadas e os portáteis tinham telas de 14 e 12 polegadas respectivamente, a distância dos participantes para a tela era de aproximadamente 55 centímetros. Todas as telas estavam com a configuração padrão de taxa de amostragem a 60hz. Como as palavras ficariam por tempo prolongado na tela e o intuito principal da tarefa não exigia muita velocidade nos tempos de reação, então nenhuma frequência foi alterada pois não havia a necessidade de maior precisão nos cálculos de TR. O *software* utilizado para a coleta dos dados foi o Psytoolkit (Stoet, 2010, 2017) e para o processamento e organização dos dados foram usados o *RStudio* (RStudio 2023.12.0+369 "Ocean Storm" Release (33206f75bd14d07d84753f965eaa24756eda97b7, 2023-12-14) for

windows Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) RStudio/2023.12.0+369 Chrome/116.0.5845.190 Electron/26.2.4 Safari/537.36) com a versão 4.3.2 do “R” e Microsoft Excel.

Todos os participantes utilizaram *Microsoft Windows* como sistema operacional e também foram unânimes no uso do *Google Chrome* como navegador, 9 participantes utilizaram teclado ligado por USB enquanto 12 utilizaram o teclado acoplado do notebook e apenas um participante indicou ter usado teclado ligado por *bluetooth*.

3.7 Criação dos estímulos da tarefa de tradução

A criação dos estímulos para a atividade principal da tarefa aconteceu de maneira colaborativa com o auxílio de outros estudos que envolvem o uso de homógrafos interlinguísticos; trabalhos feitos pelos então orientandos do grupo da prof^a. Dr. Pamela Toassi (Gadelha e Toassi, 2022; e Freitas, 2023). As palavras em inglês foram escolhidas utilizando o corpus SubtlexUS, que consiste em um banco de dados de palavras baseado em filmes e séries de TV americanos, este bando de dados tem 51 milhões de palavras registradas (Brysbaert e New, 2009). Assim, as palavras para a aferição da frequência em português foram utilizadas do corpus Léxico do Português Brasileiro “LexPorBR” (Estivalet, 2019), que tem sua origem dada pelo corpus do Núcleo Interinstitucional de Linguística Computacional de São Carlos (NILC) (Pinheiro e Aluísio, 2003). É importante pontuar que ambos os corpora (SubtlexUS e LexPorBR) são de acesso gratuito para qualquer pessoa utilizar da maneira livre e aberta.

Os estímulos montados para o estudo foi subdividido em 3 grupos, todos os estímulos são palavras em inglês. O primeiro grupo (experimental) consiste em 40 homógrafos interlinguísticos do par Inglês-Português brasileiro. A partir da seleção das palavras do grupo experimental fora criado o grupo controle, este grupo consiste em 40 palavras não homógrafas que foram selecionadas de acordo com a frequência das palavras do grupo experimental. Um terceiro grupo, também denominado como “controle” foi criado a partir da frequência da tradução para o português brasileiro dos 40 homógrafos interlinguísticos do grupo experimental. Observemos os quadros a seguir para uma melhor visualização dos estímulos de cada grupo.

Quadro 7 – Lista dos estímulos da tarefa de tradução (grupo experimental [EXP])

ESTÍMULO EXPERIMENTAL	VALOR ZIPF	TRADUÇÃO	VALOR ZIPF
COME	6,4964	VIR	4,6498

TIME	6,2913	TEMPO	5,8489
SOME	6,2367	ALGUM	5,1816
MORE	6,1128	MAIS	6,5336
CASE	5,4503	ESTOJO	2,4670
LATE	5,4303	ATRASADO	3,7776
QUITE	5,3060	BASTANTE	4,9237
NINE	4,8286	NOVE	4,9732
SUGAR	4,5676	AÇÚCAR	4,3421
BABE	4,5507	BEBÊ	4,2851
GRADE	4,4661	GRAU	4,9656
PARTICULAR	4,4453	ESPECIAL	5,1948
TEAR	4,4310	RASGAR	2,9752
DATA	4,4081	DADOS	5,2357
ROLE	4,2603	FUNÇÃO	5,0973
BALANCE	4,1754	EQUILIBRIO	2,4670
CAVE	4,1455	CAVERNA	3,2762
TIRE	4,0925	PNEU	3,7705
RETIRE	3,9908	APOSENTAR	3,6741
CARGO	3,9545	CARGA	4,7671
TENDER	3,9488	MACIA	3,4266
RANGER	3,8716	GUARDA	4,3646
URGE	3,8233	IMPULSO	4,2313
METER	3,7942	METRO	4,2523
DISCO	3,7616	DISCOTECA	3,4859
RESUME	3,7557	RETOMAR	4,1802
PASTA	3,6282	MASSA	4,8321
ALIAS	3,6119	PSEUDÔNIMO	3,3878
CHUTE	3,5590	CALHA	2,5128
PROPAGANDA	3,4959	PUBLICIDADE	4,6428
ACHE	3,3990	DOR	4,5581
RIM	3,3922	ARO	3,2981
BRIM	2,9545	ORLA	3,7276
SABER	2,9150	SABRE	2,5920
GRIPE	2,9046	APERTO	3,8028
PANE	2,6897	VIDRAÇA	2,9599
PROCURE	2,6897	ADQUIRIR	4,3128
DOZE	2,6114	SONECA	2,1148
REBATE	2,5222	DESCONTO	4,5442
GENIAL	2,3332	CORDIAL	3,6166

Fonte: Elaborado pelo autor

No quadro 7, do grupo experimental, podemos observar uma esperada assimetria entre a frequência das palavras em inglês e português brasileiro. Apesar desta assimetria, geralmente as palavras com alta frequência e também as com baixa frequência costumam permanecerem em patamares relativamente similares entre as línguas.

Quadro 8 – Lista dos estímulos da tarefa de tradução (grupo controle inglês [CI])

ESTÍMULO CONTROLE	VALOR ZIPF	TRADUÇÃO	VALOR ZIPF
WELL	6,4751	BEM	5,7516
WHEN	6,3077	QUANDO	6,1130
TAKE	6,2761	PEGAR	4,5290
VERY	6,0932	MUITO	5,9106
LOST	5,4371	PERDIDO	4,4723
CRAZY	5,4344	LOUCO	4,1784
DINNER	5,3062	JANTAR	4,4597
HOLY	4,8329	SAGRADO	3,7846
SPEECH	4,5798	DISCURSO	5,0006
BLESS	4,5444	ABENÇOAR	2,3579
BORROW	4,4667	EMPRESTAR	3,7823
ASHAMED	4,4370	ENVERGONHADO	3,2762
DRAG	4,4221	ARRASTAR	3,3453
PLATE	4,4087	PLACA	4,5470
NOON	4,2579	MEIO-DIA	3,9409
BROAD	4,1680	LARGO	3,9473
BRAND	4,1449	MARCA	5,0209
INCH	4,0925	POLEGADA	3,4266
ROAST	3,9908	ASSAR	2,6267
SOCK	3,9536	MEIA	4,4180
THRILL	3,9253	EMOÇÃO	4,3719
RECEIPT	3,8716	RECIBO	3,7915
LEAP	3,8245	SALTO	4,3873
STRAW	3,7942	CANUDO	2,4670
CURLY	3,7616	ENCARACOLADO	1,5128
TEASE	3,7557	PROVOCAR	4,4799
BATCH	3,6202	LOTE	4,2338
DONOR	3,6119	DOADOR	3,6651
NAPKIN	3,5590	GUARDANAPO	2,9107
ALLEGIANCE	3,4959	FIDELIDADE	4,0971
ATTACH	3,3956	ANEXAR	3,0041
RUNT	3,3922	NANICO	2,5920
BROTH	2,9545	CALDO	3,6527
SHREW	2,9253	MEGERA	2,7169
GRAZE	2,8938	PASTAR	2,4159
MULL	2,6897	MEDITAR	3,1360
RESPITE	2,6897	TRÉGUA	3,9615
DEFT	2,6114	HÁBIL	3,7458
RECAST	2,5222	REFORMULAR	3,3820
GOITER	2,3332	BÓCIO	2,2117

Fonte: Elaborado pelo autor

As palavras do grupo controle inglês (CI) foram escolhidas com a frequência mais próxima possível do seu “par” no grupo experimental. Ao observar o quadro 8 podemos ver

que a palavra controle em inglês “well” tem seu valor de frequência em “6,4751” enquanto seu par experimental, “come” tem valor de frequência “6,4964”.

Quadro 9 – Lista dos estímulos da tarefa de tradução (grupo controle português [CP])

ESTÍMULO EXPERIMENTAL	VALOR ZIPF	TRADUÇÃO (CONTROLE)	VALOR ZIPF
WENT	5,6138	FOI	6,5060
YOUR	6,8086	SEU	6,2978
ALSO	5,3142	TAMBÉM	6,2137
DAY	5,9034	DIA	6,0802
ACTION	4,7854	AÇÃO	5,4545
THIS	6,9013	ESSE	5,4309
FIELD	4,8458	CAMPO	5,3034
RHYTHM	4,0376	RÍTIMO	4,8287
TRUCK	4,8620	CAMINHÃO	4,5678
SOIL	3,8917	SOLO	4,5506
PRIEST	4,4179	PADRE	4,4675
CHAIR	4,6918	CADEIRA	4,4458
CONQUERED	3,4410	CONQUISTOU	4,4308
SIDES	4,2098	LADOS	4,4071
AID	4,1430	AUXÍLIO	4,2602
NOSE	4,8430	NARIZ	4,1755
SMILE	4,7629	RIR	4,1462
CHOIR	3,7264	CORO	4,0914
HEAVY	4,6743	PESADO	3,9913
FORGIVENESS	3,9003	PERDÃO	3,9553
UNHAPPY	4,2181	INFELIZ	3,9489
SOAP	4,1816	SABÃO	3,8726
NEIGHBORHOOD	4,5641	VIZINHANÇA	3,8245
STORAGE	3,9479	ARMAZENAMENTO	3,7938
VACUUM	3,7616	VÁCUO	3,7608
RELIEVE	3,7466	ALIVIAR	3,7558
BRIBE	3,7817	PROPINA	3,6267
DOLL	4,3935	BONECA	3,6097
FOG	3,9757	NEBLINA	3,5581
TURTLE	4,2313	TARTARUGA	3,4950
SWELLING	3,5013	INCHAÇO	3,3993
STRIP	4,1954	TIRA	3,3936
ESCORT	4,1668	ESCOLTA	2,9599
PADLOCK	2,8481	CADEADO	2,9107
SADDLE	3,8938	SELA	2,9107
PIE	4,4582	TORTA	2,6889
STRIKE	4,6582	GOLPEAR	2,6889
SCONCE	2,2460	CANDEEIRO	2,5920
BREATHALYZER	1,8938	BAFÔMETRO	2,5128
USABLE	2,7389	USÁVEL	2,3579

Fonte: Elaborado pelo autor

Este último quadro foi elaborado a partir do valor da frequência da palavra em português brasileiro, portanto a mais alta inicia com “foi” (6,5060) e termina com o adjetivo “usável” (2,3276). Neste quadro, a frequência das palavras em inglês não sofreu nenhum tipo de controle, a manipulação dos estímulos está na frequência da tradução.

Além dos quadros elaborados para a tarefa principal, 10 palavras foram selecionadas para a parte de treinamento. Antes de responder às 120 palavras da tarefa principal, os participantes deveriam realizar um breve treinamento e traduzir 10 palavras divididas da seguinte maneira: 5 homógrafos interlinguísticos e 5 palavras controle. Como a parte de treinamento não entraria durante a análise final dos dados, não foi feito qualquer tipo de tratamento de dado ou escolha minuciosa de palavras, a única regra posta seria que nenhuma das palavras contidas no treinamento aparecessem durante a tarefa principal para evitar um efeito de *priming* por repetição, efeito já observado e constatado ser um facilitador durante a realização de tarefas cognitivas como as de decisão lexical, tarefas de memória, teste de completar palavras, em tarefas do tipo “sim/não” e em tarefa de recordação de escrita (Scarborough, Cortese, e Scarborough, 1977; Jacoby, 1983; Jacoby e Dallas, 1981; Roediger e Blaxton, 1987; Schacter, 1990; Tulving e Schacter, 1990), vejamos adiante os quadros 10 e 11 com os estímulos usados no treinamento prévio à tarefa principal.

Quadro 10 – Lista dos homógrafos interlinguísticos do treinamento

ESTÍMULOS DO TREINAMENTO (HOMÓGRAFOS INTERLINGÜÍSTICOS)	TRADUÇÃO
AGE	IDADE
GAME	JOGO
TAPE	FITA
AGENDA	PAUTA
COSTUME	FANTASIA

Fonte: Elaborado pelo autor

Quadro 11 – Lista das palavras controle do treinamento

ESTÍMULOS DO TREINAMENTO (PALAVRAS CONTROLE)	TRADUÇÃO
ICE	GELO
GAVE	DEU
TRICK	TRUQUE
AMONGST	ENTRE
COWARD	COVARDE

Fonte: Elaborado pelo autor

Como a maioria das palavras escolhidas como estímulo são polissêmicas, isto é, elas contêm mais de um significado (ex.: *well* pode significar o advérbio “bem”, ou o substantivo “poço”, todas as duas traduções devem ser consideradas corretas), fez-se necessário uma meticulosa escolha acerca da resposta dos participantes em cada uma das palavras. Outro impasse se deu com a observação da tradução de palavras da classe gramatical dos verbos e os pronomes, como no português brasileiro nós utilizamos uma conjugação verbal com desinências diferentes para cada pessoa do sujeito, foi necessário considerar qualquer tradução da forma verbal como correta (ex.: *lost*, teve sua tradução configurada para ser corretamente aceita como “perdido”, porém as variações “**perdeu** – *He’s lost his phone*”, “**perdi**” – *I’ve lost my keys*”, “**perdemos**” – *We’ve lost the match*” também seriam consideradas corretas). Todo o processo de ajustes de respostas certas ou erradas será escrutinado em seção posterior.

3.7.1 A criação de um novo corpus e medição da frequência dos estímulos

Uma das partes principais ao montar o corpus do estudo está na frequência das palavras utilizadas (Brysbaert *et al.*, 2011), pois palavras de alta frequência são conhecida por um maior número de pessoas e são também processadas de maneira mais rápida que palavras de baixa frequência, é o que chamamos de *efeito de frequência da palavra* (Monsell, Doyle e Haggard, 1989). Em Brysbaert *et al.*, (2016) estima-se que a característica “frequência da palavra” consegue explicar até quarenta por cento das variações que acontece em tarefa de reconhecimento de palavras. Antigamente o corpus mais comumente utilizado era o de Kucera e Francys (1967) e a contagem de palavras vinha na ordem de “palavra por milhão”, este antigo corpus continha pouco mais de um milhão de palavras, circunstância essa criticada por Burgess e Livesay (1998) pela pequena variação utilizada no corpus, que apesar de mesclar várias áreas do conhecimento para catalogar suas palavras, era alimentado basicamente por textos mais científicos, e pouco contava com palavras ou expressões do nosso cotidiano.

