



Texto Livre: Linguagem e Tecnologia
E-ISSN: 1983-3652
revista@textolivre.org
Universidade Federal de Minas Gerais
Brasil

de Lima Fonseca, Norma Barbosa
Investigando adjetivos atributivos em traduções publicadas, em um sistema de tradução automática e na pós-edição monolíngue: contribuições para a pedagogia da tradução
Texto Livre: Linguagem e Tecnologia, vol. 11, núm. 1, january-april, 2018, pp. 121-143
Universidade Federal de Minas Gerais

Disponível em: <https://www.redalyc.org/articulo.oa?id=577163617008>

- Como citar este artigo
- Número completo
- Mais artigos
- Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica
Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

INVESTIGANDO ADJETIVOS ATRIBUTIVOS EM TRADUÇÕES PUBLICADAS, EM UM SISTEMA DE TRADUÇÃO AUTOMÁTICA E NA PÓS-EDIÇÃO MONOLÍNGUE: CONTRIBUIÇÕES PARA A PEDAGOGIA DA TRADUÇÃO

INVESTIGATING ATTRIBUTIVE ADJECTIVES IN PUBLISHED TRANSLATIONS, IN A MACHINE TRANSLATION SYSTEM AND IN MONOLINGUAL POST-EDITING: CONTRIBUTIONS TO TRANSLATION PEDAGOGY

Norma Barbosa de Lima Fonseca

Universidade Federal de Ouro Preto

normafonseca@gmail.com

RESUMO: Este artigo visa investigar a tradução em língua portuguesa de adjetivos atributivos em um trecho da obra em língua inglesa *Heart of Darkness* (CONRAD, 1994) em três estudos. O primeiro analisa as opções de tradução desse trecho em quatro traduções publicadas nos anos 1984, 1996, 2008 e 2011. O segundo examina o insumo da tradução automática desse mesmo trecho fornecido pelo Google Translate. O terceiro analisa textos-alvo produzidos em uma tarefa de pós-edição monolíngue, cujo texto-fonte foi esse mesmo trecho, por oito estudantes de uma disciplina de prática de tradução. Dessa maneira, integra-se uma análise do processo e do produto tradutórios para verificar padrões na tradução de grupos nominais com um, dois e três adjetivos atributivos presentes no trecho analisado. Para a realização desses estudos, utilizou-se parte do RETRAD (*Corpus* de (Re)Traduções), o qual foi complementado com o trecho traduzido automaticamente pelo Google Translate e com os textos pós-editados pelos estudantes. Os resultados apontam alguns padrões sintáticos e lexicais recorrentes nos grupos nominais com até dois adjetivos atributivos nas traduções publicadas, havendo, no caso de grupos nominais com mais de dois adjetivos, a inclusão de preposições e conjunções em português. Observou-se também uma maior variedade sintática e lexical nos grupos nominais com três ou mais adjetivos atributivos nas traduções publicadas. O fato que os estudantes fizeram poucas alterações nos textos-alvo pós-editados parece indicar que esses participantes não se sentem confiantes para produzir o texto-alvo sem consulta ao texto-fonte. Desse modo, é necessário incentivar o uso de corpora em sala de aula, expor aos estudantes as limitações de sistemas de tradução automática e as estratégias de solução de problemas, levando-os a refletir sobre os processos de tradução e de pós-edição. Isso contribuirá para o aumento da confiança desses estudantes em suas escolhas e para a produção de textos-alvo que representam a língua em uso.

PALAVRAS-CHAVE: adjetivos atributivos; sistema de tradução automática; pós-edição monolíngue; linguística de *corpus*; pesquisa experimental.

ABSTRACT: This article aims at investigating the translation in Brazilian Portuguese of attributive adjectives in an excerpt from *Heart of Darkness* (CONRAD, 1994) in three studies. The first analyses the translation of the same excerpt in four translations published in 1984, 1996, 2008 and 2011. The second examines the machine-translated output from English into Portuguese of the same excerpt as provided by Google Translate. The third analyzes the target texts produced in a monolingual post-editing task performed

by students of a translation practice subject. Thus, it integrates a process and product analysis to verify patterns of translating noun phrases with one, two and three attributive adjectives in the analyzed excerpt. For this study, part of *Corpus de (Re)traduções (RETRAD)* was used, which was complemented by the machine-translated output and students' post-edited texts. Results point to recurrent syntactic and lexical patterns in the translation of noun phrases with up to two attributive adjectives in published translations, with the addition of prepositions and conjunctions in noun phrases with more than two adjectives. Greater syntactic and lexical variety is also found in the translations of noun phrases with three or more attributive adjectives in published translations. The fact that students made a few changes in the post-edited target texts seems to indicate that these participants do not feel confident to produce the target text without referring to the source text. Therefore, it is necessary to promote the use of corpora in translation classroom, to expose students to the limitations of machine translations systems and to problem-solving strategies, leading them to reflect on the translation and post-editing processes. This will contribute to increase students' confidence in their choices and to produce target texts that represent the language in use.

KEYWORDS: attributive adjectives; machine translation system; monolingual post-editing, corpus linguistics; experimental research.

1 Introdução

Desde o surgimento da Linguística de *Corpus* na década de 80, ela tem sido utilizada em estudos de diferentes áreas de conhecimento, como na Sociolinguística (KENDALL, 2011), na Linguística Histórica (KYTO, 2011) e no Ensino-aprendizagem de Segunda Língua/Língua Estrangeira (MEUNIER, 2011), dentre outras. Esse grande interesse gerado por pesquisadores de diferentes áreas deve-se ao fato de que “[o]s dados de *corpus* são essenciais para descrever com precisão o uso da língua e mostram como o léxico, a gramática e a semântica interagem” (STUBBS, 2004, p. 106, tradução da autora).

Nos Estudos da Tradução, os corpora tradicionalmente vêm sendo utilizados em pesquisas que diferenciam, por exemplo, características próprias do texto traduzido do texto não traduzido (OLOHAN; BAKER, 2000), o estilo de traduções e/ou do tradutor (ASSIS, 2009; BARCELLOS, 2011, 2016; MAGALHÃES; BARCELLOS, 2014; MAGALHÃES; BLAUTH, 2015; MONTENEGRO, 2015; CASTRO, 2016; BLAUTH, 2014, 2015). Essas pesquisas sobre o estilo mencionadas foram desenvolvidas a partir do *Corpus de (Re)Traduções (RETRAD)*¹, construído por pesquisadores do Laboratório Experimental de Tradução (LETRA) da Faculdade de Letras da Universidade Federal de Minas Gerais. Nesses estudos, há enfoque no produto tradutório, uma das subáreas do primeiro mapeamento dos Estudos da Tradução realizado por Holmes (1972/2000). Apesar de esse mapeamento já postular uma inserção dos estudos orientados ao processo tradutório nos estudos descritivos, foi apenas a partir da segunda metade dos anos 80, com as publicações de autores como Krings (1986) e Königs (1987), dentre outros, que os estudos processuais ganharam notoriedade. Krings (2001) amplia o objeto de estudo do processo tradutório ao fazer uma investigação processual da pós-edição, ou

1 *Corpus de Estilo em Tradução (ESTRA)* era o nome anterior do *Corpus RETRAD*.

seja, do processo de correção de um texto traduzido automaticamente (KOBY, 2001).

A partir dessa ampliação, o presente trabalho busca preencher uma lacuna dos estudos que integram uma abordagem tanto do produto tradutório quanto do processo de pós-edição. Para realizar essa integração, foram comparadas traduções publicadas em português do Brasil de um trecho da obra *Heart of Darkness* (CONRAD, 1994), escrita em inglês com a tradução automática desse mesmo trecho em inglês para o português, fornecida pelo sistema de tradução automática Google Translate², e com os textos-alvo produzidos em uma tarefa de pós-edição monolíngue a partir dessa tradução automática. Essa tarefa de pós-edição foi realizada por oito estudantes do curso de Letras da Universidade Federal de Minas Gerais que cursavam uma disciplina de prática de tradução. Diferente da pós-edição bilíngue, a qual é realizada com a consulta ao texto-fonte, a pós-edição monolíngue é executada sem consulta ao texto-fonte (TF) e/ou o pós-editor desconhece a língua-fonte (KOEHN, 2010).

O trecho da obra selecionado para o desenvolvimento desta pesquisa apresenta grupos nominais com diferentes quantidades de adjetivos atributivos, buscando-se responder à seguinte pergunta de pesquisa: Como a tradução automática de grupos nominais com adjetivos atributivos se diferencia e se assemelha das opções de tradução em textos-alvo produzidos em uma tarefa de pós-edição monolíngue e em traduções publicadas?

A escolha desse tipo de estrutura gramatical deveu-se à dificuldade geralmente associada à tradução dessa estrutura do inglês para o português, uma vez que os grupos nominais do inglês possuem uma ordem diferente do português, permitindo uma maior quantidade de adjetivos atributivos em sequência – às vezes sem nenhuma palavra adicional que não seja adjetivo ou apenas com a conjunção *and*. Acredita-se que essa mesma dificuldade possa ser associada a uma tradução automática do inglês para o português e a tarefas de pós-edição monolíngue, em que o falante de língua portuguesa tem apenas um texto em sua própria língua para processar esses grupos nominais durante a execução da tarefa.