Pensando em uma alternativa para esta visão consolidada, porém que não refletia a realidade da vida das pessoas, os pesquisadores estabeleceram uma nova norma para ser usada como corpus de pesquisa. A fonte encontrada para esse novo corpus foi a análise das legendas em filmes e séries televisivas, e como os autores pontuaram, há uma grande diferença prática na própria elaboração dos corpora em razão da origem das palavras nele contidas, vejamos o que diz Brysbaert e New (2009, traduzido por Google, adaptação nossa):

“As principais fontes durante muito tempo foram livros, jornais e revistas. Em particular, para a investigação sobre o reconhecimento visual de palavras, estas foram consideradas as fontes de informação mais importantes. Os problemas com essas

formas de insumo, entretanto, são que eles geralmente foram editados (para aperfeiçoar a linguagem), que tendem a exagerar a variação lexical (para não repetir a mesma palavra indefinidamente) e que muitas vezes lidam com temas que não estão na rotina da vida das pessoas.” (Brysaert e New, 2009. p. 978-979)²⁶

Então, logo percebendo que carecia nos estudos com corpus um levantamento mais moderno e voltado para o “real” ou “natural”, os autores criaram o SUBTLEX_{US}. O corpus é baseado nas legendas inseridas no sítio www.opensubtitles.org, um sítio que agrega várias legendas enviadas pelos próprios usuários para que outras pessoas possam baixar e assistir aos seus filmes e séries com as legendas desejadas. Ao final da coleta e limpeza de todas as palavras das legendas obtidas, os autores ficaram com um total e mais de 51 milhões de palavras, um montante bem expressivo em comparação ao número visto em Kucera e Francis (1967). Brysaert e New (2009) detalham como fora feito o processo da coleta para a formação do corpus, ficando assim de fácil acesso a reprodução de um corpus similar em qualquer outra língua (que esteja presente no sítio *opensubtitles*).

Para além da melhoria de corpus, os pesquisadores (Van Heuven *et al.*, 2014) queriam também elaborar uma métrica para utilizar como cálculo da frequência de palavras, foi então que baseados no trabalho de Zipf (1949) eles criaram “escala Zipf” que leva o nome de seu idealizador, o linguista George Kingley Zipf (1902-1950). Van Heuven *et al* (2014) nos mostra a métrica até então mais utilizada que é a frequência por milhão de palavra (fpm), essa métrica surgiu precisamente do antigo corpus de Kucera e Francis (1967) pois como ele continha pouco mais de um milhão de palavras, a menor frequência observada seria de um em um milhão, conceito pouco relevante para pesquisas atuais com corpus que passam de 100 milhões de palavras, onde a ocorrência de 1 fpm poderia ser observada em várias instâncias. A ideia de um valor absoluto para a frequência por milhão não é produtiva na concepção dos pesquisadores, pois eles afirmam que este número pouco diz sobre a palavra e na verdade a real interpretação do número vai depender do tamanho do corpus a ser observado:

“[...] eles (contagem de palavras em frequência por milhão) têm uma grande desvantagem. A interpretação da medida de frequência depende do tamanho do corpus. [...] Isso contribuiu para a suposição de que 1 palavra por milhão é a frequência mais baixa possível. Obviamente, este já não é o caso de corpora maiores. Acontece que cerca de 80% dos tipos de palavras no SUBTLEX-UK têm uma frequência inferior a 1 palavra por milhão (ou seja, menos de 200 ocorrências em

²⁶ The main sources for a long time were books, newspapers, and magazines. In particular, for research on visual word recognition, these were thought to be the most important sources of input. Problems with these forms of input, however, are that they usually have been edited (to polish the language), that they tend to exaggerate lexical variation (in order not to repeat the same word over and over again), and that they often deal with topics that are not at the forefront of people’s lives. (BRYSAERT E NEW, 2009. p. 978-979)

todas as transmissões).²⁷ (Van Heuven et al., 2014. p. 1178-1179, traduzido por Google, adaptação nossa)

Os autores propõem que para facilitar a interpretação da frequência de palavras a nova métrica deve:

- a) ser uma escala logarítmica (assim como a escala sonora de decibel);
- b) deve ter relativamente poucos pontos e não conter valores negativos (assim como uma escala Likert típica variando de 1 a 7);
- c) o ponto médio da escala deve separar as palavras de baixa frequência das palavras de alta frequência.
- d) a escala deve ter uma unidade simplificada.

Para satisfazer as três primeiras demandas bastava usar uma escala logarítmica na base 10 em frequência por bilhão de palavras, já a última demanda veio como a homenagem à pesquisa de Zipf (1949) que fez com que a gradação das palavras permanecesse nessa escala simplificada que varia de 1 até 7. Por ser uma escala logarítmica, a escala Zipf varia do número 1 (palavras de baixíssima frequência) até o número 6 (palavras de altíssima frequência), podendo também chegar ao número 7 para palavras de função como pronomes e formas verbais mais simples. O cálculo desta frequência é bastante simples, como ela é derivada do logaritmo na base 10 (frequência por bilhão ou por milhão [+3]), basta entender que o valor “1” significa que estas palavras tem uma frequência de 1 por 100 milhões de palavras, o valor “2” já nos traz palavras com a frequência de 1 por 10 milhões de palavras, o valor “3” corresponde à frequência de 1 por um milhão de palavras e assim por diante. Vejamos em mais detalhes na tabela 1 exemplificada pelos autores:

²⁷ “[...] they have one big disadvantage. The interpretation of the frequency measure depends on the size of the corpus. [...] This contributed to the assumption that 1 fpmw is the lowest possible frequency. Obviously, this is no longer the case for larger corpora. As it happens, about 80% of the word types in SUBTLEX-UK have a frequency of less than 1 fpmw (i.e., fewer than 200 occurrences in all broadcasts).” (Van Heuven et al., 2014. p. 1178-1179)

Tabela 1 – A escala Zipf de frequência de palavras

Valor Zipf	FPM	Exemplos
1	0.01	antifungal, bioengineering, farsighted, harelip, proofread
2	0.1	airstream, doorkeeper, neckwear, outsized, sunshade
3	1	beanstalk, cornerstone, dumpling, insatiable, perpetrator
4	10	dirt, fantasy, muffin, offensive, transition, widespread
5	100	basically, bedroom, drive, issues, period, spot, worse
6	1000	day, great, other, should, something, work, years
7	10,000	and, for, have, I, on, the, this, that, you

Fonte: Adaptado de Van Heuven *et al.*, (2014, p. 1180).

Nota dos autores (traduzida por Google, adaptação nossa): A escala Zipf é uma escala de frequência de palavras que vai de 1 a 7. Palavras com valores Zipf de 3 ou menos são palavras de baixa frequência; palavras com valores Zipf de 4 ou superiores são palavras de alta frequência. Os exemplos são baseados nas frequências de palavras do SUBTLEX-UK. fpm = frequência por milhão de palavras.²⁸

Para termos uma ideia prática de como essa escala facilita a interpretação e até a polidez visual dos dados, vejamos o quadro montado por Borém (2023) que ilustra a diferença entre as escalas Zipf e frequência por milhão (fpm).

Quadro 12 – Representação da escala Zipf versus Frequência por milhão

Cognato	Freq. Zipf	Cognato	Freq. Zipf	Cognato	Subtlwf	Cognato	Freq_orto/M
academic	3,7	acadêmico	3,4	academic	4,7	acadêmico	2,5
accident	4,9	acidente	5	accident	81,3	acidente	101,6
alarm	4,5	alarme	3,9	alarm	29,8	alarme	7,2
actor	4,4	ator	4,9	actor	26,3	ator	72
author	3,9	autor	5,2	author	21,2	autor	169,1
confusion	3,9	confusão	4,6	confusion	7,1	confusão	38,9
drug	4,7	droga	4,5	drug	45,2	droga	34,1
effects	4,1	efeitos	4,9	effects	12,7	efeitos	86,7
exams	3,5	exames	4,8	exams	13,4	exames	30,7
fruit	4,3	fruta	3,9	fruit	21,7	fruta	7,6
galaxy	3,8	galáxia	3,4	galaxy	6,7	galáxia	2,4
insect	3,5	inseto	3,4	insect	3,2	inseto	2,6
melon	3,6	melão	3	melon	4,3	melão	1,1
memory	4,7	memória	4,8	memory	48,6	memória	64,8
minute	5,6	minuto	4,4	minute	377,5	minuto	26,3
music	5,2	música	5,3	music	151,7	música	182,3
palace	4,3	palácio	4,4	palace	19,2	palácio	23,3
panic	4,3	pânico	4,1	panic	21,8	pânico	12,6

²⁸ The Zipf scale is a word frequency scale going from 1 to 7. Words with Zipf values of 3 or lower are low-frequency words; words with Zipf values of 4 and higher are high-frequency words. Examples are based on the SUBTLEX-UK word frequencies. fpmw = frequency per million words. Van Heuven *et al.*, (2014, p. 1180).

plant	4,4	planta	4,2	plant	27,6	planta	14
protection	4,4	proteção	4,7	protection	23,5	proteção	54,4
region	3,7	região	5,5	region	5	região	297,7
revision	2,8	revisão	5,1	revision	0,6	revisão	119,3
series	4,3	série	5,2	series	20,2	série	167
tractor	3,6	trator	3,5	tractor	3,7	trator	2,8
union	4,3	união	4,9	union	21,8	união	76,5
universe	4,4	universo	4,7	universe	25,3	universo	46,8

Fonte: Borém (2023).

Nota: A parte mais à esquerda do quadro (com o fundo azul claro) está a demonstração das palavras pela escala Zipf, e a parte mais à direita do quadro (sem cor de fundo) é demonstrado a medição em palavras por milhão.

Ao analisar a diferença de gradação entre as métricas (Zipf *versus* palavra por milhão) fica evidente a fácil compreensão de utilizar a escala Zipf. Se compararmos as frequências da palavra “revision” ou “revisão” teremos respectivamente os números 2,8 e 5,1 em valores Zipf, indicando uma baixa frequência em língua inglesa e alta frequência em língua portuguesa (brasileira), já em palavras por milhão os valores ficam descritos como 0,6 e 119,3, pode-se sim inferir que 0,6 é um valor baixo, mas a relação de discrepância com o valor 119,3 é altíssima, além do mais não temos um ponto médio para indicar se um valor “x” da frequência por milhão representa uma palavra de alta ou baixa ocorrência.

Levando em consideração todo o exposto, a métrica de frequência utilizada para as palavras-estímulo neste estudo é a escala Zipf, pois como já foi demonstrado ela é uma escala mais simplificada do que “palavra por milhões” e tem mais acessibilidade de interpretações corretas pelo simples fato de não depender do tamanho do corpus utilizado e possuir uma divisão mais concreta entre palavras de baixa frequência (valor Zipf ≤ 3) e palavras de alta frequência (valor Zipf ≥ 4).

3.8 Procedimentos da pesquisa experimental

Nesta seção discorreremos acerca das etapas que envolvem o percurso global da pesquisa, desde a leitura e aceite do TCLE ao último envio de resultados ao pesquisador. As etapas serão comentadas e demonstradas com capturas de tela para cada procedimento realizado (com exceção do TCLE).

Para começar, os participantes deveriam clicar no link disponibilizado de maneira *online* pelo sistema institucional (SIGAA) da UFC a fim de iniciar a leitura do TCLE e começar a pesquisa, ou para os que optaram por participar da pesquisa no laboratório (LabFoM) os 3 dispositivos disponibilizados (1 *desktop* e 2 *notebooks*) já estavam previamente configurados

na tela inicial da pesquisa (TCLE). Todos os participantes deveriam confirmar a leitura e aceite do termo antes de iniciar a pesquisa, o termo iniciava com um convite aos participantes e logo após continha as seguintes seções:

- a) objetivo do estudo;
- b) procedimentos: (1ª etapa: Leitura do TCLE, 2ª etapa: Identificação, 3ª etapa: Questionário biográfico e linguístico, 4ª etapa: Treinamento da tarefa, 5ª etapa: Tarefa principal e 6ª etapa: Teste de vocabulário em língua estrangeira;
- c) riscos;
- d) benefícios;
- e) direitos dos participantes;
- f) compensação financeira; e
- g) utilização dos dados

Havia uma descrição clara e objetiva sobre cada etapa da pesquisa, contendo breve instruções de como proceder e o tempo estimado para cada fase do processo, ao final estavam descritos os dados de contato dos pesquisadores e instituição envolvidos. Todo o processo estava em concordância com a resolução 510/16 do Conselho Nacional de Saúde (CNS).

3.8.1 – Identificação

Logo após leitura e aceite do TCLE o participante era recepcionado com a tela para a criação do código de identificação. Nesta etapa o participante deveria criar um código alfanumérico de duas letras e dois números (ex.: TE80) que serve como a sua identificação durante toda a pesquisa, no intuito de preservar a identidade dos participantes. Foi recomendado que escolhessem duas letras aleatórias que não sejam iniciais do nome/sobrenome e que também não escolhessem números com finais do ano de nascimento (ex.: caso um participante tenha o nome “Bruno Paes Lemos” e tenha nascido em 1995, ele deverá evitar códigos de identificação como “BP95”, “BL95” ou “CA95”). Vejamos a figura 15 com a captura da tela referente a esta etapa.

Figura 15 – Identificação do participante

Identificação do participante

Insira o código alfanumérico de duas letras e dois números (ex.: HS46) para a tarefa de tradução. Evite usar iniciais do seu nome e/ou data de nascimento. LEMBRE-SE DESSE CÓDIGO PARA O RESTANTE DA PESQUISA

Clique neste botão para continuar

Fonte: Elaborado pelo autor

3.8.2 – *Questionário biográfico e linguístico*

Após a identificação os participantes deveriam responder ao questionário biográfico, as instruções de resposta estarão entre parênteses. Vejamos então a sequência de perguntas para esta parte do questionário:

- a) nacionalidade (Indique a sua nacionalidade e a sua cidade/estado atual);
- b) idade (Qual a sua idade [em anos]);
- c) sexo (Qual seu sexo [biológico]);
- d) grau de escolaridade (Superior/ especialização/ mestrado/ doutorado);
- e) formação (Indique sua área de formação (concluída ou não, ex.: "Direito", "Jornalismo" etc.) e
- f) ocupação (Qual sua profissão atual (ex.: "estudante", "arquiteto", "veterinário" etc.).

Após responder às perguntas do questionário biográfico os participantes eram prontamente requeridos a responder o questionário de nível linguístico autodeclarado. De início eles teriam que responder qual o nível percebido em cada uma das quatro habilidades (fala, compreensão auditiva, escrita e leitura) na língua inglesa como mostra a figura 16.

Figura 16 – Nível estimado de língua inglesa nas quatro habilidades

Estime seu nível de inglês em cada uma das quatro habilidades.

Fala

Muito ruim	Ruim	Razoável	Bom	Muito bom	Excelente
------------	------	----------	-----	-----------	-----------

Compreensão auditiva

Muito ruim	Ruim	Razoável	Bom	Muito bom	Excelente
------------	------	----------	-----	-----------	-----------

Escrita

Muito ruim	Ruim	Razoável	Bom	Muito bom	Excelente
------------	------	----------	-----	-----------	-----------

Leitura

Muito ruim	Ruim	Razoável	Bom	Muito bom	Excelente
------------	------	----------	-----	-----------	-----------

Fonte: Elaborado pelo autor

Decidimos utilizar um modelo de escala Likert com a descrição de cada um dos itens ao invés de números para graduar o nível autodeclarado dos participantes, a escala continha em essência 6 pontos (1 – Muito ruim, 2 – Ruim, 3 – Razoável, 4 – Bom, 5 – Muito bom e 6 – Excelente). Outra decisão foi a de utilizar um número par na seleção dos itens para evitar um efeito tendencioso de escolher o ponto mediano (Chyung *et al.*, 2017), mas ao mesmo tempo a fim de “evitar” uma decisão forçada entre um ponto inferior (1 a 3) ou um ponto superior (4 a 6) o mais alto ponto inferior (3 – Razoável) servia como o “alívio” ou um falso ponto médio para os respondentes, ficando apenas os pontos “1” e “2” (Muito ruim e Ruim) como respostas objetivamente negativas.

A próxima pergunta a respeito do questionário linguístico era sobre a maneira que os participantes aprenderam a língua inglesa, vejamos a figura 17 para maiores detalhes.