Desse modo, esta pesquisa objetiva verificar padrões de traduções de grupos nominais com um, dois e três adjetivos atributivos em três estudos distintos. O primeiro estudo analisa as traduções desses grupos nominais em um trecho da obra *Heart of Darkness*, em quatro traduções brasileiras: uma realizada por Hamilton Trevisan, em 1984, publicada pela editora Global; uma por Marcos Santarrita, em 1996, publicada pela editora Ediouro; uma por José Roberto O'Shea, em 2008, publicada pela editora Hedra; e uma por Fábio Cyrino, em 2011, publicada pela editora Landmark. O segundo estudo examina as traduções dos grupos nominais do mesmo trecho no insumo da tradução automática fornecido pelo Google Translate. Finalmente, o terceiro estudo analisa as traduções desses grupos nominais em textos-alvo produzidos em uma tarefa de pós-edição monolíngue executada por estudantes de uma disciplina de prática de tradução.

Esta pesquisa pretende ainda contribuir para a didática da tradução e da pós-edição em cursos de formação de tradutores, incentivando o uso de tradução automática e de *corpus* para a realização de tarefas de tradução e de pós-edição e a reflexão sobre o processo de tradução e de pós-edição em sala de aula. Esse incentivo é necessário para a tradução de qualquer gênero textual, mas especialmente de textos literários, que são

2 Disponível em <<https://translate.google.com/m/translate?hl=pt-BR>>. Acesso em 2 abr. 2018.

frequentemente traduzidos sem uso da tradução automática e têm os corpora como ferramentas de apoio fundamentais para tradução, contribuindo para “o entendimento da língua e do processo tradutório” (BERBER SARDINHA, 2003, p. 48).

Para o desenvolvimento desta investigação, este artigo apresenta, além desta Introdução e do Referencial teórico, três estudos a serem relatados na seção Metodologia. O Estudo 1 concentra-se na análise das traduções de grupos nominais com adjetivos atributivos em um trecho das traduções publicadas em língua portuguesa de *Heart of Darkness*: Trevisan (1984), Santarrita (1996), O'Shea (2008) e Cyrino (2011). O Estudo 2 verifica a tradução desses grupos nominais na tradução automática fornecida pelo Google Translate, e o Estudo 3 investiga como estudantes de uma disciplina de prática de tradução processam esses grupos nominais em uma tarefa de pós-edição monolíngue sem acesso ao TF.

2 Estudos da Tradução e Linguística de *Corpus*

Este estudo integra conceitos e teorias da Linguística de *Corpus* e dos Estudos da Tradução para a sua realização, conforme destacado nas subseções a seguir.

2.1 O conceito de tradução

Uma das definições mais conhecidas da palavra tradução é a de Jakobson (1959/2004). O autor apresenta o conceito de tradução classificando-a em três tipos:

1. A tradução intralingual ou *reformulação* é a interpretação dos signos verbais por meio de outros signos da mesma língua.
2. A tradução interlingual ou *tradução propriamente dita* é a interpretação dos signos verbais por meio de alguma outra língua.
3. A tradução intersemiótica ou *transmutação* é a interpretação dos signos verbais por meio de sistemas de signos não verbais. (JAKOBSON, 1959/2004, p. 139, ênfase no original, tradução da autora).

O conceito de tradução usado neste artigo refere-se ao segundo tipo destacado por Jakobson (1959/2004), qual seja a tradução de uma língua-fonte (LF) para outra língua, a língua-alvo (LA). Entretanto, “mais do que um texto escrito na página, o produto do processo tradutório” (MUNDAY, 2009, p. 6), a tradução pode ser compreendida como o que Hatim e Munday (2004) chamam de “âmbito da tradução”, que abarca uma visão da tradução tanto como processo quanto como produto.

1. O processo de transferência de um texto escrito a partir de uma LF para uma LA, conduzido por um tradutor ou tradutores, em um contexto sociocultural específico.
2. O produto escrito, ou TA, que resulta desse processo e que funciona no contexto sociocultural da LA.
3. Os fenômenos cognitivo, linguístico, visual, cultural e ideológico que são parte integrante de 1 e 2. (HATIM; MUNDAY, 2004, p. 6, tradução da autora)

Conforme ressalta Munday (2009), apesar de ser uma área de investigação relativamente recente, o estudo da tradução baseou-se inicialmente na prática da tradução, com escritos de tradutores clássicos, como Cícero e São Jerônimo. Percebe-se já nesses escritos iniciais uma preocupação não só com o produto tradutório, mas também com o processo tradutório, em que os autores explicavam e discutiam suas estratégias de tradução e suas escolhas, que poderiam optar por uma tradução palavra por palavra, configurando o que se chama de tradução literal, ou pela tradução pelo sentido, em que o tradutor busca traduzir o texto-fonte pelo sentido, com uma tradução mais livre (SÃO JERÔNIMO, 1995).

Com o trabalho de Vinay e Darbenelt (1958), as discussões teóricas, inseridas em uma abordagem linguística da tradução, tiveram início e se concentraram nos sete procedimentos da tradução: empréstimo, decalque, tradução literal, transposição, modulação, equivalência e adaptação, dando impulso para o surgimento dos Estudos da Tradução. Passados 60 anos do início dessas discussões teóricas, a Linguística de *Corpus* têm permitido inovações principalmente metodológicas nos Estudos da Tradução.

2.2 A Linguística de *Corpus* nos Estudos da tradução

No início dos anos 60, a análise de *corpus*, uma palavra inicialmente utilizada em referência a “uma coletânea ou reunião de textos escritos de natureza semelhante” (McCARTHY; O’KEEFFE, 2010, p. 4), passava por grandes transformações com o advento da tecnologia, tornando possível o armazenamento de grandes quantidades de textos e o tratamento automático deste material. Surge então o primeiro *corpus* eletrônico, o *Brown Corpus*, criado no início dos anos 60, por W. Nelson Francis e Henry Kefera, que “são agora considerados pioneiros e visionários na comunidade da Linguística de *Corpus*” (MEYER, 2004, p. 1, tradução da autora).

A partir dos anos 80, os Estudos da Tradução consolidam-se como uma disciplina acadêmica, e diferentes abordagens passam a discutir a tradução: abordagens textuais, abordagens comunicativas/socioculturais, abordagens filosóficas/hermenêuticas, abordagens cognitivas e abordagens linguísticas (HURTADO ALBIR, 2001). Esse grande avanço nos Estudos da Tradução coincide com a difusão de *corpora* eletrônicos, ocorrida entre 1980 e 1990, que passaram a ter prestígio junto a várias comunidades acadêmicas, dentre elas a linguística, fazendo surgir a disciplina Linguística de *Corpus* (KENNY, 2001), a qual estuda a língua em uso. A palavra *corpus* refere-se a um conjunto de textos falados e/ou escritos de tamanho finito, produzidos espontaneamente e formatados eletronicamente para serem analisados por ferramentas computacionais (McENERY; WILSON, 2001).

A partir da sugestão de Lindquist (1984) de uso de *corpora* para o treinamento de tradutores, Baker (1993) propõe uma interface entre Estudos da Tradução e Linguística de *Corpus*. A autora argumenta que “a disponibilidade de grandes *corpora* tanto de textos originais quanto de textos traduzidos, juntamente com o desenvolvimento de uma metodologia baseada em *corpus*, permitirá aos acadêmicos de tradução revelar a natureza do texto traduzido enquanto um evento comunicativo mediado” (BAKER, 1993, p. 243, tradução da autora).

Baker (1995) define *corpus* principalmente como “uma coleção de textos compilados em formato eletrônico e capazes de serem analisados semiautomaticamente

ou automaticamente em uma variedade de formas” (BAKER, 1995, p. 225, tradução da autora). A autora apresenta uma classificação para os *corpora*, dividindo-os em três categorias: i) corpora paralelos, que contêm textos-fonte em uma língua A e suas traduções para uma língua B; ii) corpora multilíngues, que contêm grupo de corpora monolíngues, que são conjuntos de textos não traduzidos em determinada língua; iii) corpora comparáveis, que contêm dois subcorpora: um *corpus* monolíngue, contendo textos originalmente produzidos em uma língua X e um *corpus* de textos traduzidos para a língua X. Essa interface entre estudos descritivos da tradução e Linguística de *Corpus* foi posteriormente chamada de Estudos da Tradução Baseados em Corpus (ETBC), proporcionando novas possibilidades de investigação nos Estudos da Tradução.

2.3 Abordagem processual da tradução e da pós-edição nos Estudos da Tradução

Até a década de 80, o campo disciplinar Estudos da Tradução havia se dedicado quase exclusivamente ao estudo do produto tradutório, com enfoque em aspectos linguísticos, literários etc. Entretanto, em 1986, o campo disciplinar passa por uma virada processual, com a publicação da tese de doutorado de Krings (1986), em que o autor apresenta um mapeamento do processo tradutório. Esse mapeamento baseia-se na análise de dados de protocolos verbais concomitantes (TAPs, *Think Aloud Protocols*) e de estratégias utilizadas por oito estudantes de francês durante a realização de tarefas de tradução direta e inversa.

Krings (2001) investiga os processos cognitivos envolvidos em processos de tradução e de pós-edição de textos traduzidos automaticamente em sua pesquisa de pós-doutorado. Esse estudo envolvia os pares linguísticos inglês-alemão, francês-alemão e alemão-inglês, sendo participantes 52 estudantes de língua alemã de um programa de tradução técnica. Os métodos utilizados foram os protocolos verbais retrospectivos e a gravação de vídeos. O objetivo de Krings (2001), ao realizar esse estudo, foi fornecer uma avaliação objetiva e empírica da tradução automática e da pós-edição, considerando o ponto de vista psicolinguístico e o custo e esforço despendido em tarefas de pós-edição.

Uma das tarefas que os participantes de Krings (2001) executaram foi a pós-edição monolíngue, entretanto o autor não usou esse nome para referir-se a esse tipo de pós-edição em que o pós-editor realiza a tarefa sem acesso ao TF e/ou desconhece a língua-fonte. Além disso, o produto da tradução automática do inglês-alemão e o produto da pós-edição monolíngue foram submetidos à avaliação por instrutores de tradução e por tradutores em formação. A média das notas atribuídas por esses profissionais ao texto traduzido automaticamente foi 2,39, e o texto pós-editado recebeu a nota média de 3,38, em um total de 5,0. Mesmo realizando a pós-edição sem a consulta ao TF, quase 80% dos erros foram corrigidos com sucesso pelos estudantes, mas Krings (2001) concluiu que os erros não corrigidos eram tipicamente aqueles que levariam o leitor a interpretar erroneamente o sentido do TF. De acordo com Krings (2001), a ausência do TF pode ter levado os participantes, que executaram a pós-edição monolíngue, a apresentarem frases no TA final com sentidos bem diferentes daquelas que estavam no TF. Esses erros podem não ocorrer na pós-edição bilíngue, uma vez que, nesse tipo de pós-edição, o participante tem acesso ao TF e pode recorrer a ele sempre que precisar no decorrer do processo de pós-edição.

Diversos programas são utilizados atualmente para traduzir automaticamente os textos para serem posteriormente pós-editados, com ou sem consulta ao TF. Dentre esses programas, o Google Translate, um sistema de tradução automática cuja base de processamento iniciou-se como predominantemente estatística em 2006 e que recentemente foi alterada para um mecanismo de tradução automática por redes neurais (GNMT, [Google Neural Machine Translation](#)), que traduz frases inteiras de uma vez, em vez de palavras isoladas.

Ressalta-se que o trabalho de Krings (2001) retoma um tema que havia sido quase totalmente abandonado na década de 60 pela maioria dos pesquisadores, a saber, a tradução automática, que havia sido objeto de estudo desde a década de 40. Nessa época, de acordo com Santos (1995), Booth e Weaver desenvolveram uma aplicação inicial, impulsionados pela busca de informações da União Soviética, durante a Guerra Fria o mais rápido possível. Essa aplicação desenvolvida pelos autores consistia em uma calculadora científica tradutora que realizava uma tradução palavra por palavra, proporcionando apenas uma ideia do conteúdo do TF. Em 1948, para aprimorar essa aplicação, Richens introduziu informações relativas à análise gramatical das desinências russas, o que proporcionou uma maior compreensão do TF, e posteriormente testou a validade do método em diferentes línguas (RICHENS; BOOTH, 1955). Em 1950, Reifler (*apud* Santos, 1995) apontou a necessidade da preparação dos textos a serem submetidos à tradução automática, o que, nos termos atuais, poderia ser chamado de pré-edição.

Conforme Santos (1995), no início dos anos 50, Weaver tentou solucionar problemas de ambiguidade semântica da aplicação. Nessa mesma época, surgem estudos e conferências sobre a tradução automática, como o congresso do Instituto de Tecnologia de Massachusetts em 1952, que abordou os problemas da tradução automática, e no qual se decidiu investigar inicialmente a frequência das palavras, as equivalências linguísticas, as memórias eletrônicas e outros aspectos técnicos, para então se passar para análise sintática e a construção de programas de tradução automática. Além disso, nesse evento, determinou-se como objetivo de curto prazo desenvolver um programa que realizasse a tradução em uma única direção entre duas línguas, sem uma preocupação inicial com a produção de traduções multilíngues.

Ainda de acordo com Santos (1995), em 1954, na Universidade de Georgetown, ocorreu a primeira experiência de tradução automática do russo para o inglês com um computador IBM 701, e a mesma foi bem-sucedida. Posteriormente, foram realizados diversos congressos (americanos e ingleses) e houve diversas publicações de 1955 a 1964 sobre tradução automática. Entretanto, em 1966, a publicação do relatório ALPAC com uma avaliação negativa desses estudos de tradução automática culminou com a redução de financiamentos de projetos nessa área. Consequentemente, diversos pesquisadores viram-se obrigados a abandonar suas pesquisas, sendo poucos os estudiosos, como Peter Toma, desenvolvedor do programa Systran, que continuaram suas investigações. Apenas na década de 80 é que ressurgiu o interesse pela tradução automática, destacando-se a aprovação em 1982 do EUROTRA, o maior projeto europeu de tradução automática.

Desde então, ocorreram muitas evoluções em sistemas de tradução automática com a criação de diversas ferramentas estejam elas disponíveis na Web ou como aplicativos para instalação em computadores pessoais. Atualmente, com o uso dessas

ferramentas, é possível ter uma ideia do TF, mas, se o usuário requer alguma qualidade dessa tradução automática, será necessário pós-editar esse texto. A próxima seção deste artigo aborda, dentre outros aspectos metodológicos, o uso de uma dessas ferramentas, o Google Translate, para traduzir automaticamente um trecho da obra *Heart of Darkness*.

3 Metodologia

3.1 Descrição do *corpus* e dos estudos

Partindo de uma perspectiva de múltiplas abordagens de um mesmo tema, a metodologia apresentada a seguir foi empregada em três estudos distintos. O Estudo 1 consistiu na observação de padrões de tradução de grupos nominais com adjetivos atributivos em um parágrafo nas traduções publicadas do livro *Heart of Darkness* e na busca e comparação das traduções dos substantivos desses grupos nominais com as traduções presentes no *corpus* Compara³, que é um *corpus* paralelo bidirecional de português e inglês criado pela Linguateca, um projeto em rede formalmente chamado de Centro de Recursos – distribuído – para a Língua Portuguesa, do qual fazem parte pesquisadores de diferentes instituições brasileiras e estrangeiras. O Estudo 2 abrangeu a análise das traduções de grupos nominais com adjetivos atributivos do trecho indicado após ele ser submetido à tradução no sistema de tradução automática Google Translate e na verificação de colocados com as traduções automáticas fornecidas para os substantivos no *Corpus* do Português⁴, um *corpus* do Português do Brasil criado por Mark Davies e Michael Ferreira em 2006. O Estudo 3 investigou o processamento de grupos nominais com adjetivos atributivos do mesmo trecho de *Heart of Darkness* em textos-alvo produzidos em uma tarefa de pós-edição monolíngue por estudantes de Letras (Bacharelado em Tradução) em uma aula de prática de tradução.

Conforme a classificação de Baker (1995), o *corpus* utilizado nesta pesquisa é um *corpus* escrito paralelo bilíngue, contendo o TF, a obra de ficção *Heart of Darkness* (CONRAD, 1994) em inglês, que continha 38.803 palavras, e quatro traduções publicadas para o português do Brasil em 1984, 1996, 2008 e 2011, conforme a descrição do Estudo 1 a seguir. Além desses textos, o *corpus* foi complementado com a tradução automática do trecho selecionado para análise, conforme o Estudo 2, e com os textos-alvo pós-editados pelos participantes desta pesquisa, conforme detalhes fornecidos no Estudo 3.

3.1.1 Estudo 1: traduções publicadas

Os principais dados de quatro traduções publicadas da obra *Heart of Darkness* analisadas neste estudo são indicados no Quadro 1 e foram contabilizados no WordSmith Tools⁵, uma ferramenta de anotação criada por Scott (2012) para analisar aspectos da língua em uso. Esse quadro apresenta o nome do tradutor, o ano de publicação, a editora que publicou a tradução e o número de palavras de cada uma das traduções publicadas, com um total de 150.406 palavras ($M=37.601$) e 38.803 palavras no TF.

3 Disponível em: <<https://www.linguateca.pt/COMPARA>>. Acesso em: 13 mai. 2018.

4 Disponível em: <<http://www.corpusdoportugues.org>>. Acesso em: 9 mai. 2018.

5 Disponível em: <<http://lexically.net/wordsmith/research>>. Acesso em: 10 mai. 2018.

Quadro 1: Dados de publicação de quatro traduções brasileiras da obra *Heart of Darkness*.

Tradutor	Ano	Editores	Palavras
Hamilton Trevisan	1984	Global	37.085
Marcos Santarrita	1996	Ediouro	36.639
José Roberto O'Shea	2008	Hedra	36.004
Fábio Cyrino	2011	Landmark	40.678

Fonte: Elaborado pela autora utilizando o WordSmith Tools.