Figura 17 – Forma(s) de aprendizado da língua inglesa

Indique a(s) forma(s) que você aprendeu inglês. (Selecione quantas opções forem verdade para você)

Imersão (morou em lugar onde a língua falada era o inglês)

Com amigos (treinando a língua inglesa)

Em casa (aulas particulares/reforço)

Na escola (tradicional ou bilíngue)

No trabalho (necessidade do cargo ex.: viagem para o exterior, apresentações, etc.)

Autoaprendizagem (Vendo filmes/séries, lendo livros, jogos digitais, etc.)

Fonte: Elaborado pelo autor

Para cada uma das opções os participantes tinham ao lado um exemplo simplificado do que aquele item significava. Aqui o participante estava livre para marcar quantas opções fossem verdadeiras acerca de como ele aprendeu inglês. O vocábulo *aprender* é aqui utilizado de maneira abrangente e engloba todas as definições teóricas da dualidade “aquisição/assimilação x aprendizado/estudo” de Krashen (1987 e 1988). Outra consideração a ser feita, porém já explicada pela imagem, é que o item “imersão” foi colocado de maneira exclusiva para aquela pessoa que morou em algum lugar onde a língua falada era o inglês, e não como o método imersivo de ensinar/aprender a língua em questão.

A próxima pergunta solicitava que os participantes respondessem com a idade que iniciaram a aprender inglês, a pergunta estava formulada da seguinte maneira: “Com qual idade você iniciou o aprendizado da língua inglesa?”. Havia uma caixa de diálogo para a inserção da resposta, que poderia ser tanto em número quanto por extenso. Subsequente a esta pergunta os participantes deveriam responder se tinham conhecimento em outras línguas e qual o nível geral como mostrado na figura 18:

Figura 18 – Questionário de conhecimento de línguas estrangeiras

Conhecimento de outras línguas estrangeiras

Caso você tenha conhecimento em outras línguas além do inglês e português, indique a(s) língua(s) e estime o seu nível de conhecimento (Muito ruim, Ruim, Razoável, Bom, Muito bom, Excelente). Caso contrário apenas passe para a próxima pergunta.

Fonte: Elaborado pelo autor

Aqui consideramos como *língua estrangeira* qualquer língua que o participante detivesse conhecimento que não fosse sua língua materna (L1), isto é, qualquer outra língua

que não fosse o português brasileiro no caso da pesquisa em questão. Um dos participantes respondeu que detinha conhecimento em Libras, e por mais que seja um idioma reconhecidamente brasileiro ela não deixaria de ser uma língua estrangeira, assim como o português brasileiro o é para o falante nativo de Libras, vejamos um trecho de Gesser (2006) que discorre sobre o assunto:

“[...] e em outros momentos como “língua estrangeira”, ou seja, como uma língua alheia, estranha à língua tradicionalmente oral desses ouvintes. É uma relação de contato/conflito não só de línguas mutuamente ininteligíveis, mas também de modalidades distintas para a comunicação. [...] portanto, de uma “língua estrangeira” em seu sentido mais amplo, pois sabemos que a comunidade majoritária ouvinte pertence a uma tradição oral – e aqui não me refiro em oposição à modalidade escrita – que concebe a língua no sentido vocal-auditivo e não espaço-visual. Ao tratar a relação dos ouvintes com a LS como “estrangeira” não estou levando em consideração somente questões de modalidades distintas, bem como o fato de a LS pertencer a uma minoria lingüística “invisível”, e que não é falada e entendida na sociedade brasileira [...]”. (Gesser, 2006, p. 65-67).

É sabido que a nomenclatura para língua materna/ primeira língua (LM/L1), segunda língua (L2) e língua estrangeira (LE) têm suas definições delimitadas, porém, as definições tanto de segunda língua (L2) quanto língua estrangeira (LE) são intercambiáveis dentro do escopo desta pesquisa.

3.8.3 – *Questionário de tecnologia utilizada*

Após as perguntas de cunho lingüístico os participantes responderam às perguntas relacionadas às tecnologias utilizadas durante a realização da pesquisa. No total haviam 4 perguntas:

- a) que tipo de máquina você está usando nesta pesquisa (*Desktop* ou *Notebook*);
- b) qual sistema operacional você está utilizando (Windows, macOS ou Linux);
- c) qual navegador você está utilizando (Chrome, Edge, Firefox, Safari, Opera ou outros) e
- d) qual o tipo de conexão do seu teclado (Conectado via cabo usb, conexão sem fio ou integrado ao *notebook*).

Estas perguntas de base puramente tecnológica servem para uma análise mais aprofundada em caso de irregularidades na análise dos dados. Ao tomar como exemplo da última pergunta (d) que questionava sobre o tipo de conexão do teclado, podemos por consequência de uma conexão mais lenta (conexão sem fio “*wireless*” ou *bluetooth*) nos deparar com maiores tempos de reação, pois normalmente teclados com conexão sem fio têm um maior atraso na resposta e maiores perda de transmissão de dados fazendo com que estes possam gerar dados distorcidos no estudo.

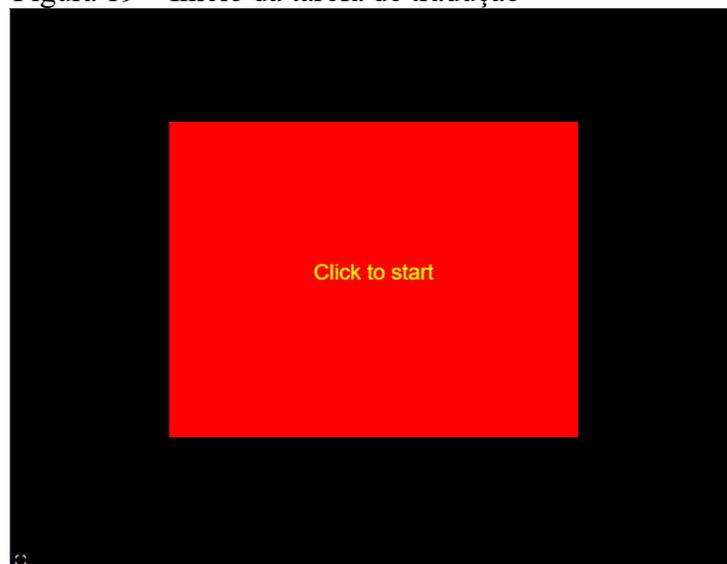
Vale ressaltar que as divisões “questionário biográfico e linguístico” e “questionário de tecnologia utilizada” foram criadas apenas para melhor segmentar a escrita deste estudo, e que na realização da pesquisa pelos participantes não havia nenhum tipo de divisão explícita separando estas etapas.

Ao terminar os questionários, o participante estava pronto para realizar a tarefa de tradução. Vejamos na próxima subseção os detalhes referentes a esta tarefa.

3.9 Tarefa de tradução

Antes de iniciar qualquer tipo de leitura sobre a tarefa o participante era exposto a uma tela padrão do *software* com os dizeres em amarelo “click to start” (clique para iniciar) no centro da tela, com fundo vermelho seguido por um plano de fundo preto, como mostra a figura 19.

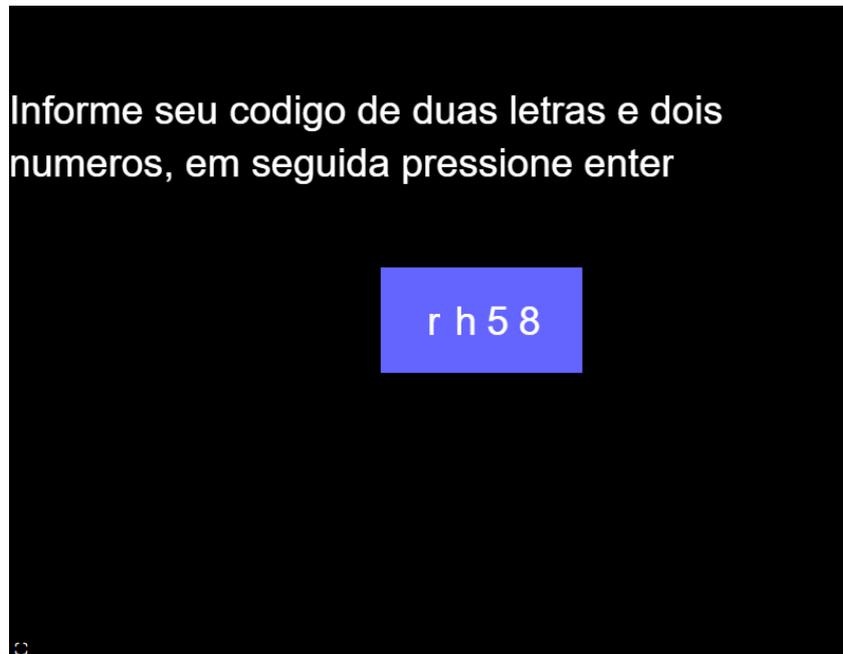
Figura 19 – Início da tarefa de tradução



Fonte: Elaborado pelo autor

Após dar início à etapa da tarefa de tradução o participante deveria repetir o seu código criado na etapa *identificação* para assim os seus dados ficarem atrelados àquele número e ser possível a sua identificação em alguma ocorrência futura (FIGURA 20).

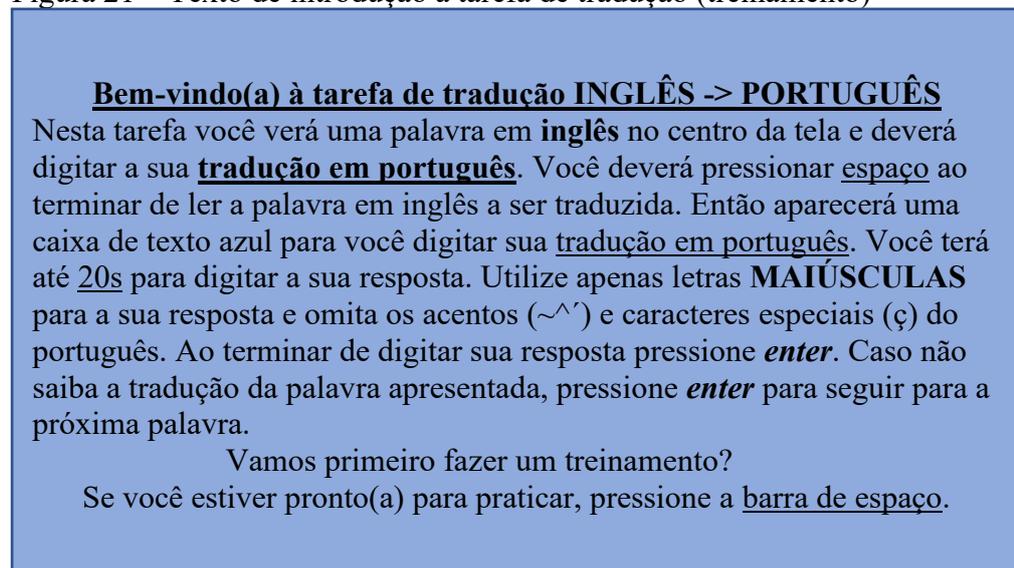
Figura 20 – Identificação do participante



Fonte: Elaborado pelo autor

Logo que o participante indicasse o código e apertasse a tecla *enter*, ele seria direcionado para a tela das instruções da tarefa que continha o seguinte texto (FIGURA 21).

Figura 21 – Texto de introdução à tarefa de tradução (treinamento)

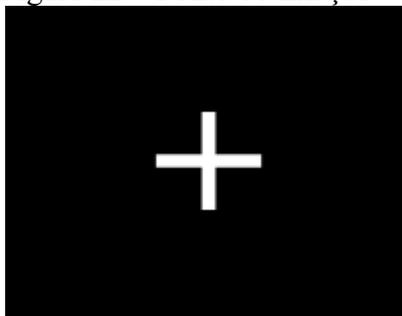


Fonte: Elaborado pelo autor

Todas as informações estavam expostas de maneira bem sucinta, e as informações mais pertinentes à pesquisa em questão estavam enfatizadas com algum tipo de destaque textual como a utilização de palavras em negrito, sublinhadas, em itálico ou maiúsculas. O tipo de fonte utilizado nas etapas da pesquisa foi a “Arial” com o tamanho 28. Logo que o participante pressionasse a barra de espaço ele seria levado ao treinamento.

Antes da primeira palavra do treinamento aparecer, um ponto de fixação era exposto no centro da tela por 500 milissegundos (ms). Este ponto de fixação é presente durante todo o resto da tarefa de tradução e sempre reaparece na ocorrência de um novo estímulo a ser traduzido (FIGURA 22).

Figura 22 – Ponto de fixação



Fonte: Elaborado pelo autor

Apesar do ponto de fixação estar disposto no centro da área útil da tarefa, a palavra a ser traduzida era mostrada um pouco acima deste ponto, porém sempre centralizada a ele. O ponto de fixação já foi mostrado ser um elemento facilitador de reconhecimento de palavras de acordo com a sua posição inicial em relação ao estímulo, tudo isso devido a especialização hemisférica, hábitos de leitura ou à restrição léxica (capacidade de inferir uma sequência que forma a palavra, vide como exemplo o modelo de ativação interativa bilíngue em Dijkstra, T., & van Heuven, W. J. B., 1998), como visto em Farid e Grainger (1996). A apresentação dos estímulos seguia a ordem “ponto de fixação → estímulo → resposta”.

Após o ponto de fixação sumir da tela o participante teria acesso à primeira palavra a ser traduzida do treinamento. Todas as palavras apareciam de maneira aleatória para todos os participantes em quaisquer que fossem as condições, treinamento ou tarefa. Veja a figura 23 com a demonstração de um estímulo na fase de treino.

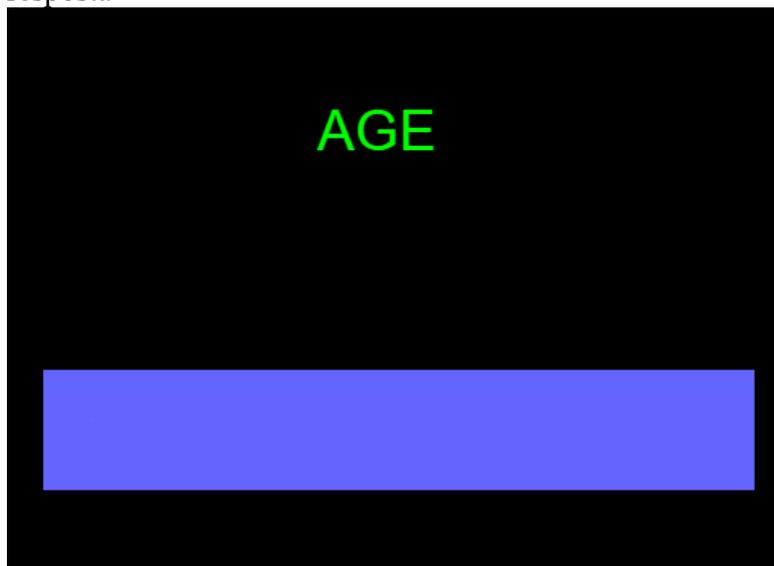
Figura 23 – Estímulo (treinamento)



Fonte: elaborado pelo autor

A palavra em questão, o homógrafo “age” (idade) deveria ser lido mentalmente pelo participante, e logo em seguida ele deveria apertar a tecla *espaço* para digitar a sua resposta (tradução). A palavra permanecia na tela por um total máximo de 25 segundos, sendo os primeiros 5 segundos reservados para a leitura e o restante para a que fosse digitada a sua tradução. Mesmo que o participante não apertasse a tecla espaço, a caixa de texto para inserir a resposta seria automaticamente liberada com o passar dos 5 segundos (FIGURA 24).