3.1.2 Estudo 2: tradução automática

Um parágrafo do livro *Heart of Darkness* com 189 palavras, conforme o Anexo 1, foi selecionado para análise com base na presença de grupos nominais com diferentes quantidades de adjetivos atributivos. Esse parágrafo selecionado foi submetido à tradução automática no Google Translate em 12/09/2012. O insumo traduzido com 201 palavras é apresentado no Anexo 2 e também passou a fazer parte do *corpus* paralelo bilíngue, que foi criado inicialmente por pesquisadores do LETRA, constando o TF de *Heart of Darkness* e diversas traduções publicadas, além das quatro traduções analisadas no Estudo 1 desta pesquisa.

3.1.3 Estudo 3: pós-edição monolíngue

Outro subcorpus desta pesquisa consistiu nos textos-alvo pós-editados por 8 estudantes de Letras (Bacharelado em Tradução) que cursavam uma disciplina de prática de tradução na Universidade Federal de Minas Gerais no segundo semestre de 2012. Para a realização desse estudo, foi criado um projeto de pós-edição monolíngue utilizando o programa Translog 2000⁶, um programa desenvolvido por Jakobsen e Schou (1999) que registra todos os pressionamentos de teclas e movimentos de mouse realizados durante um processo de produção textual.

Antes da execução da tarefa, os estudantes receberam instruções sobre como utilizar o programa e fizeram uma pequena tarefa de treinamento para que se familiarizassem com as funcionalidades dele, como, por exemplo, o Dicionário incorporado no programa. Durante a realização do experimento realizado nesse estudo, esse Dicionário apresentava as outras opções de tradução que o Google Translate fornece, além daquelas presentes no TA traduzido automaticamente, para que o participante pudesse usá-las, caso desejasse, durante o processo de pós-edição.

Após a conclusão da tarefa de pós-edição monolíngue, os participantes foram solicitados a responder, por escrito, perguntas específicas sobre o processo em um protocolo retrospectivo guiado escrito, conforme o Anexo 3.

Conforme se observa no Quadro 2 a seguir, a média de palavras dos textos-alvo pós-editados por esses participantes foi de 196 palavras, com um total de 1.567 palavras

6 O *download* do Translog II, uma versão atualizada do programa com rastreamento ocular integrado, pode ser feito gratuitamente no site <<https://sites.google.com/site/centretranslationinnovation/translog-ii>>.

nesse subcorpus.

Quadro 2: Número de palavras dos textos-alvo pós-editados pelos participantes.

Participante	Palavras
P01	199
P02	194
P03	201
P04	200
P05	190
P06	202
P07	192
P08	189

Fonte: Elaborado pela autora.

Como mencionado anteriormente, o TA traduzido automaticamente pelo Google Translate continha 201 palavras. Observa-se assim que, conforme dados apresentados no Quadro 2, houve pouca alteração na quantidade de palavras nos textos-alvo pós-editados pelos participantes na comparação com o insumo da tradução automática. Além disso, ao somar esse número de palavras com o número de palavras do TF, das traduções publicadas e da tradução automática do trecho sob análise, o tamanho do *corpus* desta pesquisa atinge um total de 190.977 palavras, que é considerado um *corpus* pequeno-médio, segundo Berber Sardinha (2004, p. 26).

3.2 Procedimentos de análise

A ferramenta WordSmith Tools foi utilizada para selecionar o trecho a ser analisado nos Estudos 1, 2 e 3, a partir da lista de palavras que ela gera ao ler o arquivo .txt do TF de *Heart of Darkness*. Essa lista contém todas as palavras do TF e inclui também a frequência de ocorrência dessas palavras, dentre outros dados. O Quadro 3 a seguir lista as 10 palavras mais frequentes desse TF.

Quadro 3: Lista das 10 palavras mais frequentes no TF.

Número	Palavra	Frequência	Porcentagem
1	The	2.292	5,91%
2	Of	1.372	3,54%
3	A	1.149	2,96%
4	I	1.133	2,92%
5	And	991	2,55%
6	To	894	2,30%
7	Was	670	1,73%
8	In	616	1,59%
9	He	593	1,53%
10	Had	503	1,30%

Fonte: Elaborado pela autora utilizando o WordSmith Tools.

Nesse quadro, observa-se que nenhum adjetivo ou substantivo é apresentado dentre as palavras mais frequentes, uma vez que as categorias gramaticais fechadas, como artigos, verbos e preposições, correspondem sempre aos itens de maior frequência nas análises de *corpus*. Entretanto, observou-se na lista completa – que contém um total de 5.506 diferentes palavras e que não é reproduzida aqui por restrições de espaço – que os adjetivos são muito frequentes na obra, o que demandou bastante tempo para a pesquisa das linhas de concordância que continham esses adjetivos e a posterior escolha do trecho a ser submetido à análise. Após a consulta a essas linhas de concordância, selecionou-se o trecho que consta no Anexo 1, no qual estão presentes os seguintes grupos nominais com diferentes quantidades de adjetivos atributivos que serão analisados nesta pesquisa: *hot desire*; *strong, lusty, red-eyed devils*; *blinding sunshine*; *flabby, pretending, weak-eyed devil* e *rapacious and pitiless folly*.

Além do WordSmith Tools, o TreeTagger⁷, um programa para anotação de textos com informações de categoria gramatical e lema, foi utilizado para classificar esses grupos nominais presentes no TF, nas traduções publicadas, na tradução automática fornecida pelo Google Translate e nos textos-alvo pós-editados pelos participantes. O TreeTagger é um programa gratuito desenvolvido por Schmid (1995).

Como se pode observar nos grupos nominais indicados, em inglês os adjetivos atributivos são usados em apenas um tipo de estrutura: adjetivo + substantivo nos grupos nominais, em que os adjetivos são antepostos aos substantivos, independentemente da quantidade de adjetivos nos grupos nominais. Em português, os adjetivos podem vir pospostos ou antepostos aos substantivos, como em “linda flor” e “flor linda”, predominando assim duas estruturas: substantivo + adjetivo e adjetivo + substantivo. O escopo deste artigo não abrange discussões a respeito de prováveis mudanças semânticas advindas com a mudança da ordem desses adjetivos. Em vez disso, ele concentra-se na análise das traduções de grupos nominais com diferentes quantidades de adjetivos atributivos em traduções publicadas, conforme descrito no Estudo 1, em um sistema de tradução automática, de acordo com detalhes do Estudo 2, e na pós-edição monolíngue, conforme o Estudo 3.

3.3 Estudo 1: traduções publicadas

Os procedimentos de análise desse estudo consistiram no processamento do *corpus* (TF e quatro traduções publicadas), na pesquisa das ocorrências de adjetivos na lista de palavras do WordSmith Tools e na seleção de um parágrafo do TF contendo grupos nominais com um número variado de adjetivos atributivos. Em seguida, usando o WordSmith Tools, foi feito o alinhamento do parágrafo do TF com o parágrafo correspondente em quatro traduções publicadas de *Heart of Darkness* para identificar como foram traduzidos os grupos nominais.

Após esse alinhamento, o programa TreeTagger foi utilizado para etiquetar os grupos nominais do trecho selecionado no TF e nas traduções publicados. Em seguida, foi feita uma consulta no *corpus* Compara das traduções mais frequentes para os substantivos que compõem os grupos nominais do parágrafo selecionado.

7 Disponível em: <<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>>. Acesso em: 20 abr. 2018.

3.4 Estudo 2: tradução automática

O procedimento de análise consistiu inicialmente na identificação das traduções automáticas fornecidas para os grupos nominais do trecho selecionado que foi traduzido automaticamente do inglês para o português utilizando o Google Translate. Assim como ocorreu no Estudo 1, utilizou-se em seguida o TreeTagger para classificar as palavras desses grupos nominais. Posteriormente, foi feita uma consulta em um *corpus* de referência de português, o *Corpus* do Português, da ocorrência dos colocados mais frequentes dos substantivos desses grupos nominais traduzidos automaticamente.

3.5 Estudo 3: pós-edição monolíngue

Concluída a tarefa de pós-edição monolíngue pelos participantes, identificou-se no TA como esses participantes traduziram os adjetivos atributivos dos grupos nominais do trecho selecionado. Em seguida, foi utilizado o TreeTagger para classificar esses grupos nominais.

4 Resultados e discussão

4.1 Estudo 1: traduções publicadas

Os resultados apresentados no Quadro 4 indicam as traduções de grupos nominais que continham adjetivos atributivos em quatro traduções publicadas de *Heart of Darkness*.

Quadro 4: Traduções dos grupos nominais com adjetivos atributivos em traduções publicadas.

GN	Trevisan (1984)	Santarrita (1996)	O'Shea (2008)	Cyrino (2011)
<i>exact cost</i> (JJ + NN)	medir as consequências (V + DET + NOM)	custo exato (NOM + ADJ)	qualquer custo (ADJ + NOM)	custo exato (NOM + ADJ)
<i>hot desire</i> (JJ + NN)	demônio da luxúria (NOM + PRP + DET + NOM)	desejo ardente (NOM + ADJ)	desejo ardente (NOM + ADJ)	desejo ardente (NOM + ADJ)
<i>blinding sunshine</i> (VVG + NN)	sol ofuscante (NOM + ADJ)	sol cegante (NOM + ADJ)	luminosidade ofuscante (NOM + ADJ)	luz ofuscante do sol (NOM + ADJ + PRP + DET + NOM)
<i>rapacious and pitiless folly</i> (JJ + CC + JJ + NN)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura rapace e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	estupidez opressora e cruel (NOM + ADJ + CONJ + ADJ)
<i>strong, lusty, red-eyed devils</i> (JJ + JJ + JJ + NNS)	demônios vigorosos, potentes, sanguíneos (NOM + ADJ + ADJ + ADJ)	demônios fortes, sequiosos, de olhos sanguinolentos (NOM + ADJ + ADJ + PRP + NOM + ADJ)	demônios fortes, viris, de olhos avermelhados (NOM + ADJ + ADJ + PRP + NOM + ADJ)	demônios de olhos vermelhos, fortes e robustos (NOM + PRP + NOM + ADJ + ADJ + CONJ + ADJ)
<i>flabby, pretending, weak-eyed devil</i> (JJ + VVG + JJ + NN)	demônio balofo, pretencioso e velhaco (NOM + ADJ + ADJ + CONJ + ADJ)	demônio mole, de olhar débil (NOM + ADJ + PRP + NOM + ADJ)	demônio flácido, dissimulado, de olhar débil (NOM + ADJ + ADJ + PRP + NOM + ADJ)	demônio flácido, dissimulado, de olhos débeis (NOM + ADJ + ADJ + PRP + NOM + ADJ)

Fonte: Elaborado pela autora a partir dos resultados do programa TreeTagger.

As siglas apresentadas para cada categoria gramatical nesse quadro são indicadas conforme a língua da versão utilizada do programa. A versão em inglês foi utilizada para

anotar o TF. Nela, JJ refere-se ao adjetivo, CC à conjunção coordenada, NN ao nome, ou seja, ao substantivo, no singular, VVG a uma forma verbal no gerúndio ou particípio e NNS a um nome no plural. A versão em português do TreeTagger foi utilizada para a anotação dos grupos nominais com adjetivos atributivos presentes no trecho sob análise em cada uma das traduções publicadas, na tradução automática e nos textos-alvo produzidos pelos participantes. Nessa versão em português, NOM refere-se a nome, ADJ ao adjetivo, PRP à preposição, NOM ao nome (singular e plural), V ao verbo principal, CONJ à conjunção e DET ao determinante.

Não se observa nos resultados apresentados nesse quadro um padrão de traduções dos grupos nominais nas traduções publicadas do grupo nominal (GN) *exact cost* apesar de dois tradutores terem optado pela tradução mais literal “custo exato”, com a tradução de adjetivo + nome em inglês por nome + adjetivo em português. Outro tradutor optou pela tradução do GN no TA utilizando as mesmas classes do TF, entretanto ele usou uma tradução pelo sentido “qualquer custo”. Uma única tradução publicada foi feita de maneira diferente, utilizando-se verbo + determinante + nome em “medir as consequências”, passando parte do sentido para o verbo “medir”.

No caso do GN *hot desire*, também composto por adjetivo + nome, a tradução dele por nome + adjetivo foi a mesma “desejo ardente” em três das quatro traduções publicadas. Apenas um dos tradutores preferiu traduzir esse GN por “demônio da luxúria”, utilizando, portanto, nome + preposição + determinante + nome.

Assim como no GN *hot desire*, o terceiro exemplo de GN com adjetivo atributivo *blinding sunshine* possui adjetivo + nome; observa-se que, nas traduções publicadas, três dos quatro tradutores utilizaram as classes nome + adjetivo. Entretanto, diferente do que ocorreu com a opção de tradução para *hot desire*, esses tradutores usaram substantivos e/ou adjetivos diferentes ao traduzirem esses adjetivos (“sol ofuscante”, “sol cegante”, “luminosidade ofuscante”). O único tradutor que utilizou classes diferentes dos demais na tradução do GN *blinding sunshine* optou por “luz ofuscante do sol”, mudando o adjetivo + nome no TF por nome + adjetivo + preposição + determinante no TA. Nesse caso, o tradutor usou três palavras nessa tradução, precisando recorrer à palavra “do” para complementar o GN em português.

No GN *rapacious and pitiless folly*, surge um elemento diferente, a conjunção “and”. Ela é posposta ao primeiro adjetivo nesse GN, com o adjetivo + conjunção + adjetivo + nome, sendo traduzido com as mesmas classes gramaticais em todas as traduções publicadas (nome + adjetivo + conjunção + adjetivo), incluindo uma mesma opção de tradução de dois tradutores diferentes (“loucura voraz e impiedosa”, “loucura rapace e impiedosa”, “estupidez opressora e cruel”).

Apesar de esses exemplos de grupos nominais que contém até dois adjetivos atributivos não possuírem um padrão rígido de tradução, observa-se uma tendência a traduzir esses grupos nominais alterando-se a ordem dos adjetivos com o nome, como é de se esperar na tradução desse tipo de GN da língua inglesa para a língua portuguesa, e a inclusão de conjunção e de determinante. Entretanto, quando se tem mais de dois adjetivos, já se observa maior variação na tradução desses grupos nominais com adjetivos atributivos, em que há inserção de conjunções e preposições.

No GN *strong, lusty, red-eyed devils*, por exemplo, em que há três adjetivos atributivos seguidos de um nome no plural, observam-se três configurações diferentes nas traduções publicadas: nome + adjetivo + adjetivo + adjetivo (“demônios vigorosos,

potentes, sanguíneos”), nome + adjetivo + adjetivo + preposição + nome + adjetivo (“demônios fortes, sequiosos, de olhos sanguinolentos” e “demônios fortes, viris, de olhos avermelhados) e nome + preposição + nome + adjetivo + adjetivo + conjunção + adjetivo (“demônios de olhos vermelhos, fortes e robustos”).

Com relação ao GN *flabby, pretending, weak-eyed devil*, em que há adjetivo + verbo no gerúndio + adjetivo + nome no singular, observou-se também uma maior variação, havendo também três configurações diferentes nas quatro traduções publicadas: nome + adjetivo + adjetivo + conjunção + adjetivo (“demônio balofo, pretencioso e velhaco”), nome + adjetivo + preposição + nome + adjetivo (“demônio mole, de olhar débil”) e nome + adjetivo + adjetivo + preposição + nome + adjetivo (“demônio flácido, dissimulado, de olhar débil” e “demônio flácido, dissimulado, de olhos débeis”).

Ao verificar no *corpus* paralelo bilíngue Compara as traduções mais comuns para os substantivos dos grupos nominais, os resultados apresentados no Quadro 5 foram encontrados. Nesse quadro, a sigla NA significa “não se aplica” e foi utilizada para indicar que não houve tradução menos frequente, pois as traduções mais frequentes apresentaram a mesma frequência (1).

Quadro 5: Traduções mais comuns dos substantivos dos grupos nominais.

Substantivo	Tradução mais frequente e frequência	Linhas de concordância e ocorrências	Outras traduções menos frequentes e frequência
<i>cost</i>	custo: 8	30 linhas de concordância, com 31 ocorrências de <i>cost</i>	outros substantivos: 5 (à custa: 1, despesa: 1, preço: 3) custar e variantes desse verbo: 12
<i>desire</i>	desejo: 57	70 linhas de concordância, com 87 ocorrências de <i>desire</i>	outros substantivos: 17 (vontade: 16, ímpeto: 1) desejar e variantes desse verbo: 8 outros verbos: “latejava-lhe, reclamando-a”; quiser, quero, quis: 4 omissão: 1
<i>devils</i>	demônios: 1 coitados: 1 eles: 1 infelizes: 1	4 linhas de concordância, com 4 ocorrências de <i>devils</i>	NA
<i>devil</i>	diabo: 34	58 linhas de concordância, com 64 ocorrências de <i>devil</i>	demônio: 12 outros substantivos e adjetivos: 16 (raio: 3, brejeiro, sujo, miserável: 2, inferno, demo, diabinho, Apre, estupor, bruto, desgraçado: 1) outras traduções: 1 (Que vida!) omissão: 1
<i>sunshine</i>	calor do sol: 1 sol: 1	2 linhas de concordância, com 2 ocorrências de <i>sunshine</i>	NA

Fonte: Elaborado pela autora a partir da consulta ao *corpus* Compara.

Nesses resultados, observa-se uma predominância de traduções mais literais para os substantivos dos grupos nominais sob análise nesta pesquisa, não havendo muita variação nas possíveis traduções para os substantivos que aparecem nesses grupos nominais. Em números absolutos, o substantivo que teve maior variação foi *devil*, com 13 variações de tradução no Compara em 64 ocorrências desse substantivo em inglês.

Entretanto, em números relativos, a maior variação foi na tradução do substantivo *devils*, que apresentou 4 traduções diferentes para 4 ocorrências em inglês.

4.2 Estudo 2: tradução automática

No Quadro 6, são apresentados resultados referentes ao Estudo 2, em que se verificou como os grupos nominais foram traduzidos automaticamente pelo Google Translate. A classificação das categorias gramaticais dos grupos nominais indicadas entre parênteses nesse quadro foi feita automaticamente pelo TreeTagger.