Figura 24 – Estímulo (treinamento) com a caixa de texto para resposta



Fonte: Elaborado pelo autor

Com o surgimento da caixa de texto o participante teria agora 20 segundos para inserir sua resposta, no exemplo mostrado anteriormente o participante deveria digitar a palavra “idade” que seria a correta resposta para o homógrafo interlinguístico “age”. A fase de treinamento contava com 10 estímulos, e o tempo máximo para a realização desta etapa seria limitado a 255 segundos (soma-se aos 250 segundos do intervalo de leitura mais resposta os 5 segundos obrigatórios do total de dez pontos de fixação entre os estímulos). Ao término da fase de treinamento o participante se deparava com a tela de finalização de treinamento e mais uma vez era reforçado a ele as instruções da pesquisa de acordo com o texto a seguir (FIGURA 25).

Figura 25 – Texto de início da tarefa de tradução

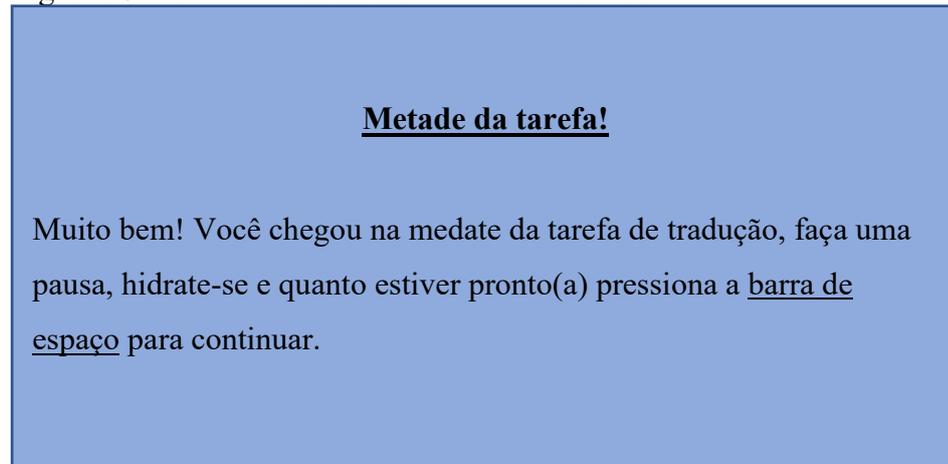
Muito bem!
Você terminou o treinamento!

Lembre-se:
Você deverá pressionar espaço ao terminar de ler a palavra em **inglês** a ser traduzida. Então aparecerá uma caixa de texto azul para você digitar sua **tradução em português**. Você terá até 20s para digitar sua resposta. Utilize apenas letras **MAIÚSCULAS** para a sua resposta e omite os acentos (~^´) e caracteres especiais (ç) do português. Ao terminar de digitar a sua resposta pressione ***enter***.
Caso não saiba a tradução da palavra apresentada pressione ***enter*** para seguir para a próxima palavra. Serão 120 palavras no total, apresentadas uma de cada vez.
Se você estiver pronto(a) para começar, pressione a barra de espaço.

Fonte: Elaborado pelo autor

A partir deste ponto o participante daria início à parte experimental da tarefa de tradução. O funcionamento ocorre da mesma forma do treino, porém como agora ele deveria traduzir 120 palavras ao invés de apenas 10, na metade da tarefa (após a sexagésima palavra) era mostrado ao participante uma tela indicando que ele atingiu esta marca (FIGURA 26). Após fazer uma breve pausa, ou caso o participante estivesse se sentindo disposto, ele poderia concluir a outra metade das palavras para terminar a tarefa de tradução.

Figura 26 – Tela de aviso de metade da tarefa concluída

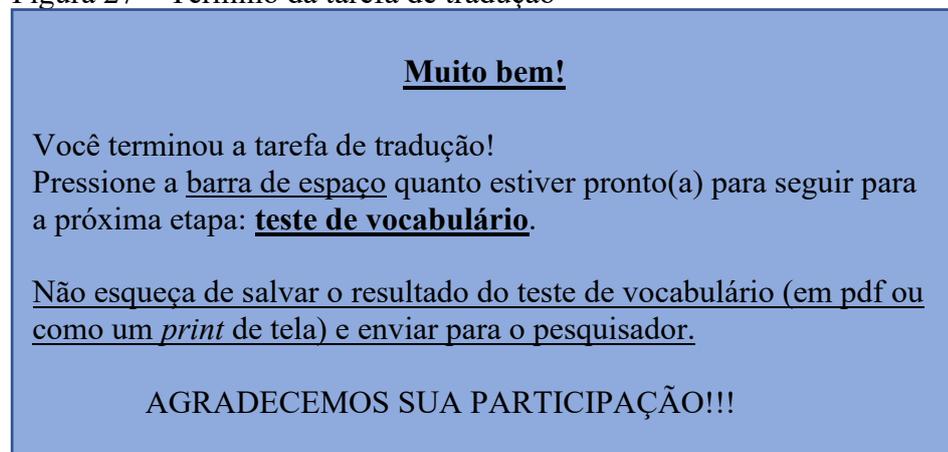


Fonte: Elaboração própria

Nota: A palavra “metade” ficou grafada de maneira incorreta (medate) porém não influenciava no resultado final da pesquisa, assim como a palavra “quando” foi erroneamente grafada como “quanto”

Os três grupos de estímulos (QUADRO 7, QUADRO 8 e QUADRO 9) que compunham a parte experimental da pesquisa foram todos aleatoriamente distribuídos entre grupos e entre participantes, de modo que nenhum participante teria a chance de encontrar a mesma sequência de palavras, pois o total de combinações possíveis gera um número demasiadamente grande que é melhor expresso por notação científica (combinações possíveis = $n!$, $120! = 6,68950291 * 10^{198}$). Após terminar a tarefa de tradução o participante teria mais um texto explicativo parabenizando-o e agradecendo pela sua participação (FIGURA 27) e após a tela de agradecimentos era solicitado ao participante que clicasse em um *hyperlink* para seguir à última tarefa referente ao nível de conhecimento de vocabulário em língua inglesa.

Figura 27 – Término da tarefa de tradução



Fonte: Elaborado pelo autor

Nota: A palavra “quando” foi grafada novamente de maneira incorreta (quanto), porém também não apresentava efeito prejudicial aos participantes por não entrar em contagem de tempo.

3.10 Teste de produção de vocabulário em língua inglesa

A última etapa da parte experimental da pesquisa ocorria logo após a realização das etapas anteriormente descritas. Aqui os participantes realizaram um teste nível de conhecimento de vocabulário em língua inglesa. O sítio que permite realizar esta atividade é o https://www.itt-leipzig.de/static/vltenglish_01p/index.html, este era o *hyperlink* associado ao término da tarefa de tradução.

Ao entrar no devido sítio para realizar o teste de vocabulário o participante tinha acesso às instruções da tarefa que dizia o seguinte: “Bem-vindo ao teste produtivo de vocabulário em inglês. Teste seu vocabulário produtivo em inglês. Quantas das 5000 palavras mais frequentes em inglês você conhece? Seu tempo limite para esse *quiz* é de 30 minutos. Não utilize dicionário. Boa sorte!”. Abaixo das instruções havia a referência para a montagem da lista de palavras utilizadas neste teste e a data da última atualização (que consta do dia 6 de janeiro de 2021 na data da escrita desta dissertação). Vale ressaltar também que o sítio em questão traz testes de vocabulário com suporte para mais de 10 línguas diferentes, incluindo a língua portuguesa.

Os testes de vocabulário são disponibilizados de maneira totalmente gratuita e existem em duas categorias: receptivo e produtivo. O teste *receptivo* (ou *passivo*), como o próprio sítio denomina, é aquele que “faz referência às palavras que o aprendiz consegue entender ou aquelas que ele ou ela consegue deduzir ao usar estratégias” e , já o termo *produtivo* (ou *ativo*) ficou denominado como “vocabulário referente às palavras que o aprendiz está apto a usar ativamente”. O teste receptivo consta com um total de 150 palavras enquanto o produtivo tem um total de 90 palavras, ambos têm o mesmo tempo limite de 30 minutos. O teste receptivo, como o é destacado na página, geralmente reflete um alcance maior no conjunto de palavras do que o teste produtivo e por isso há uma diferença de 40 palavras entre os testes. Cada um dos dois testes contem 5 blocos de palavras divididos entre níveis, cada um baseado nas 1000, 2000, 3000, 4000 e 5000 palavras mais frequentes daquela língua.

A escolha do tipo de teste para este estudo foi baseada nas tarefas anteriores. Como os participantes tiveram que ativamente digitar a tradução de uma palavra então eles tiveram que produzir, fazendo assim com que a escolha fosse pelo teste *produtivo* de vocabulário. O sítio ainda informa que o teste produtivo consiste em um teste de 18 itens cloze por subtteste (níveis 1, 2, 3, 4 e 5 para cada 1000 palavras mais frequentes na língua) e que são fornecidas quantas letras necessárias para que o item a ser respondido não cause ambiguidade.

Ao finalizar o teste produtivo de vocabulário em língua inglesa o participante teria acesso ao resultado de imediato (em porcentagem e em número inteiro de itens corretos), este resultado deveria então ser enviado para o e-mail do pesquisador através de um *print* da tela ou que o participante guardasse o arquivo e enviasse da maneira que lhe fosse mais conveniente.

3.11 Tratamento dos dados

A princípio, o *software* Psytoolkit (Stoet, 2010, 2017) gerou arquivo tabulado no bloco de notas contendo os dados pré-selecionados (ID, palavra, tradução, tempo de leitura, tempo de resposta e acurácia) para cada participante. Para termos uma noção geral e mapear melhor a pesquisa, foi criada uma tabela no Microsoft Excel contendo os seguintes dados de cada participante:

- a) PARTICIPANTE – número de identificação atribuído na ordem que o participante terminava as tarefas, este número difere do código do participante que servia para identificação de eventuais anormalidades;
- b) ID – ordem (sempre aleatória) em que as palavras-estímulos apareceram para cada participante;
- c) PALAVRA – Aqui encontram-se as palavras escolhidas para compor os estímulos da tarefa de tradução (ver quadros 7, 8 e 9 no capítulo 3, seção 3.7);
- d) TRADUÇÃO – Opção de tradução dada como certa para a correção automatizada;
- e) TL (Tempo de Leitura) – Tempo que cada participante levou para ler o estímulo e pressionar a barra de espaço para iniciar a sua resposta;
- f) TR (Tempo de Resposta) – Tempo que cada participante levou para digitar a tradução do estímulo mostrado na tarefa;
- g) ACURÁCIA – Referente à precisão da resposta dos participantes, aqui era subdividida em 3 pontos (1 = resposta correta, 2 = resposta errada e 3 = tempo de resposta fora do limite)
- h) RESPOSTA – Aqui situa-se a resposta que cada participante deu para os estímulos da tarefa de tradução
- i) CONDIÇÃO – Dividida em 3 partes (EXP – condição experimental, ou os homógrafos interlinguísticos; CI – controle inglês, ou seja, palavras não homógrafas que tinham sua frequência Zipf relacionadas diretamente com as da condição experimental; e CP – controle português, condição em que a

frequência Zipf foi tratada pela frequência da tradução da palavra experimental para o português e depois versada para inglês²⁹);

- j) TR_NORMAL – Tempo de resposta normalizado para cada resposta³⁰; e
- k) ZIPF – Frequência na escala Zipf para cada palavra estímulo da pesquisa.

A primeira parte do tratamento de dados iniciou-se na correção do item “resposta”, isso é, todas as respostas que a correção automatizada destacou como erro (atribuição do numeral “2”) tiveram que ser manualmente investigadas uma por uma. Respostas com apenas erros de grafia eram mais fáceis de corrigir, e caso estivessem de acordo com a tradução seriam convertidas para corretas (atribuição do número “1”), a exemplo podemos pegar a palavra “vacuum” que foi respondida como “vaco”, “vaco” e “vacou” além da resposta registrada como correta “vácuo”, outro exemplo pode ser visto no estímulo “tire” que foi respondido como “pineu” e ainda assim todas estas respostas foram consideradas como correta pois a proposta do estudo estava focada na acurácia semântica e não ortográfica.

A segunda parte do tratamento exigiu muito mais esforço pois a correção passou da parte ortográfica para a semântica, a que de fato teria mais peso para os resultados estatísticos, o total de 2.640 itens foi apurado nesta etapa. Aqui a correção foi auxiliada por dicionários *online* como o *Merriam Webster*, *Google Tradutor* e dicionários com agregadores de significados como o *Reverso*. Pegando como exemplo a mesma palavra anterior “tire”, temos como opção certa de correção automática uma única resposta que era “pneu”, porém a tarefa não especificava a classe gramatical das palavras em questão, portanto todos os participantes que responderam com a forma verbal “cansar” também tiveram suas respostas convertidas para correta. O critério de avaliação nesta etapa naturalmente não pode ser imparcial, por mais que critérios fossem elaborados para dirimir a atuação do componente humano era inevitável um pouco de subjetividade na correção. Um critério utilizado era que a resposta não poderia ser um jargão muito específico para uma determinada área, como por exemplo a palavra “case” que tinha como resposta padrão “estojo” e foi traduzida como “capa”, esta resposta foi considerada incorreta por ser de uso específico (“capa” para aparelho celular), além de existirem outras palavras mais utilizadas para o substantivo “capa” como “cover” ou “cloak”. No geral, o critério adotado nesta parte do tratamento resumiu-se a “caso o participante respondesse com um uma

²⁹ Exemplificando a condição “CP”: Pega-se a palavra da lista experimental (ex.: *late*) e verifica a sua frequência zipf (5,4303), em seguida usa-se essa frequência para achar uma palavra em português (ex.: “esse”, zipf = 5,4309), e por final faz a versão da palavra para o inglês para montar um estímulo da tabela “CP” (*this*. zipf = 6,9013).

³⁰ A normalização foi feita de maneira simplificada, dividindo o número de caracteres digitado pelo tempo de resposta.

palavra que fosse semanticamente próxima, porém de pouco uso ou uso restrito a um contexto só, essa palavra não seria considerada como correta”; outro exemplo pode ser visto na palavra “ranger” que estava configurada corretamente como “guarda” e o participante que tenha traduzido como “policial” não teve sua resposta considerada como correta.

A terceira parte do tratamento dos dados ficou focada nos critérios técnicos do próprio *software* Psytoolkit (Stoet, 2010, 2017). Algumas das respostas tiveram sua acurácia designada com o número 3, isto é, o participante havia esgotado o tempo para responder tal item, porém ao observar os tempos de resposta descobrimos alguns destes itens estavam marcados incorretamente pela correção automática. Um bom exemplo aconteceu com a palavra “well” que estava designada como tradução padrão “bem”, porém o participante que escrevesse “poço” também estaria correto; o participante P14 respondeu a este item com “bom” e demorou um total de 5781 milissegundos, e ainda assim a acurácia dele estava como tempo esgotado (3) e não como resposta errada (2). Como as respostas fora do tempo não entraram para a análise estatística então excluí-las da tabela de dados, ou adicionar como foi o caso, alteraria o resultado das análises.

Tendo feito todas as correções com a maior cautela possível para evitar exageros ao corrigir as respostas dos participantes, temos uma planilha com um total de 1392 respostas corretas (1), 1224 incorretas (2) e 24 fora do tempo (3) totalizando 2640 respostas ($n = 22$, sendo assim: 22 [participantes] * 120 [estímulos] = 2640 [respostas individuais]). Para a análise feita no RStudio a tabela final desconsiderava as 24 respostas obtidas fora do tempo, e o montante final de observações ficou em 2616 respostas individuais divididas assimetricamente entre os participantes.

4 RESULTADOS

Neste capítulo, o presente estudo expõe a análise aprofundada dos resultados da pesquisa com homógrafos interlinguísticos entre as línguas inglesa e portuguesa, buscando compreender as nuances pragmáticas que permeiam esses termos compartilhados. No âmbito deste estudo, desvelamos algumas das complexidades inerentes à tradução e interpretação desses homógrafos, mais precisamente no que concerne à acurácia e os tempos de reação, sempre atentando às distintas estruturas linguísticas que caracterizam cada idioma e levando em consideração como se dá o processo do acesso lexical. A pesquisa adentrou nas peculiaridades semânticas dos homógrafos, explorando como suas acepções podem variar ou convergir nas duas línguas, bem como examinou as implicações pragmáticas dessas divergências.