Conforme o Quadro 6, a tradução de grupos nominais com até dois adjetivos na tradução automática segue uma das possíveis estruturas descritas para o português: o nome seguido do(s) adjetivo(s). No caso de haver dois adjetivos no GN em inglês em que há a ligação por meio de uma conjunção coordenada, observa-se que a conjunção é mantida na tradução automática e que a inversão da ordem do substantivo, que passa para a posição inicial, é bem-sucedida, apesar de não haver uma concordância do adjetivo com o substantivo do GN. Constata-se também que, mesmo o português permitindo tanto a posição anteposta quanto a posição posposta do adjetivo em um GN, na tradução automática há uma preferência pela posição posposta do adjetivo.

Quadro 6: Traduções dos grupos nominais com adjetivos atributivos na tradução automática fornecida pelo Google Translate.

GN	Tradução automática
<i>exact cost</i> (JJ + NN)	custo exato (NOM + ADJ)
<i>hot desire</i> (JJ + NN)	desejo quente (NOM + ADJ)
<i>strong, lusty, red-eyed devils</i> (JJ + JJ + JJ + NNS)	fortes, corpulentos, de olhos vermelhos demônios (ADJ + ADJ + PRP + NOM + ADJ + NOM)
<i>blinding sunshine</i> (WG + NN)	luz do sol ofuscante (NOM + PRP+DET + NOM + ADJ)
<i>flabby, pretending, weak-eyed devil</i> (JJ + VVG + JJ + NN)	frouxo, fingindo, diabo fraco de olhos (ADJ + V + NOM + ADJ + PRP + NOM)
<i>rapacious and pitiless folly</i> (JJ + CC + JJ + NN)	loucura voraz e impiedoso (NOM + ADJ + CONJ + ADJ)

Fonte: Elaborado pela autora a partir dos resultados do programa TreeTagger.

Ainda de acordo com o Quadro 6, percebe-se que a tradução de grupos nominais com três adjetivos atributivos na tradução automática já apresenta problemas, já que o Google Translate, apesar de reconhecer um substantivo, não conseguiu fazer uma inversão bem-sucedida do substantivo. Isso pode provocar um erro de interpretação para o pós-editor de um texto com grupos nominais com mais de dois adjetivos atributivos se ele estiver fazendo uma pós-edição monolíngue, ou seja, sem acesso ao TF, ou mesmo na pós-edição bilíngue, caso o pós-editor não esteja atento.

Para complementar o Estudo 2, foi feita uma consulta ao *Corpus* do Português dos colocados mais frequentes na seção *Fiction* do *corpus* usando como palavra de busca as traduções fornecidas automaticamente pelo Google Translate para os substantivos dos grupos nominais. Essa consulta foi feita usando o horizonte de 2:2, ou seja, até duas palavras à esquerda e à direita do nóculo (substantivo). Os resultados dessa consulta são

apresentados no Quadro 7 a seguir⁸.

Quadro 7: Colocados mais frequentes com as traduções automáticas fornecidas pelo Google Translate para os substantivos dos grupos nominais sob análise

Substantivo	10 colocados mais frequentes (2:2) na seção <i>Fiction</i>	Mais frequente ponderado	Maior MI ponderada
custo	muito, todo, qualquer, plúmbeo, coadjuvado, balbuciadas, flácidos, extras, latentes	muito (28)	todo (4,92)
desejo	seu, enorme, irreprimível, ardente, absurdo, único, louco, imenso, secreto, furioso	seu (11 7)	irreprimível (7,84)
demônios	vingadora, entalados, XVI, desses, alguns	vingadora (1), entalados (1), XVI (1), desses (1), alguns (1)	vingadora (12,52)
luz	acesa, clara, elétrica, mortíça, crua, vermelha, meia, baça, fraça, frouxa	acesa (27)	mortíça (8,31)
diabo	pobre, leve, semelhante, encastado, crismado, cianídrico, engarrafado, samento, zangados, abomináveis	pobre (27)	pobre (4,60)
loucura	aquela, essa, mansa, rematada, desastrada, polar, senil, irremediável, verdadeira, humana	aquela (12)	rematada (10,17)

Fonte: Elaborado pela autora a partir dos resultados na consulta ao Corpus do Português.

Nesses resultados, constata-se que nenhum dos substantivos dos grupos nominais apresentou os adjetivos usados na tradução automática fornecida pelo Google Translate como um dos colocados mais frequentes e/ou com a maior MI (*Mutual Information*) ponderada. Ressalta-se, entretanto, que uma mudança do horizonte de pesquisa poderia apresentar algum desses adjetivos da tradução automática como colocado mais frequente ou com maior média ponderada.

A partir desses resultados, pode-se afirmar que, apesar de o Google Translate possuir uma grande base de dados, a primeira opção fornecida como tradução automática para o adjetivo não correspondia ao uso, ou seja, não era um colocado. Observa-se, no entanto, que o adjetivo “ardente”, que aparece como um dos 10 colocados mais frequentes para o substantivo “desejo”, foi o adjetivo utilizado por três dos quatro tradutores cujas traduções publicadas foram analisadas.

4.3 Estudo 3: pós-edição monolíngue

Conforme mencionado anteriormente, o Estudo 3 concentrou-se na análise da tradução de grupos nominais com adjetivos atributivos nos textos-alvo pós-editados por estudantes sem consulta ao TF. O Quadro 7 apresenta as traduções fornecidas pelos estudantes para os grupos nominais analisados.

Observa-se, no Quadro 7, que os grupos nominais com apenas um adjetivo atributivo em inglês: *hot desire*, *exact cost* e *blinding sunshine*, traduzidos automaticamente como “custo exato”, “desejo quente” e “luz do sol ofuscante” respectivamente não foram alterados por sete participantes; apenas P08 os alterou. Há, portanto, um padrão de processamento desses grupos na pós-edição monolíngue, com

⁸ Horizonte refere-se à distância máxima estabelecida pelo pesquisador entre o nóculo e a palavra de contexto para definir os colocados.

predominância de nome + adjetivo (“custo exato”, “desejo quente” e “desejo quente”) e nome + preposição + determinante + nome + adjetivo (“luz do sol ofuscante”).

Esse padrão só não ocorre na pós-edição realizada por P08, que mantém nome + adjetivo usando “desejo ardente” para a tradução de *hot desire*. Nesse caso, altera-se apenas a tradução automática “quente” para “ardente”. Não se observa, no entanto, um padrão na tradução dos grupos nominais *exact cost* e *blinding sunshine* (adjetivo + nome), uma vez que o participante alterou a tradução automática que seguia o padrão nome + adjetivo (“custo exato”) e nome + preposição + determinante + nome + adjetivo (“luz do sol ofuscante”) para advérbio + pronome + verbo + pronome + verbo (“quanto isso ia me custar”) e nome + adjetivo + preposição + determinante + nome (“luz ofuscante do sol”).

Quadro 7: Traduções dos grupos nominais com adjetivos atributivos nos textos pós-editados.

GN	P01	P02	P03	P04	P05	P06	P07	P08
<i>exact cost</i> (JJ + NN)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	custo exato (NOM + ADJ)	quanto isso ia me custar (ADV + P + V+ P + V)
<i>hot desire</i> (JJ + NN)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo quente (NOM+ ADJ)	desejo ardente (NOM+ ADJ)
<i>strong, lusty, red-eyed devils</i> (JJ + JJ + JJ + NNS)	fortes, corpulentos, demônios de olhos vermelhos (ADJ + ADJ + NOM + PRP + NOM + ADJ)	fortes, corpulentos, de olhos vermelhos, demônios (ADJ + ADJ + PRP + ADJ + NOM)	fortes, corpulentos, demônios de olhos vermelhos (ADJ + ADJ + NOM + PRP + NOM + ADJ)	fortes, corpulentos, de olhos vermelhos demônios (ADJ + ADJ + PRP + ADJ + NOM)	fortes, corpulentos, de olhos vermelhos demoníacos (ADJ + ADJ + PRP+ NOM + ADJ + ADJ)	fortes, corpulentos, de olhos vermelhos demoníacos (ADJ + ADJ + PRP+ NOM + ADJ + ADJ)	fortes, corpulentos, de olhos vermelhos demoníacos (ADJ + ADJ + PRP + ADJ + ADJ)	Esses possuíam olhos vermelhos, eram fortes e encorpados (P + V+ NOM + ADJ + V + ADJ + CONJ + ADJ)
<i>blinding sunshine</i> (NN)	luz do sol ofuscante (NOM + PRP+DET + NOM + ADJ)	luz do sol ofuscante (NOM + PRP + DET + NOM + ADJ)	luz do sol ofuscante (NOM + PRP+DET + NOM + ADJ)	luz do sol ofuscante (NOM + PRP + DET + NOM + ADJ)	luz do sol ofuscante (NOM + PRP + DET+ NOM + ADJ)	luz do sol ofuscante (NOM + PRP + DET+ NOM+ADJ)	luz do sol ofuscante (NOM + PRP+DET + NOM + ADJ)	luz ofuscante do sol (NOM + ADJ + PRP+DET + NOM)
<i>flabby, pretending, weak-eyed devil</i> (JJ + VVG + JJ + NN)	frouxo, fingido, diabo fraco de olhos (ADJ + ADJ + NOM + ADJ + PRP + NOM)	diabo frouxo e fingido, fraco de olhos (NOM + ADJ + CONJ + ADJ + PRP + NOM)	frouxo, fingido, com olhos fracos de um diabo (frouxo (ADJ + V + PRP + NOM + ADJ + PRP + NOM)	frouxo, fingido, diabo fraco de olhos (ADJ + V + NOM + ADJ + PRP + NOM)	frouxo, fingido, diabo fraco de olhos (ADJ+V+N OM + ADJ + PRP + NOM)	frouxo, fingido, diabo fraco de olhos (ADJ + V + NOM + ADJ + PRP + NOM)	frouxo, um diabo fraco de olhos (ADJ + DET + NOM + ADJ + PRP + NOM)	demônio impiedoso se fingindo de fraco (NOM + ADJ + P + V+PRP + ADJ)
<i>rapacious and pitiless folly</i> (JJ + CC + JJ + NN)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedoso (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedoso (NOM + ADJ + CONJ + ADJ)	loucura voraz e impiedosa (NOM + ADJ + CONJ + ADJ)	loucura voraz estampada nos olhos (NOM + ADJ + V + PRP + DET + NOM)