Além disso, devemos enfatizar que o escopo deste estudo se delimita à análise comparativa dos tempos de resposta e da acurácia associados aos homógrafos interlinguísticos, em contraste com palavras controle cuidadosamente selecionadas. Por meio de experimentos controlados e metodologias rigorosas, avaliamos como falantes (bilíngues) nativos do português brasileiro (L1) e com conhecimento de inglês (L2) processam cognitivamente esses homógrafos em tarefa específica de tradução. Ao incluir palavras controle que compartilham similaridades morfológicas e frequência na escala Zipf, mas não apresentam a ambiguidade intrínseca dos homógrafos, buscamos verificar se as diferenças nos tempos de resposta e na acurácia podem ser atribuídas especificamente à natureza interlinguística desses termos ambíguos. Este enfoque metodológico visa oferecer uma análise abrangente da interpretação cognitiva de homógrafos interlinguísticos, destacando os pormenores que permeiam a compreensão desses termos em contextos de tradução de palavras.

A análise detalhada dos resultados de acurácia e os tempos de resposta associados aos homógrafos interlinguísticos e às palavras controle foi minuciosamente explorada e será demonstrada nas seções subsequentes desta dissertação. A próxima seção abordará a visualização e interpretação dos dados estatísticos descritivos referentes à acurácia, seguida por outra seção que nos mostrará os dados dos tempos de reação para cada condição e por fim, a análise estatística inferencial.

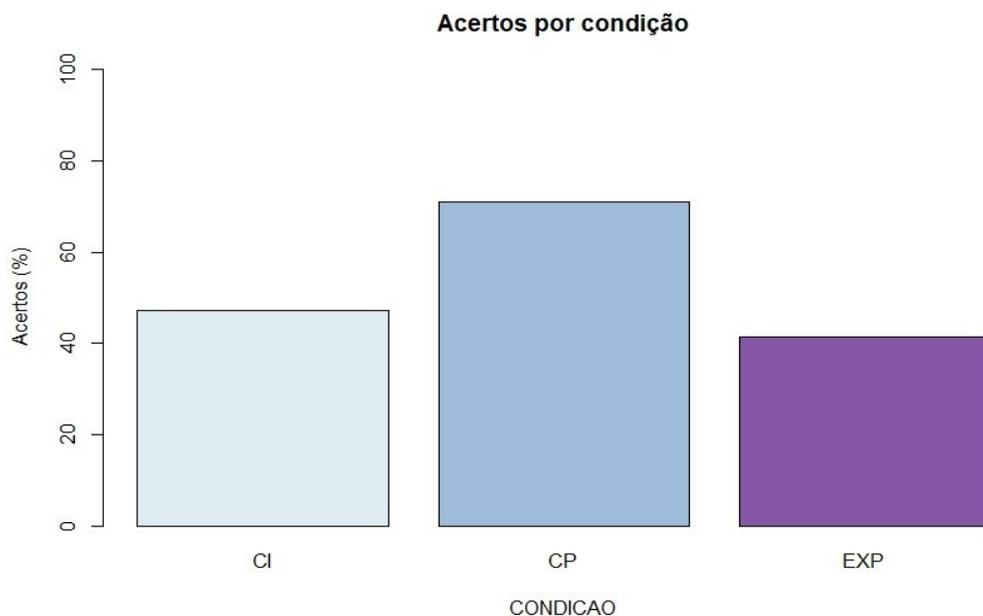
4.1 Análise estatística descritiva: Acurácia

A fase inicial da análise estatística evidencia a avaliação comparativa da acurácia entre as condições dos homógrafos interlinguísticos e das palavras controle. Utilizamos o modelo generalizado misto como teste estatístico para investigar as diferenças significativas, quando existentes, no desempenho dos participantes em relação a essas duas categorias de palavras (homógrafos interlinguísticos e não homógrafos).

Os dados de tempo de leitura, tempo de resposta e acurácia foram captados de forma automática pelo *software* Psytoolkit (Stoet, 2010, 2017) e posteriormente inseridos na planilha do Excel. Com todos os dados da planilha organizados fez-se necessário o tratamento deles para cada resposta de cada participante, esta etapa foi realizada manualmente e será escrutinada em seção posterior.

Com a tabela de dados pós-tratamento, utilizamos o *software* RStudio para criar os *scripts* e gerar os dados estatísticos referentes a acurácia entra as condições da pesquisa. Vejamos então a porcentagem de acurácia para cada uma das três condições no gráfico 1 e o valor exato na tabela 2.

Gráfico 1 – Acurácia por condição



Fonte: Elaborado pelo autor

Legenda: Controle pela frequência em inglês (CI); Controle pela frequência da tradução em português (CP), Grupo experimental (EXP)

Tabela 2 – Porcentagem da acurácia por condição

	CI	CP	EXP
1	47,18	70,85	41,51
2	52,81	29,14	58,48

Fonte: Elaborado pelo autor

Legenda: Controle pela frequência em inglês (CI); Controle pela frequência da tradução em português (CP), Grupo experimental (EXP)

Com a exclusão das respostas fora do tempo ficamos apenas com duas condições verificáveis, tornando assim uma escolha binária. Foi atribuído o número “1” para a palavra que estivesse correta e o número “2” para a palavra errada. Ao verificar a acurácia das respostas de acordo com a condição dos estímulos (“1” ou “2”) foi possível constatar uma vantagem para a condição controlada “CP”, na qual o número total de acertos foi de 620 (70,85%), em segundo lugar podemos observar o grupo controle “CI” com um total de 410 (47,18%) acertos, e por último ficou o grupo experimental “EXP” dos homógrafos interlinguísticos com um total de 362 (41,51%) respostas corretas. Em valores percentuais a condição experimental apresentou um valor aproximadamente 5,67% menor de acertos em comparação à condição diretamente controlada “CI”.

Com base nos resultados apresentados da acurácia em relação à condição avaliada, podemos observar um efeito de interferência dos homógrafos interlinguísticos em relação às palavras não homógrafas. Vejamos na próxima seção os valores dos resultados referentes ao tempo de reação (TR) para cada condição.

4.1.2 Análise estatística inferencial: Acurácia

A análise estatística inferencial dos dados realizada nesta pesquisa abrangeu o uso do modelo generalizado misto para avaliar a relação entre as variáveis de interesse. Inicialmente, uma análise exploratória foi conduzida, incluindo a visualização da distribuição dos dados e tabelas descritivas. Em seguida, o modelo de regressão linear foi empregado no RStudio com o pacote “lme4” (Bates *et al.*, 2015) para examinar as associações entre as variáveis dependentes e independentes, considerando também efeitos aleatórios para acomodar a estrutura hierárquica dos dados. Os resultados desse modelo foram interpretados, considerando coeficientes estimados, significância estatística e medidas de ajuste do modelo. A utilização do modelo misto generalizado forneceu uma compreensão mais aprofundada das

relações pertencentes aos dados coletados, permitindo *insights* significativos que serão abordados capítulo de discussão como a exploração de resultados particulares por estímulo.

Após os dados da tabela geral do estudo serem tratados, utilizamos o RStudio para rodar o modelo generalizado misto com a função *link* de regressão linear, para tal utilizamos a fórmula mais simplificada “<- glmer(data = d2, ACURACIA ~ CONDICAO + (1|PARTICIPANTE), family=binomial(link = "logit"))” a fim de observar os efeitos aleatórios para os participantes. Ao incluir na fórmula um outro modelo mais complexo com a adição dos efeitos aleatórios para as palavras (ID) ficamos com a fórmula “<- glmer(data = d2, ACURACIA ~ CONDICAO + (1|PARTICIPANTE) + (1|ID), family=binomial(link = "logit"))”. Um teste ANOVA entre os modelos mostrou que o modelo mais complexo (segundo) obteve evidências estatísticas mais significativas para explicar a variação da acurácia do que o modelo simplificado. Vejamos o modelo na tabela 3 a seguir.

Tabela 3 – Modelo generalizado de efeitos mistos (acurácia por condição)

<i>Preditores</i>	Acurácia por condição		
	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(EXP)	2.10	1.32 – 3.35	0.002
CI	0.50	0.33 – 0.77	0.002
CP	0.15	0.10 – 0.23	<0.001
Efeitos Aleatórios			
σ^2	3.29		
τ_{00} ID	1.73		
τ_{00} PARTICIPANTE	0.42		
ICC	0.40		
$N_{PARTICIPANTE}$	22		
N_{ID}	120		
Observações	2616		
R^2 Marginal / R^2 Condicional	0.102 / 0.458		

Fonte: Elaborado pelo autor

Nesta tabela do modelo apresentado, podemos notar que a diferença da acurácia entre as condições foi significativa, pois o valor *p* de todas as condições situou-se em um patamar de relevante significância ($p < 0,05$) sendo a condição CP significativa. É visível também o decréscimo na razão de possibilidades (odds ratios) quando passamos da condição

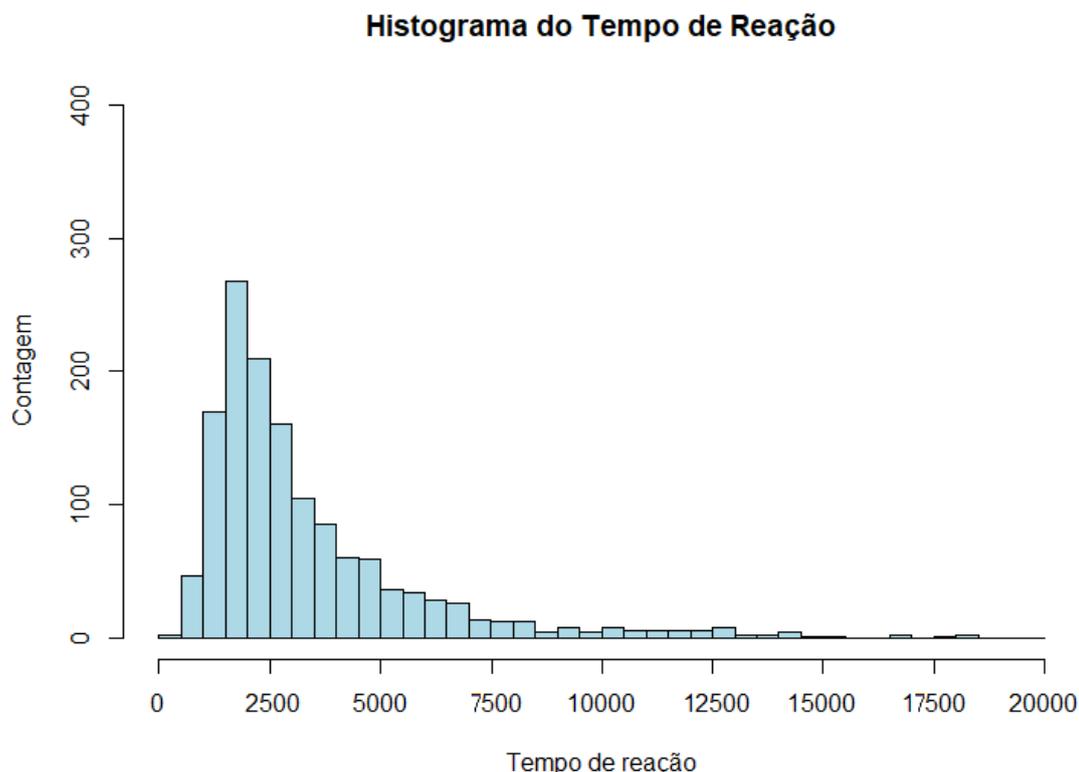
EXP para CI ou CP, reforçando a hipótese do efeito de interferência em acurácia dos homógrafos interlinguísticos na tarefa de tradução.

Os efeitos aleatórios mostram pouca variação entre participantes nas 2616 iterações observadas. O R^2 Marginal que leva em consideração a variância total dos efeitos fixos atingiu uma pontuação de 0.102, explicando assim 10% destes efeitos, já o R^2 condicional que nos mostra a variância total após os efeitos fixos e aleatórios apresentou uma pontuação de 0.458, explicando assim aproximadamente 46% das interações.

4.2 Análise estatística descritiva: Tempo de reação

Para ter uma base mais fundamentada acerca da diferença de tempo de processamento entre os estímulos do estudo analisamos além da acurácia, o tempo de reação (TR) que cada participante levou para responder cada um dos estímulos. Vejamos então o histograma da distribuição dos tempos de reação perante os 120 estímulos da pesquisa (GRÁFICO 2).

Gráfico 2 – Histograma do tempo de reação da tarefa de tradução



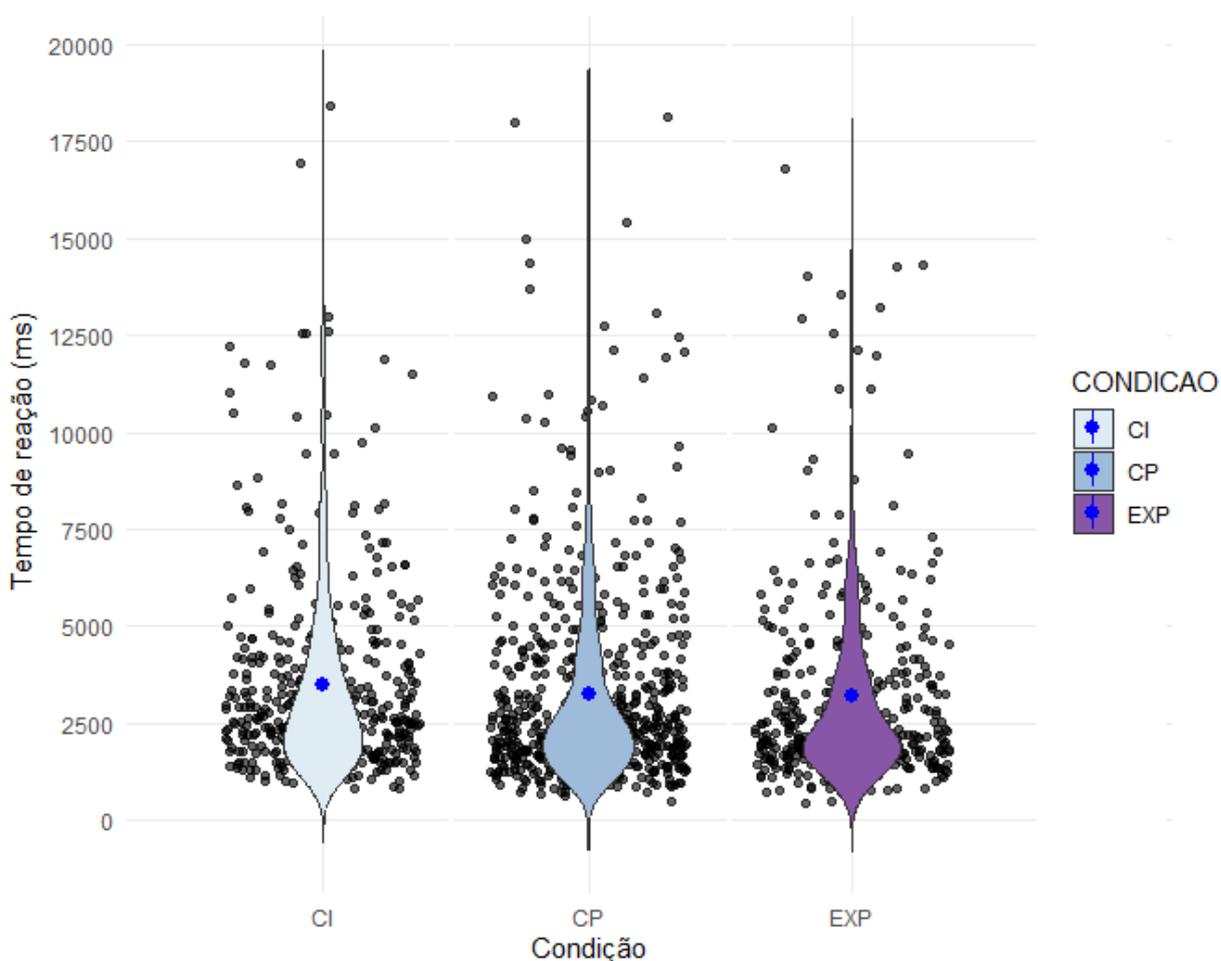
Fonte: Elaborado pelo autor

Legenda: Cada barra está separada na ordem de 500 milissegundos

Como visto no histograma, a maioria das respostas aos estímulos estão situadas entre 1500 e 2000 milissegundos. Houve uma predominância de tempo de reação até a faixa dos 2000 milissegundos e a partir deste ponto a quantidade de respostas com tempo superior foi diminuindo cada vez mais até chegar ao limite de 20.000 milissegundos. O total de respostas corretas na faixa mais ampla (entre 1500 e 2000ms) foi de 478 (34,33%), o que ficou ligeiramente acima de um terço do total dos itens corretos. O histograma nos mostra um padrão geral das respostas dos participantes em todas as condições, porém é necessária uma análise por condição para dissertar melhor sobre os resultados do grupo experimental.

Com a visualização do gráfico 3 podemos ver a dispersão das respostas por cada condição. Neste gráfico de diagrama em nuvem de chuva (*raincloud plot*) podemos notar pontos fora da curva (*outliers*) em todas as condições.

Gráfico 3 – Diagrama de nuvem de chuva do tempo de reação (TR) da tarefa de tradução por cada condição



Fonte: Elaborado pelo autor

Legenda: Controle pela frequência em inglês (CI); Controle pela frequência da tradução em português (CP), Grupo experimental (EXP).

Assim, após verificar as relações com as outras medidas de tempo, ficou evidente que a condição experimental (EXP) apresentou os maiores valores em todos os tempos de reação, seguida pelo controle em inglês (CI) e por último com os menores tempos de reação ficou a condição controlada pelo português (CP). Vejamos na tabela 4 a distribuição exata para cada condição.

Tabela 4 – Tempo de reação por condição da tarefa de tradução

Condição	Média	Mediana	DP	TR Mínimo	TR Máximo	Variação
EXP	3207,08	2347,5	2524,68	432	16798	16366
CI	3517,81	2732,5	2532,08	839	18430	17591
CP	3267,37	2436,5	2518,95	512	18110	17598

Fonte: Elaborado pelo autor.

Legenda: DP = Desvio Padrão. Número total de participantes = 22.

A tabela 4 nos mostra algumas diferenças entre as condições e os valores analisados. A condição experimental apresentou os menores valores para a média e mediana, já o desvio padrão da condição experimental ficou entre as duas outras condições. O tempo de resposta mínimo obteve o menor valor na condição experimental (EXP) porém este deve ser examinado com cautela pois o participante respondeu com a tecla “9” ao estímulo “nine” fazendo assim com que o tempo de digitar “nove” diminuísse drasticamente, na contagem com o tempo de resposta normalizado (TR normal) a condição “CP” foi a que obteve o menor valor para o TR mínimo. A variação (range) da condição experimental apresentou o menor valor (16366), o que pode ser explicado também pela menor quantidade de respostas corretas para esta condição, já as outras duas condições (CI e CP) apresentaram valores de variação praticamente iguais, com uma diferença de apenas 7 milissegundos, tempo este irrelevante para a pesquisa em questão.

4.2.1 Análise estatística inferencial: tempo de reação

A análise estatística inferencial do tempo de reação (TR) proporcionou uma investigação detalhada das relações entre as variáveis preditoras e a variável resposta. Após uma exploração inicial dos dados, empregamos modelos de regressão linear e técnicas de modelagem mista para avaliar como as condições experimentais influenciam o TR. A utilização de modelos mistos permitiu incorporar tanto efeitos fixos quanto aleatórios, considerando a estrutura hierárquica dos dados. Os coeficientes estimados e os resultados dos testes de significância foram interpretados para compreender a proporção dessas influências. Essa

abordagem visa não apenas descrever as diferenças médias no TR entre condições, mas também considerar a variabilidade entre participantes e possíveis interações. Os resultados desta análise contribuem para uma compreensão mais profunda dos fatores que moldam o tempo de reação, fornecendo *insights* valiosos que serão abordados posteriormente da parte de discussões dos resultados, pois ao analisar o estudo incluindo as respostas erradas obtivemos resultados divergentes dos achados nesta seção.

Utilizando o software RStudio compilamos todos os dados da tabela tratada e utilizamos a fórmula “`lmer(data = d3, TR ~ CONDICAO + (1|PARTICIPANTE) + (1|ID))`” após testá-la e verificar que o modelo mais complexo, assim como na análise da acurácia, apresentava resultados mais significativos. Ao rodar a fórmula obtivemos os dados mostrados na tabela 5 a seguir.

Tabela 5 – Modelo linear de efeitos mistos (TR por condição)

<i>Preditores</i>	<i>Estimativa</i>	TR por condição	
		<i>CI</i>	<i>p</i>
EXP	3371.73	2732.82 – 4010.64	<0.001
CI	323.14	-157.18 – 803.46	0.187
CP	98.70	-340.42 – 537.82	0.659
Efeitos Aleatórios			
σ^2	4420243.43		
τ_{00} ID	591922.71		
τ_{00} PARTICIPANTE	1584778.29		
ICC	0.33		
$N_{PARTICIPANTE}$	22		
N_{ID}	116		
Observações	1392		
R^2 Marginal / R^2 Condicional	0.002 / 0.332		

Fonte: elaborado pelo autor

A condição EXP foi atribuída como o intercepto da função (*baseline*) então a estimativa para o tempo de reação dos homógrafos interlinguísticos ficou em 3371.73. O valor *p* mostrou resultado significativo apenas para a condição EXP, já o intervalo de confiança (*CI*) apresentou valores negativos para os dois grupos controlados (CI e CP) o que pode ser

interpretado como uma incerteza do modelo para essas condições em relação ao TR da condição experimental e às respostas corretas aos estímulos. Para a condição CI há uma variação considerável nos dados e, com base na análise, não podemos afirmar com confiança se o efeito da condição é positivo ou negativo e para a condição CP, da mesma forma, a variação nos dados não nos permite afirmar com confiança se o efeito da condição é positivo ou negativo. Uma possível explicação para esses índices está na assimetria dos dados, visível na cauda alongada para a esquerda no histograma dos tempos de resposta, outra explicação pode ser dada a alta variância dos efeitos aleatórios. A variação para os efeitos fixos (R^2 marginal) ficou em 0.002, o que nos mostra um efeito de 2% destes para explicar a variação da nossa variável dependente e para a análise pós efeitos aleatórios (R^2 condicional) ficou em 0.332, ou 33,2% da possível variação da dependente variável.

Com o resultado estatístico das condições (acurácia e tempo de resposta) não podemos concordar totalmente com a hipótese H1 da pesquisa que previa um efeito de interferência da condição experimental (homógrafos interlinguísticos) em relação à condição controle (não homógrafos). Tanto o grupo controlado pela frequência quanto a adição de um outro grupo controle, independente da frequência da palavra, mostraram resultados significativos em relação a condição experimental apenas para a acurácia e não para os tempos de resposta. A hipótese H2 também não pode ser plenamente confirmada pois o grupo EXP apesar de ter mostrado menores valores para os TR apresentou ineficácia para a acurácia em relação aos outros dois grupos. Abordaremos as análises exploratórias do estudo na próxima seção de discussões para compreender o fator da exclusão das respostas erradas na análise dos TR e como isso afeta a interpretação dos dados, além de interpretar também os resultados com a inclusão das respostas incorretas.

4.3 Discussão dos resultados

Nesta seção, discutimos detalhes dos resultados obtidos na análise estatística, explicando as particularidades reveladas pelos modelos aplicados às variáveis de interesse. Inicialmente, concentraremos nossa atenção nas conclusões extraídas da análise de *acurácia*, que trouxe à luz padrões distintos de desempenho sob diferentes condições experimentais. Em seguida, voltaremos nossa atenção para as implicações derivadas da análise do *tempo de reação* (TR), oferecendo uma visão aprofundada das dinâmicas observadas às variáveis estudadas e terminaremos a sessão com a análise do tempo de reação contabilizando todas as respostas (corretas e incorretas). Ao discutir cada resultado, consideraremos a significância estatística, as

particularidades que podem influenciar os efeitos e as possíveis interações entre os fatores analisados. Este enfoque busca contextualizar os achados dentro do arcabouço teórico e fornecer entendimento valioso para futuras investigações.

4.3.1 A acurácia na tarefa de tradução

Para compreender melhor a proporção de acertos na tarefa de tradução, podemos associar parcialmente os resultados observados ao efeito de frequência da palavra. No início das pesquisas em processamento/tradução de palavras acreditava-se que a frequência seria uma variável preditora quase incontestável para o processamento de palavras (Preston, 1935) porém novas técnicas de medição e corpora mais robustos e atualizados mostraram que a diversidade contextual tem seu peso significativo e foi considerado até sendo mais preditor do que a própria frequência de palavra (Adelman; Brown e Quesada, 2006; Toassi *et al.*, 2023).

Reiteramos essa explicação parcial do efeito de frequência de palavras, porém a condição experimental dos homógrafos interlinguísticos teve uma significância maior nos resultados do que a própria frequência da palavra, a média da frequência Zipf dos estímulos da condição EXP era tecnicamente igual à da condição CI ($\bar{X} f_{Zipf}[EXP] = 4,0285$ e $\bar{X} f_{Zipf}[CI] = 4,0814$) portanto a frequência sozinha não poderia explicar a divergência encontrada. Outro fator que coaduna ainda mais para a aceitação da hipótese H1 é ao verificar a média da frequência Zipf das traduções entre as condições EXP e CI, pois como o estudo tinha foco no participante bilíngue e a tarefa realizada seria uma tradução, logo a frequência da palavra em português brasileiro também poderia ter relevância pois ela é o “alvo” que o participante deve produzir. Para o nosso estudo obtivemos médias diferentes entre as condições ($\bar{X}^{tradução} f_{Zipf}[EXP] = 4,0812$, $\bar{X}^{tradução} f_{Zipf}[CI] = 3,7906$ e $\bar{X}^{tradução} f_{Zipf}[CP] = 4,0816$), ficando a condição CI com a menor média de todas, tanto na frequência do estímulo (inglês) quanto na frequência da tradução para o português brasileiro dos estímulo e, obstante a isso, obteve resultados de acurácia superiores à condição experimental.

4.3.2 O tempo de reação na tarefa de tradução

Como dito anteriormente, os estudos cognitivos em processamento de palavras consideram a frequência da palavra como a variável mais relevante, e predizem até que tanto a acurácia como o tempo de reação são reflexos dessa variável (Forster e Chambers, 1973). Em Balota *et al.* (2004) a frequência da palavra mostrou ser mais eficiente para explicar as variações

no tempo de reação do que o efeito de vizinhança e mostrou-se pareado com a condição do tamanho da palavra.

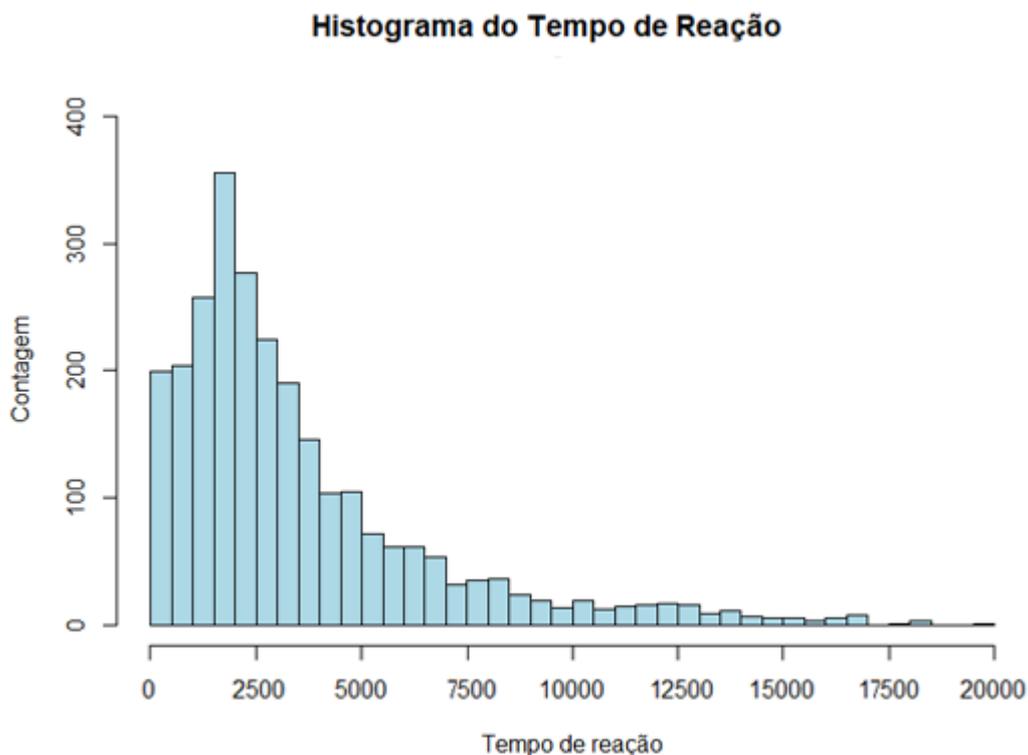
Nosso estudo isolou apenas a característica da frequência da palavra *versus* a condição (EXP, CI ou CP), o *tamanho da palavra* não foi utilizado como critério e por isso o tempo de reação foi normalizado de maneira simples dividindo o número de caracteres digitados no estímulo pelo tempo de reação, encontramos então resultados que corroboram mais uma vez com a análise estatística, a condição EXP teve uma média do TR_{normal} de aproximadamente 618 milissegundos por caractere digitado, enquanto as condições CI e CP tiveram respectivamente 498 e 497 milissegundos por caractere digitado.

Ao avaliar os resultados da variação de acurácia e dos tempos de reação em relação às condições, conseguimos concretizar os fatores que aceitam parcialmente a hipótese H1. Vejamos na próxima sessão como a adição das respostas incorretas revela um outro viés para as hipóteses e reforça ainda mais a aceitação da hipótese H1 como verdadeira.

4.3.3 Análise do tempo de reação com respostas incorretas

A completude da análise deste estudo se dá com a adição das respostas incorretas na análise dos tempos de reação para cada condição. Para averiguar o experimento em sua totalidade, as respostas incorretas foram inseridas na observação do tempo de reação em todas as condições e posteriormente as mesmas análises estatísticas (descritiva e inferencial) foram utilizadas para a análise dos resultados. Para começar, vejamos o histograma da distribuição dos tempos de reação perante os 120 estímulos da pesquisa no gráfico 4.