Fonte: Elaborado pela autora a partir dos textos-alvo pós-editados e do TreeTagger.

Os resultados indicam que, durante o processo de pós-edição monolíngue, os estudantes tiveram dificuldade para processar os grupos nominais com mais de dois adjetivos atributivos, sendo esses grupos os mais alterados pelos participantes. Desse modo, essa dificuldade pode ser observada por meio de uma análise do produto em concomitância com os protocolos retrospectivos guiados sobre o processo, uma vez que, nas respostas a perguntas desse protocolo, a maioria dos estudantes relatou ter dificuldades para pós-editar grupos nominais para a produção dos textos-alvo.

O GN *strong, lusty, red-eyed devils* não foi alterado apenas por P04, que manteve a tradução automática. Entretanto, algumas alterações realizadas pelos participantes nesse GN foram pequenas, como a inclusão de uma vírgula por P02. Foram constatados que, em três textos-alvo pós-editados (P05, P06 e P07), o substantivo do GN no inglês tornou-se em adjetivo, produzindo, portanto, um novo GN em português “olhos demoníacos”.

Com relação ao GN *flabby, pretending, weak-eyed devil* traduzido automaticamente como “frouxo, fingindo, diabo fraco de olhos”, o mesmo não foi alterado por três (P04, P05 e P06) dos oito participantes, os quais mantiveram a tradução desse GN tal como fornecida pelo Google Translate. Esse GP pode ter sido o mais difícil para processar e pós-editar, conforme relato de um dos participantes: “A última parte do diabo foi mais difícil”. A não alteração da tradução automática é referida por um dos participantes como uma estratégia para lidar com esses trechos considerados difíceis: “Mantive os trechos tal como estavam porque não consegui compreendê-los”, o que justifica o fato de apenas três participantes terem modificado a tradução automática desse GN.

O GN que apresentava dois adjetivos atributivos, *rapacious and pitiless folly*, foi alterado por cinco dos oito participantes; quatro deles alteraram apenas a concordância do substantivo com o adjetivo. Uma das alterações realizadas nesse GN afetou o GN anterior, uma vez que a referência aos “olhos” que aparecia em *flabby, pretending, weak-eyed devil* foi transferida para esse GN traduzido como “loucura voraz estampada nos olhos”. A dificuldade para processar e pós-editar esses dois grupos nominais pode ser constatada no relato de outro participante quando ele afirma que uma das partes que ele considerou mais difícil pós-editar foi essa que continha esses dois grupos nominais “que eu gostaria de familiarizar com um frouxo, fingindo, diabo fraco de olhos de uma loucura voraz e impiedoso”, apresentando também uma justificativa para tal dificuldade: “foi difícil encontrar o grupo nominal [...] eu não sabia se era ‘os olhos’, ou ‘diabo’”. Essas dúvidas poderiam ser esclarecidas, segundo esse participante, com o acesso ao TF: “possivelmente eu mudaria os grupos nominais que pós-editei, pois aí seria mais fácil ver a que eles realmente se referiam”.

5 Considerações finais

Na realização deste estudo, constataram-se padrões sintáticos e lexicais recorrentes nas traduções de grupos nominais com até dois adjetivos atributivos nas traduções publicadas. No caso de grupos nominais com mais de dois adjetivos, observou-se que a estratégia mais frequentemente utilizada pelos tradutores é a inclusão de preposições e conjunções em português ao traduzirem esses adjetivos, revelando um texto-alvo mais fluido em português.

Observou-se uma maior variedade sintática e lexical nas traduções de grupos

nominais com três ou mais adjetivos atributivos nas traduções publicadas. Além disso, as traduções mais frequentes de substantivos no *corpus* Compara foram as traduções usadas pelos tradutores da obra *Heart of Darkness*, o que pode indicar que esses tradutores fazem suas escolhas tradutórias com base na observação do uso. Além disso, as pesquisas nos corpora de referência consultados reforçam a necessidade de ponderação nas frequências e nos dados de MI.

Com relação ao Estudo 2, a realização desta pesquisa demonstrou a necessidade de expor para os estudantes de tradução a dificuldade que os sistemas de tradução automática têm para traduzir grupos nominais com mais adjetivos, como *strong*, *lusty*, *red-eyed devils* e *flabby*, *pretending*, *weak-eyed devil*. Ademais, deve-se ensinar os alunos estratégias para solucionar essa dificuldade e outras que surgirem para que os estudantes se familiarizem com os sistemas e se sintam mais seguros ao executarem tarefas de pós-edição bilíngue ou monolíngue.

No que se refere ao Estudo 3, as poucas alterações realizadas pelos estudantes nos textos pós-editados parecem indicar que eles não se sentem confiantes para fazer alterações que podem levar a interpretações incorretas do TA, o que pôde ser constatado no relato de um dos participantes que afirmou utilizar a estratégia de manter inalterado o trecho considerado difícil. Essas interpretações incorretas podem ser provocadas pelo próprio sistema de tradução automática, por isso é necessário que os estudantes sejam treinados para usá-lo.

É necessário ressaltar que o uso de *corpora* paralelo bilíngue e em português em sala de aula pode ajudar os estudantes de cursos de tradução a sentirem mais confiantes em suas escolhas tradutórias para assim produzir textos-alvo que representam a língua em uso. Desse modo, este estudo busca contribuir para o delineamento de atividades a serem desenvolvidas com os estudantes usando *corpus*, a fim de levá-los a refletir sobre o próprio processo de tradução e pós-edição e sobre o uso de sua própria língua nessas tarefas. Nesse caso, propõe-se utilizar a metodologia apresentada no Estudo 1, juntamente com a exposição aos estudantes das dificuldades que os sistemas de tradução automática possuem conforme relatado no Estudo 2 e das estratégias para solução dessas dificuldades para assim treiná-los a produzir um texto pós-editado tão fluido quanto uma tradução publicada.

Referências

ASSIS, R. C. *A representação de europeus e de africanos como atores sociais em Heart of darkness (O coração das trevas) e em suas traduções para o português: uma abordagem textual da tradução*. 2009. 267 f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.

BAKER, M. *Corpora in translation studies: an overview and some suggestions for future research*. *Target*, Amsterdam: John Benjamins, v. 7, n. 2, p. 223-243, 1995.

BAKER, M. *Corpus linguistics and translation studies: implications and applications*. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (Ed.). *Text and technology: in honour of John Sinclair*. Amsterdã/Filadélfia: John Benjamins Publishing Company, p. 233-250,

1993.

BARCELLOS, C. P. *Estilo da tradução, convencionalidade e mudanças na tradução: um estudo de caso sobre os padrões de escolhas do tradutor Paulo Henrique Britto*. 2016. 196 f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.

BARCELLOS, C. P. *O estilo de tradutores: apresentação do discurso no corpus paralelo Heart of Darkness/(No) Coração das Trevas*. 2011. 154 f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.

BERBER SARDINHA, A. *Linguística de corpus*. Barueri: Manole, 2004.

BERBER SARDINHA, A. Uso de corpora na formação de tradutores. *DELTA*. São Paulo, 2003, v. 29, p. 43-70.

BLAUTH, T. P. O indizível em *Heart of Darkness* e duas traduções brasileiras: um estudo de estilística tradutória. In: ZAVAGLIA, C.; STUPIELLO, E. N. A. (Org.). *Tendências contemporâneas dos Estudos da Tradução* - volume 2. São José Do Rio Preto: Universidade Estadual Paulista, 2015.

BLAUTH, T. P. A paisagem indescritível em *Heart of Darkness* e duas traduções brasileiras: um estudo exploratório de estilística tradutória com base em corpus. *Belas Infieis*, v. 3, n. 1, p. 185-197, 2014.

CASTRO, M. C. *O perfil estilístico de quatro tradutores de Heart of Darkness para o espanhol: uma investigação de mudanças de tradução (shifts in translation) baseada em padrões de itens lexicais de um corpus paralelo*. 2015. 181 f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.