Gráfico 4 – Histograma do tempo de reação da tarefa de tradução (todas as respostas)



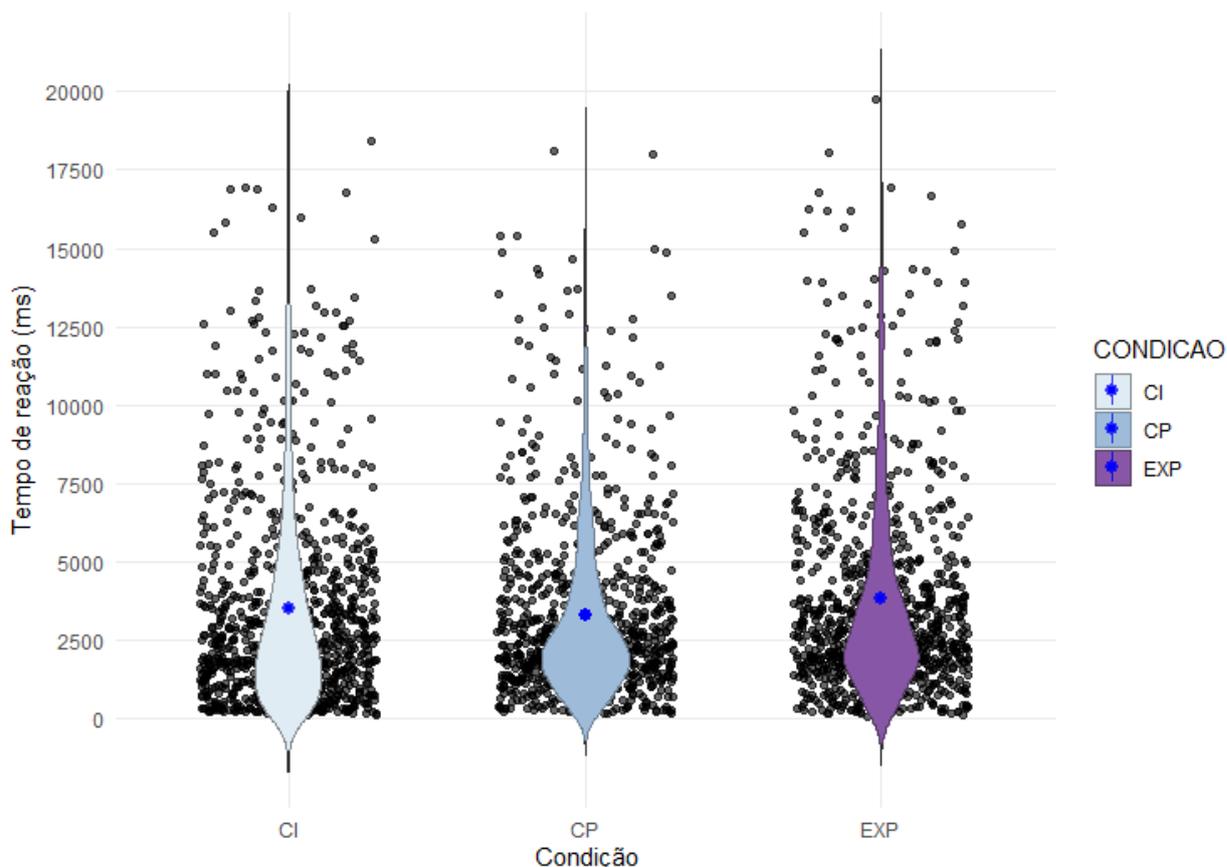
Fonte: Elaborado pelo autor

Legenda: Cada barra está separada na ordem de 500 milissegundos

Assim como no histograma que contabiliza apenas as respostas corretas (vide gráfico 2), a faixa em que podemos encontrar a maioria das respostas situa-se entre 2000 e 2500 milissegundos e o final da cauda de distribuição das respostas também segue o mesmo ritmo, vão decrescendo gradualmente até o limite de 20.000 milissegundos.

O histograma nos mostra apenas os resultados globais do tempo de reação. A condição que mostrou o maior ponto atípico (resultado fora da média) foi a condição experimental dos homógrafos interlinguísticos, além de ter apresentado também a maior variação em tempo de reação. Para visualizar melhor a distribuição do tempo de resposta por condição, vejamos também um gráfico em distribuição de nuvem de chuva com o TR separado por cada condição. Vejamos então no gráfico 5 a distribuição de cada resposta por cada condição no diagrama de nuvem de chuva.

Gráfico 5 – Diagrama de nuvem de chuva do tempo de reação (TR) da tarefa de tradução por cada condição



Fonte: Elaborado pelo autor

Legenda: Controle pela frequência em inglês (CI); Controle pela frequência da tradução em português (CP), Grupo experimental (EXP).

Assim, após verificar as relações com as outras medidas de tempo, ficou evidente que a condição experimental (EXP) apresentou os maiores valores em todos os tempos de reação, seguida pelo controle em inglês (CI) e por último com os menores tempos de reação ficou a condição controlada pelo português (CP). Vejamos na tabela 6 a distribuição exata para cada condição.

Tabela 6 – Tempo de reação por condição da tarefa de tradução (geral)

Condição	Média	Mediana	DP	Mínimo	Máximo	Varição
EXP	3797,857	2801	3276,614	63	19748	19658
CI	3472,625	2493	3249,063	64	18430	18366
CP	3261,688	2382	2835,955	127	18110	17983

Fonte: Elaborado pelo autor

A tabela 6 nos mostra algumas diferenças relevantes entre as condições e os valores analisados. A condição experimental apresentou os maiores valores para a média, mediana, desvio padrão, máximo e variação (range), ficando com uma média de 325,323 milissegundos acima da condição controle inglês o que reflete em uma diferença percentual de aproximadamente 8,5% entre as condições, o desvio padrão entre as condições EXP e CI não apresentou diferença relevante. A condição controle português (CP) apresentou os melhores resultados em todas as análises, acurácia e tempo de reação, possuindo também os menores valores em todas as situações, exceto o TR mínimo, e apresentou um desvio padrão bem menor que as outras duas condições.

O menor tempo de reação na condição experimental pode ser explicado por uma maior dificuldade de traduzir aquela palavra, pois como os participantes tinham 5 segundos para ler cada estímulo antes de responderem e haviam sido instruído a passar o mais rápido possível para a próxima caso não soubessem a tradução da palavra, é possível que esse fator explique um menor tempo de reação mínimo para esta condição. Este resultado pode ser observado de maneira antagônica para a condição em que os indivíduos tiveram a melhor acurácia e os menores tempos de resposta (CP), em que o menor tempo de reação obteve o maior valor dentre as condições. Por uma observação análoga ao exposto sobre o tempo mínimo de reação, os valores do tempo máximo de reação também respeitaram a mesma regra.

Para realizar a estatística inferencial, recorreremos novamente ao software RStudio para compilar todos os dados da tabela tratada com todas as respostas e utilizamos a fórmula “`lmer(data = d2, TR ~ CONDICAO + (1|PARTICIPANTE) + (1|ID))`” por assumir que a distribuição dos valores seja normal entre as condições. Ao executar a fórmula obtivemos os dados mostrados na tabela 7.

Tabela 7 – Modelo linear de efeitos mistos (TR por condição - geral)

<i>Preditores</i>	TR por condição		
	<i>Estimativa</i>	<i>CI</i>	<i>p</i>
EXP	3782.14	3222.98 – 4341.30	< 0.001
CI	-344.51	-706.25 – 17.22	0.062
CP	-457.14	-815.53 – -98.76	0.012
Efeitos Aleatórios			
σ^2	8168375.68		
τ_{00} ID	309427.70		
τ_{00} PARTICIPANTE	1383847.72		
ICC	0.17		
$N_{\text{PARTICIPANTE}}$	22		
N_{ID}	120		
Observações	2616		
R^2 Marginal / R^2 Condicional	0.004 / 0.175		

Fonte: elaborado pelo autor

A condição EXP foi atribuída como o intercepto da função (*baseline*) então a estimativa para o tempo de reação dos homógrafos interlinguísticos ficou em 3782.14. A condição CI apresentou uma estimativa de TR 344.51 pontos abaixo da EXP, e a condição (CP) que havia apresentado os menores valores de TR ficou com uma estimativa de 457.17 pontos abaixo da condição EXP. O valor *p* mostrou resultados significativos para as condições EXP e CP, e um valor próximo do significativo para a condição CI, já o intervalo de confiança (*CI*) apresentou valores negativos para os dois grupos controlados (CI e CP) o que pode ser interpretado como uma incerteza do modelo para essas condições em relação ao TR da condição experimental e às respostas corretas aos estímulos.

Parte dessa incerteza pode ser atribuída ao alto índice de respostas incorretas nas condições EXP e CI em contrapartida ao elevado número de respostas corretas na condição CP. A variação para os efeitos fixos (R^2 marginal) ficou em 0.004, o que nos mostra um efeito de 4% destes para explicar a variação da nossa variável dependente, e para a análise pós efeitos aleatórios (R^2 condicional) ficou em 0.175, ou 17,5% da possível variação da dependente variável.

Com o resultado estatístico das condições (acurácia e tempo de resposta) podemos concordar com a hipótese H1 da pesquisa que previa um efeito prejudicial da condição

experimental (homógrafos interlinguísticos) em relação à condição controle (não homógrafos). A adição de um outro grupo controle, independente da frequência da palavra, também mostrou resultados significativos para efeito de interferência da condição experimental. A hipótese H2 fica então descartada com a introdução da análise geral do tempo de resposta.

No próximo capítulo discorreremos sobre as considerações finais do estudo e falamos sobre possíveis implicações do achado para o mundo acadêmico e profissional.

5. CONSIDERAÇÕES FINAIS

Ao abordarmos as considerações finais deste estudo, renovamos o escopo que foi investigar por um viés cognitivo o efeito dos homógrafos interlinguísticos numa tarefa de tradução. Os resultados derivados das análises estatísticas fornecem uma compreensão mais precisa e fundamentada das relações entre as variáveis de interesse. Este estudo contribui tanto para a esfera teórica quanto para possíveis aplicações práticas, destacando implicações subjetivas derivadas das observações. Ao refletir sobre essas descobertas, é essencial contextualizá-las em relação às teorias existentes já demonstradas e considerar sua relevância para a literatura científica. As considerações finais servirão como base sólida para futuras discussões e investigações, delineando caminhos potenciais para o avanço do conhecimento nesta área específica de pesquisa e correlatas. Iniciaremos esta seção com as delimitações inerentes à pesquisa, depois abordaremos algumas implicações dos resultados obtidos e por fim mostramos caminhos para futuras pesquisas na área.

A começar pelas limitações, os participantes observados apresentaram um padrão distinto na tarefa final (conhecimento produtivo de vocabulário), logo este padrão indicava que haviam 2 grupos dissemelhantes em proficiência na língua inglesa. A média de acertos nesta tarefa ficou em torno de 46% e a mediana também retornou o mesmo valor, indicando que a pesquisa tinha praticamente um grupo “acima da média” e outro abaixo. É provável que o tipo de teste (produtivo) tenha contribuído para estes resultados pouco favoráveis devido ao grau de dificuldade inerente da tarefa. Outra limitação fora a escolha das palavras, a língua portuguesa difere em vários aspectos da língua inglesa, e a nossa ortografia tem uma variabilidade maior quando se trata de verbos (flexão verbal) além disso, os substantivos em português possuem gênero enquanto no inglês essa não é uma característica evidenciada, o que leva a palavras como “holy” terem a possibilidade de duas traduções, “sagrado” ou “sagrada”.

Durante a análise individual do tratamento dos dados de cada participante (n = 22) foi possível perceber também uma ativação do componente fonético na tradução de determinadas palavras, a exemplo podemos citar os estímulos “nose” que foi traduzido como “nariz”, “deft” traduzido como “surdo”, “noon” como “freira”, “broth” traduzido como “trouxe”, “ache” como “cinza”, “role” como “buraco”, “aid” como anúncio/propaganda dentre outros que seguiram essa mesma lógica de ter uma ortografia e ou fonética semelhante a outras palavras da língua inglesa. Sousa³¹ (em conclusão) abordará o tema de homófonos

³¹ Dissertação em conclusão intitulada até então de: *Portuguese-English interlingual homophones in word reading and translation*. Previsão de conclusão para 2024.

interlinguísticos com bilíngues do par inglês-português brasileiro e os resultados podem demonstrar interações de relevância significativa para o fato observado.

A escolha das palavras é um fator limitante crucial pois ela carrega consigo uma gama de variáveis que devem ser controladas ao máximo para impedir uma aleatoriedade exagerada na pesquisa. Outro caso observado durante o tratamento fora a co-ativação das duas línguas, isto é, a L1 e a L2 estavam competindo durante a tarefa de tradução pois alguns estímulos eram “traduzidos” ou melhor dizer, “versados” do português (o participante assumiu indiscriminadamente que a palavra estava em português) para o inglês como em “doze” que virou “twelve”, “chute” foi traduzido como “kick” e “balance” ficou como “shake”, e a palavra “ache” apareceu mais uma vez agora versada como “find”. este efeito da não-seletividade também pode ser visto em Toassi *et al.* (2023).

Muitos são os contextos que podemos analisar os homógrafos interlinguísticos, temos por exemplo Freitas (2023) que fez sua análise contextualizada em frases e em palavras isoladas para reconhecimento de tradução e o resultado do custo de processamento diferiu entre as duas situações. Os resultados de nossa pesquisa também concordam com o observado em Lameira *et al.* (2023) que utilizou análise de Potenciais de Eventos Relacionados numa tarefa de julgamento semântico de homógrafos interlinguísticos com multilíngues (português brasileiro, inglês e alemão) e observou uma queda de acurácia e aumento nos tempos de reação para a condição experimental.

Na seara da educação, podemos utilizar os resultados da pesquisa de maneira didática. Professores, por exemplo, podem utilizar-se dessa característica do efeito de interferência dos homógrafos interlinguísticos para modelar textos de acordo com o nível observado dos alunos em sala de aula, a fim de dificultar (com a adição dos homógrafos) ou facilitar (não utilizando os homógrafos) a compreensão de um determinado texto, a sistemática da aula deve levar em consideração as particularidades de cada obstáculo que possa acarretar num maior esforço cognitivo para o aluno. No contexto da tradução profissional, os homógrafos interlinguísticos podem complicar durante a parte da pós-edição pois o tradutor nem sempre tem acesso ao texto de partida, o que dificulta e toma tempo para interpretar ou compreender aquela palavra que pode ter sido mal traduzida no texto (informação verbal)³².

Apesar das características biográficas dos participantes do estudo apresentarem resultados semelhantes, um novo estudo com a divisão entre níveis de proficiência e um maior número de participantes poderia suceder estatísticas mais significativas, podendo ser uma

³² Informação fornecida pela prof^a Dr^a. Diana Costa Fortier Silva em 26 fev. 2024.

proposta de pesquisa a divisão dos participantes por nível aferido no teste de vocabulário da língua inglesa. Todas as observações ficam como alerta para estudos futuros, que podem agregar também outras métricas e meios de intervenções como o ERP, EEG ou até mesmo o método de rastreamento ocular (*eye-tracking*) que tem atualmente uma grande capacidade de fornecer dados robusto e precisos (devido aos avanços tecnológicos no dispositivo) sobre tempo de reação, padrões de leitura e uma sucessão de outras variáveis que podem impactar de maneira positiva a produção científica na área.

Este estudo objetivou verificar o processo tradução dos homógrafos interlinguísticos de maneira isolada e compará-los com a tradução de palavras não homógrafas. O efeito de interferência foi evidenciado parcialmente ao contabilizar apenas as palavras traduzidas corretas e de maneira total ao contabilizar todas as tentativas de tradução. Espera-se que este estudo advirta para como as diferentes classes de palavras podem interferir no processamento de uma tradução e promova futuras pesquisas sob a ótica da psicolinguística acerca do processamento de palavras/frases do falante bilíngue.