CONRAD, J. *Heart of Darkness*. London: Penguin Books, 1994.

HATIM, B.; MUNDAY, J. *Translation: an advanced resource book*. New York: Routledge. 2004.

HOLMES, J. S. The Name and Nature of Translation Studies. In: VENUTI, L. (Ed.). *The Translation Studies Reader*. Londres e Nova York: Routledge, 1972/2000. p. 172-185.

HURTADO ALBIR, Amparo. *Traducción y Traductología: Introducción a la Traductología*. Madrid: Cátedra, 2001.

JAKOBSEN, A. L.; SCHOU, L. Logging target text production with Translog. *Copenhagen Studies in Language*, Copenhagen: Samfundslitteratur, v. 24, p. 9-20, 1999.

JAKOBSON, R. On linguistic aspects of translation. In: VENUTI, L. (Ed.). *The Translation*

Studies Reader, 1959/2004. p. 138-143.

KENDALL, T. Corpora from a sociolinguistic perspective. In: GRIES, S. T. (Ed.), *Corpus Studies: future directions*, Special Issue of *Revista Brasileira de Linguística Aplicada*, v. 11, n. 2, p. 361-389. 2011.

KENNY, D. *Lexis and creativity in translation: a corpus-based study*. Manchester, UK e Northampton MA: St Jerome Publishing, 2001, p. 22-72.

KOBY, G. S. Editor's Introduction – Post-editing of Machine Translation Output: Who, What, Why, and How (much). In: KOBY, G. (Ed.). *Repairing texts: empirical investigations of machine translation post-editing processes*. Tradução de G. S. Koby; G. M. Shreve; K. Mischericow; S. Litzer. Kent, Ohio, London: Kent State University Press, 2001.

KOEHN, P. Enabling monolingual translators: post-editing vs. options. In: *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL*. Los Angeles, Califórnia, p. 537-545, 2010. Disponível em: <<http://www.aclweb.org/anthology/N10-1078>>. Acesso em: 4 jan. 2014.

KÖNIGS, F. G. Was beim Übersetzen passiert. Theoretische Aspekte, empirische Befunde und praktische Konsequenzen. *Die Neueren Sprachen* 2, p. 162-185. 1987.

KRINGS, H. *Repairing texts: empirical investigations of machine translation Post-editing processes*. Tradução de G. S. Koby; G. M. Shreve; K. Mischericow; S. Litzer. Kent, Ohio, London: Kent State University Press, 2001.

KRINGS, H. P. *Was in den Klipfen von Übersetzern vorgeht: eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern*. Tübingen: Narr. 1986.

KYTÖ, M. Corpora and historical linguistics. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v. 11, n. 2, p. 417-457. 2011.

LINDQUIST, H. The use of corpus-based studies in the preparation of handbooks for Translators. In: WILSS, W.; THOME, G. (Ed.). *Translation theory and its Implementation in the teaching of translating and interpreting*. Tübingen: Narr, 1984, p. 260-270.

MAGALHÃES, C. M.; BLAUTH, T. P. Estilo do tradutor: um estudo do uso do itálico, palavras estrangeiras e itens culturais específicos por seis tradutores do português de *Heart of Darkness*. In: VIANA, Vander; TAGNIN, S. E. O. (Org.). *Corpora na tradução*. 2015. p. 171-209.

MAGALHÃES, C. M.; BARCELLOS, C. P. Estilo de tradutores: estudo baseado no *corpus Heart of Darkness*/(No)Coração Das Trevas. In: IBANOS, A. M. T.; MOTTIN, L. P.; SARMENTO, S.; BERBER SARDINHA, T. (Org.). *Pesquisa e perspectivas em linguística de corpus*. Campinas, SP: Mercado de Letras, 2014. p. 441-468.

McCARTHY, M. J.; O'KEEFFE, A. Historical perspective: what are corpora and how have they evolved? In: O'KEEFFE, A.; McCARTHY, M. J. (Ed.). *The Routledge Handbook of Corpus Linguistics*. Londres: Routledge, 2010. p. 3-13.

McENERY, T.; WILSON, A. *Corpus Linguistics: an introduction*. Edinburgh: Edinburgh UP, 2001.

MEUNIER, F. Corpus linguistics and second/foreign language learning: exploring multiple paths. *Revista Brasileira de Linguística Aplicada*, Belo Horizonte, v. 11, n. 2, p. 459-477, 2011. Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1984-63982011000200008>. Acesso em: 5 mar. 2018.

MEYER, C. F. *English Corpus Linguistics: an introduction*. Cambridge: Cambridge University Press, 2004.

MONTENEGRO, M. S. *O perfil de quatro tradutores portugueses de Heart of darkness: um estudo do estilo do tradutor com base em corpus*. 2015. 181 f. Tese (Doutorado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte.

MUNDAY, J. *The Routledge Companion to Translation Studies*, London: Routledge, 2009.

OLOHAN, M.; BAKER, M. Reporting that in translated English: evidence for subconscious processes of explicitation?. *Across Languages and Cultures*, v. 1, n. 2, p. 141-158, 2000.

RICHERS, R. H.; BOOTH, A. D. Some methods of mechanized translation. In: LOCKE, W.; BOOTH, A. D. (Ed.). *Machine translation of languages: fourteen essays*. Cambridge, Mass.: MIT Press, 1955, p. 24-46. Acesso em 15 maio 2018. Disponível em: <<https://pdfs.semanticscholar.org/b7d6/3112b3da746efa8ae7effb33609fd9a5e85c.pdf>>. Acesso em: 25 jun. 2018.

SANTOS, P. Tradução automática. In: MATEUS, M. H.; BRANCO, A. H. (Org.). *Engenharia da linguagem*. Lisboa: Colibri, 1995. p. 121-128.

SÃO JERÔNIMO. *Carta a Pamáquio sobre os problemas da tradução*, ep. 27. Introdução, revisão de edição, tradução e notas de Aires A. Nascimento. Lisboa: Edições Cosmos, 1995.

SCOTT, M. *WordSmith Tools version 6*, Stroud: Lexical Analysis Software, 2012.

STUBBS, M. Language corpora. In: DAVIES, A; ELDER, C. (Ed.). *The Handbook of Applied Linguistics*. Oxford: Blackwell, 2004. p. 106-132.

Recebido em dia 28 de março de 2018.
Aprovado em dia 20 de abril de 2018.

ANEXOS

Anexo 1 – Trecho da obra *Heart of Darkness* selecionado para análise

Instead of going up, I turned and descended to the left. My idea was to let that chain-gang get out of sight before I climbed the hill. You know I am not particularly tender; I've had to strike and to fend off. I've had to resist and to attack sometimes—that's only one way of resisting—without counting the exact cost, according to the demands of such sort of life as I had blundered into. I've seen the devil of violence, and the devil of greed, and the devil of hot desire; but, by all the stars! these were strong, lusty, red-eyed devils, that swayed and drove men—men, I tell you. But as I stood on this hillside, I foresaw that in the blinding sunshine of that land I would become acquainted with a flabby, pretending, weak-eyed devil of a rapacious and pitiless folly. How insidious he could be, too, I was only to find out several months later and a thousand miles farther. For a moment I stood appalled, as though by a warning. Finally I descended the hill, obliquely, towards the trees I had seen.

Anexo 2 – Trecho da obra *Heart of Darkness* traduzido automaticamente pelo Google Translate

Em vez de subir, eu me virei e desceu para a esquerda. Minha idéia era deixar que a cadeia de gangues sair de vista antes que eu subia a colina. Você sabe que eu não sou particularmente sensível, eu tive que atacar e para defender-se. Eu tive de resistir e atacar, por vezes - que é apenas uma maneira de resistir - sem contar o custo exato, de acordo com as exigências de tal tipo de vida que eu tinha tropeçado em. Eu vi o diabo da violência, eo diabo da ganância, eo diabo do desejo quente, mas, por todas as estrelas! estes eram fortes, corpulentos, de olhos vermelhos demônios, que oscilavam e dirigiu homens - homens, eu lhe digo. Mas como eu estava nesta encosta, eu previa que na luz do sol ofuscante de terra que eu gostaria se familiarizar com um frouxo, fingindo, diabo fraco de olhos de uma loucura voraz e impiedoso. Como insidioso que ele poderia ser, também, eu era apenas para descobrir vários meses depois e milhares de quilômetros mais longe. Por um momento eu fiquei chocado, como se de um aviso. Finalmente desci do monte, obliquamente, em direção às árvores que eu tinha visto.

Anexo 3 – Protocolo retrospectivo guiado escrito

- 1) Qual é o seu nível de conhecimento da língua-fonte do parágrafo que foi traduzido automaticamente?
- 2) Qual(is) parte(s) do parágrafo você considerou mais difícil pós-editar? Por quê?
- 3) Que estratégia você utilizou para pós-editar essa parte?
- 4) Qual(is) parte(s) do parágrafo você considerou mais fácil pós-editar? Por quê?
- 5) Se você tivesse acesso ao texto-fonte, você mudaria alguma escolha de pós-edição que fez? Qual? Por que você mudaria?