REFERÊNCIAS

- ADELMAN, J. S.; BROWN, G. D. A.; QUESADA, J. F. Contextual diversity, not word frequency, determines word-naming and lexical decision times. **Psychological science**, v. 17, n. 9, p. 814–823, 2006.
- AHRENBERG, L. Comparing machine translation and human translation: A case study. In: **RANLP 2017: The First Workshop on Human-Informed Translation and Interpreting Technology (HiT-IT)**. Association for Computational Linguistics, 2017. p. 21-28.
- ALVES, F., MAGALHÃES, C.M., PAGANO, A. **Traduzir com autonomia: estratégias para o tradutor em formação**. São Paulo: Contexto, 2000.
- ALVSTAD, C.; TISELIUS, E.; HILD, A. **Methods and strategies of process research: Integrative approaches in Translation Studies** [Benjamins Translation Library 94]. Amsterdam: John Benjamins, p. 1-390, 2011.
- ARROJO, Rosemary (org.). **O signo desconstruído: implicações para a tradução, a leitura e o ensino**. Campinas: Pontes, 2003.
- ARROJO, Rosemary. **Tradução, desconstrução e psicanálise**. Rio de Janeiro: Imago, 1993.
- AUBERT, F.H. **As (in)fideliades da tradução: servidões e autonomia do tradutor**. Campinas, Campinas, SP: Editora da UNICAMP, 1993.
- AVAILABLE ONLINE. **About the vocabulary tests - Institut für Testforschung und Testentwicklung e.V. (ITT)**. Disponível em: <https://itt-leipzig.de/about-the-vocabulary-tests-2-2/?lang=en>. Acesso em: 16 jan. 2024.
- BALOTA, D. A. et al. Visual word recognition of single-syllable words. **Journal of experimental psychology. General**, v. 133, n. 2, p. 283–316, 2004.
- BATES, D. et al. (2015). Fitting linear mixed-effects models using lme4. **Journal of Statistical Software**, 67, 1-48.
- BERNARDINI, S. Think-aloud protocols in translation research: Achievements, limits, future prospects. **Target. International Journal of Translation Studies**, v. 13, n. 2, p. 241-263, 2001.
- BOUKRERIS, L.; LABED, Z. **Human vs Machine Translation: A Comparative Study**. 2016.
- BRYSSBAERT, M. et al. The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. **Journal of experimental psychology. Human perception and performance**, v. 42, n. 3, p. 441–458, 2016.
- BRYSSBAERT, M. et al. The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German: A review of recent

developments and implications for the choice of frequency estimates in German. **Experimental psychology**, v. 58, n. 5, p. 412–424, 2011.

BRYLSBAERT, M.; BORIS NEW. Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. **Behavior research methods**, v. 41, n. 4, p. 977–990, 2009.

BURGESS, C.; LIVESAY, K. The effect of corpus size in predicting reaction time in a basic word recognition task: Moving on from Kučera and Francis. **Behavior Research Methods, Instruments, & Computers**, v. 30, p. 272–277, 1998.

BUTLER, Y.G.; HAKUTA, K. Bilingualism and second language acquisition. **The handbook of bilingualism**, p. 114-144, 2006.

CADWELL, P.; O'BRIEN, S.; TEIXEIRA, C.S.C. Resistance and accommodation: factors for the (non-) adoption of machine translation among professional translators. **Perspectives**, v. 26, n. 3, p. 301-321, 2018.

CANTOR, Jacob. **An Objective Psychology of Grammar**. 1936.

CHESTERMAN, A. Translation, epistemology and cognition. In: **The Routledge handbook of translation and cognition**. New York: Taylor And Francis, 2020.

CHO, K. et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. **arXiv preprint arXiv:1406.1078**, 2014.

CHYUNG, S. Y. Y. et al. Evidence-based survey design: The use of a midpoint on the likert scale. **Performance improvement**, v. 56, n. 10, p. 15–23, 2017.

COSTA, A.; CARAMAZZA, A.; SEBASTIAN-GALLES, N. The cognate facilitation effect: implications for models of lexical access. **Journal of experimental psychology. Learning, memory, and cognition**, v. 26, n. 5, p. 1283–1296, 2000.

DIJKSTRA, T. et al. More on interlingual homograph recognition: language intermixing versus explicitness of instruction. **Bilingualism (Cambridge, England)**, v. 3, n. 1, p. 69–78, 2000.

DIJKSTRA, T. et al. Multilink: A computational model for bilingual word recognition and word translation. **Bilingualism: Language and Cognition**, v. 22, n. 4, p. 657-679, 2019.

DIJKSTRA, T.; GRAINGER, J.; VAN HEUVEN, W. J. B. Recognition of cognates and interlingual homographs: The neglected role of phonology. **Journal of memory and language**, v. 41, n. 4, p. 496–518, 1999.

DIJKSTRA, T.; VAN HEUVEN, W.J.B. The architecture of the bilingual word recognition system: From identification to decision. **Bilingualism: Language and cognition**, v. 5, n. 3, p. 175-197, 2002.

- DIJKSTRA, T.; VAN HEUVEN, W. J. B.; GRAINGER, J. Simulating cross-language competition with the bilingual interactive activation model. **Psychologica belgica**, v. 38, n. 3–4, p. 177, 1998.
- DIJKSTRA, T.; VAN JAARSVELD, H.; BRINKE, S. T. Interlingual homograph recognition: Effects of task demands and language intermixing. **Bilingualism** (Cambridge, England), v. 1, n. 1, p. 51–66, 1998.
- DUMPER, M.; STANLEY, B. E. **Cities of the Middle East and North Africa: a historical encyclopedia**. Santa Barbara, Calif.: Abc-Clio, 2007.
- DUÑABEITIA, J. A.; PEREA, M.; CARREIRAS, M. Masked translation priming effects with highly proficient simultaneous bilinguals. **Experimental psychology**, v. 57, n. 2, p. 98–107, 2010.
- ELSTON-GÜTTLER, K. E.; GUNTER, T. C.; KOTZ, S. A. Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. **Brain research. Cognitive brain research**, v. 25, n. 1, p. 57–70, 2005.
- ESTIVALET, G. L.; MEUNIER, F. Corpus psicolinguístico Léxico do Português Brasileiro. **Revista SOLETRAS**, v. 0, n. 33, 2017.
- ESTIVALET, Gustavo Lopez. **Léxico do Português Brasileiro – LexPorBR**. 2019. Disponível em: <http://www.lexicodoportugues.com/>. Acesso em: 26 abr. 2023.
- EVEN-ZOHAR, I. POLYSYSTEM STUDIES. **POETICS TODAY: International Journal for Theory and Analysis of Literature and Communication**. Vol. 11, num. 1 (1990).
- FARID, M.; GRAINGER, J. How initial fixation position influences visual word recognition: a comparison of French and Arabic. **Brain and language**, v. 53, n. 3, p. 351–368, 1996.
- FISCHER, L.; LÄUBLI, S. What's the Difference Between Professional Human and Machine Translation? A Blind Multi-language Study on Domain-specific MT. **arXiv preprint arXiv:2006.04781**, 2020.
- FORSTER, K. I.; CHAMBERS, S. M. Lexical access and naming time. **Journal of verbal learning and verbal behavior**, v. 12, n. 6, p. 627–635, 1973.
- FRAZIER, S.; BROWN, H. D. Teaching by principles: An interactive approach to language pedagogy. **TESOL quarterly**, v. 35, n. 2, p. 341, 2001.
- GESSER, A. “Um olho no professor surdo e outro na caneta”: **Ouvintes aprendendo a Língua Brasileira de Sinais**. Tese de doutorado inédita, Campinas: Unicamp, 2006
- GOLLAN, T. H. et al. The bilingual effect on Boston Naming Test performance. **Journal of the International Neuropsychological Society: JINS**, v. 13, n. 2, p. 197–208, 2007.
- GÖPFERICH, S.; ALVES, F.; MEES, I.M. (Ed.). **New approaches in translation process research**. Samfundslitteratur, 2010.

GÖPFERICH, S.; JAKOBSEN, A.L.; MEES, I.M. (Ed.). **Behind the mind: Methods, models and results in translation process research**. Samfundslitteratur, 2009.

GRAINGER, J. Visual word recognition in bilinguals. Em: **Studies in Bilingualism**. Amsterdam: John Benjamins Publishing Company, 1993. p. 11.

GUERINI, A.; COSTA, W. **Introdução aos Estudos de Tradução**. Universidade Federal de Santa Catarina. Florianópolis, 2006. ISBN: 978-85-60522-00-2.

GUTTENTAG, R. E. et al. Semantic processing of unattended words by bilinguals: A test of the input switch mechanism. **Journal of verbal learning and verbal behavior**, v. 23, n. 2, p. 178–188, 1984.

HAMERS, J.F.; BLANC, M. **Bilinguality and bilingualism**. Cambridge University Press, 2000.

HASIBUAN, Z. A comparative study between human translation and machine translation as an interdisciplinary research. **Journal of English Teaching and Learning Issues**, v. 3, n. 2, p. 115-130, 2020.

HERMANS, D. et al. Producing words in a foreign language: Can speakers prevent interference from their first language? **Bilingualism (Cambridge, England)**, v. 1, n. 3, p. 213–229, 1998.

HOLMES, J. The Name and Nature of Translation Studies. Em: Holmes, J.S., 1988. **Translated!** Papers on Literary Translation and Translation Studies. Amsterdam: Rodopi, p. 67-80. Reprinted e.g. in VENUTI, L. (ed.). *The Translation Studies Reader*. London: Routledge, 2000, p. 172-185.

HUTCHINS, J. **The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954**. 2004.

JACOBY, L. L. Perceptual enhancement: persistent effects of an experience. **Journal of experimental psychology. Learning, memory, and cognition**, v. 9, n. 1, p. 21–38, 1983.

JACOBY, L. L.; DALLAS, M. On the relationship between autobiographical memory and perceptual learning. **Journal of experimental psychology. General**, v. 110, n. 3, p. 306–340, 1981.

KECSKES, I.; ALBERTAZZI, L. **Cognitive Aspects of Bilingualism**. Nova Iorque, NY, USA: Springer, 2007.

KERKHOFS, R. et al. Testing a model for bilingual semantic priming with interlingual homographs: RT and N400 effects. **Brain research**, v. 1068, n. 1, p. 170–183, 2006.

KRASHEN, S. D.; TERRELL, T. D. **The Natural Approach. Language Acquisition in the Classroom**. 1983.

KRATZ, D. An interview with Norman Shapiro. **Translation review**, v. 19, n. 1, p. 27–28, 1986.

KROLL, J.F.; STEWART, E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. **Journal of memory and language**, v. 33, n. 2, p. 149-174, 1994.

LAMEIRA, M. et al. Evidence of non-selective lexical access to second and third language in unbalanced multilinguals: an N400 study on the processing of interlingual homographs. **Pandaemonium Germanicum**, v. 26, n. 49, p. 35–67, 2023.

LEVELT, Willem JM. **A history of psycholinguistics: The pre-Chomskyan era**. Oxford University Press, 2013.

LEVY, J.; ALTHOFF, G.; VIDAL, C. Translation as a Decision Process / A Tradução como um Processo de Tomada de Decisão. **Scientia traductionis**, v. 0, n. 11, 2012.

LIGHTBOWN, P. M.; SPADA, N. **How Languages Are Learned**. Oxford University Press. 1998.

MACIZO, P.; BAJO, T.; CRUZ MARTÍN, M. Inhibitory processes in bilingual language comprehension: Evidence from Spanish–English interlexical homographs. **Journal of memory and language**, v. 63, n. 2, p. 232–244, 2010.

MONSELL, S.; DOYLE, M. C.; HAGGARD, P. N. Effects of frequency on visual word recognition tasks: where are they? **Journal of experimental psychology. General**, v. 118, n. 1, p. 43–71, 1989.

MUFTAH, M. Machine vs human translation: a new reality or a threat to professional Arabic–English translators. **PSU Research Review**, n. ahead-of-print, 2022.

OSIMO, B. **LOGOS - Multilingual Translation Portal**. Disponível em: http://courses.logos.it/IT/1_19.html. Acesso em: 1 dez. 2022.

PATKOWSKI, M. S. Age and accent in a second language: A reply to James Emil Flege. **Applied linguistics**, v. 11, n. 1, p. 73–89, 1990.

POORT, E. D.; WARREN, J. E.; RODD, J. M. Recent experience with cognates and interlingual homographs in one language affects subsequent processing in another language. **Bilingualism** (Cambridge, England), v. 19, n. 1, p. 206–212, 2016.

POTTER, M. C. et al. Lexical and conceptual representation in beginning and proficient bilinguals. **Journal of verbal learning and verbal behavior**, v. 23, n. 1, p. 23-38, 1984.

PRECUP-STIEGELBAUER, L.-R. Automatic translations versus human translations in nowadays world. **Procedia, social and behavioral sciences**, v. 70, p. 1768–1777, 2013.

PRESTON, K. A. The speed of word perception and its relation to reading ability. **The Journal of general psychology**, v. 13, n. 1, p. 199–203, 1935.

ROEDIGER, H. L.; BLAXTON, T. A. Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. **Memory & cognition**, v. 15, n. 5, p. 379–388, 1987.

RONAI, P. **Como aprendi o português e outras aventuras**. Porto Alegre: Globo, 1956.

RONAI, P. **Escola De Tradutores**. 6. Ed. Rio de Janeiro, 1987.

SCARBOROUGH, D. L.; CORTESE, C.; SCARBOROUGH, H. S. Frequency and repetition effects in lexical memory. **Journal of experimental psychology. Human perception and performance**, v. 3, n. 1, p. 1–17, 1977.

SCHACTER, D. L. Perceptual representation system and implicit memory: Toward a resolution of the multiple memory debate. Em: DIAMOND, A. (Ed.). **Annals of the New York Academy of Sciences**. [s.l.] New York Academy of Sciences, 1990. v. 608p. 543–571.

SCHULPEN, B. et al. Recognition of interlingual homophones in bilingual auditory word recognition. **Journal of experimental psychology. Human perception and performance**, v. 29, n. 6, p. 1155–1178, 2003.

SHAHROKHI, M. The Impact of Think-aloud on the Speed of Translation. **International Journal of English Language & Translation Studies**, v. 4, n. 1, p. 60-68, 2016.

SHREVE, G.M.; ANGELONE, E. **Translation and cognition**. Amsterdam/Philadelphia, v. 10, 2010.

SMITS, E. et al. Naming interlingual homographs: Variable competition and the role of the decision system. **Bilingualism** (Cambridge, England), v. 9, n. 3, p. 281–297, 2006.

TAKAKUSAGI, Y. et al. Validation of the reliability of machine translation for a medical article from Japanese to English using deepl translator. **Cureus**, v. 13, n. 9, 2021.

TOASSI, P.F.P. et al. Effect of interlingual homographs and word frequency on bilingual lexical access. **Ilha do Desterro**, v. 76, n. 3, p. 67–91, 2023.

TOASSI, P.F.P; MOTA, M.B. Acesso lexical de bilíngues e multilíngues. **Acta Scientiarum: Language and Culture**, Maringá, v. 37, p. 393-404, 2015. DOI: <https://doi.org/10.4025/actascilangcult.v37i4.25234>.

TULVING, E.; SCHACTER, D. L. Priming and human memory systems. **Science (New York, N.Y.)**, v. 247, n. 4940, p. 301–306, 1990.

TURNER, A.M. et al. A comparison of human and machine translation of health promotion materials for public health practice: time, costs, and quality. **Journal of public health management and practice: JPHMP**, v. 20, n. 5, p. 523, 2014.

VAN ASSCHE, E. et al. Does bilingualism change native language reading? Cognate effects in a sentence context. **Psychological science**, v. 20, n. 8, p. 923–927, 2009.

VAN HEUVEN, W. J. B.; DIJKSTRA, T.; GRAINGER, J. Orthographic neighborhood effects in bilingual word recognition. **Journal of memory and language**, v. 39, n. 3, p. 458–483, 1998.

WYLER, Lia. Entrevista com Lia Wyler. **Cadernos de Tradução**, Florianópolis, v. 2, n. 8, p. 205-231, jan. 2001. Disponível em: <https://periodicos.ufsc.br/index.php/traducao/article/view/5894>. Acesso em: 2 dez. 2022.