



Jorge Mario Davidson

**Tradução automática em ambientes
de memória de tradução
um estudo comparativo de dois métodos
de trabalho**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Letras/Estudos Estudos da Linguagem pelo Programa de Pós-graduação em Estudos da Linguagem da PUC-Rio.

Orientadora: Profa. Maria Cláudia de Freitas

Rio de Janeiro
setembro de 2021



Jorge Mario Davidson

**Tradução automática em ambientes
de memória de tradução
um estudo comparativo de dois métodos
de trabalho**

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de Pós-
graduação em Estudos da Linguagem da PUC-Rio.
Aprovada pela Comissão Examinadora abaixo:

Profa. Maria Cláudia de Freitas

Orientadora

Departamento de Letras – PUC-Rio

Profa. Marcia do Amaral Peixoto Martins

Departamento de Letras – PUC-Rio

Profa. Adriana Silvina Pagano

Faculdade de Letras – UFMG

Todos os direitos reservados. A reprodução, total ou parcial do trabalho, é proibida sem a autorização da universidade, da autora e do orientador.

Jorge Mario Davidson

Professor de História pelo *Instituto Superior del Profesorado Joaquín V. González* (Argentina). Mestre e Doutor em História Social pela Universidade Federal Fluminense. Professor das disciplinas Tradumática e Tradução Técnica da Especialidade em Tradução (latu sensu) da Universidade Estácio de Sá. Tradutor desde 2004, com livros publicados nas áreas de História, Saúde Pública, Arquitetura e Educação.

Ficha Catalográfica

Davidson, Jorge Mario.

Tradução automática em ambientes de memória de tradução: um estudo comparativo de dois métodos de trabalho / Jorge Mario Davidson; orientadora: Maria Cláudia de Freitas. – Rio de Janeiro: PUC, 2021.

118 f.; il.; 29,7 cm.

Dissertação (mestrado) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2017.

Inclui referências bibliográficas.

1. Letras – Dissertações. 2. Tradução. 3. Tradução automática. 4. Pós-edição. 5. CAT Tools. 6. Avaliação de tradução. I. Freitas, Maria Cláudia de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. III. Título.

CDD: 400

À Graciela e Martina, as duas mulheres incríveis com quem
tenho a imensa sorte de compartilhar a minha vida.

Agradecimentos

À Graciela, que me incentivou a sair do “conforto desconfortável” da minha mesa de trabalho para enfrentar um novo desafio, e pelo amor e apoio no mais amplo sentido da palavra.

À Martina, pela predisposição para me ajudar e por emprestar ouvidos às minhas neuroses e pensamentos de maneira incondicional.

À minha orientadora, Cláudia Freitas, pelo apoio e a grande disponibilidade para me guiar nessa pesquisa, e por me apresentar nas suas aulas um mundo de possibilidades para refletir sobre meu trabalho cotidiano.

À PUC-Rio, sem cujo apoio esse trabalho não teria sido possível, e que prorrogou gentilmente a minha permanência no programa em um momento muito difícil.

Aos funcionários do Departamento de Letras, em especial a Chiquinha, pela simpatia.

Às professoras do PPGEL que ampliaram meus horizontes, Marcia Martins e Letícia Sicuro Corrêa. Um agradecimento especial à Helena Martins, quem chacoalhou a minha cabeça com leituras e debates sempre instigantes.

Aos tradutores profissionais que de alguma maneira me inspiraram a voltar para a Universidade, Meritxell Almarza, Daniel Estill e Ricardo Souza. Um agradecimento especial à Michel Teixeira, pelo incentivo e pelos conselhos valiosos.

Aos meus alunos, que me fazem querer ser cada dia um profissional mais completo.

À Elvis de Souza pela sua inestimável colaboração.

À Cari, Fabi, Andrés, Pablo, Lina, Martín, Julieta, Sol, Gaby, pela ajuda desinteressada.

À Juno, porque não teve um dia só nesta interminável e difícil pandemia que não tenha provocado umas boas e necessárias risadas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Davidson, Jorge Mario; Freitas, Maria Cláudia de. **Tradução automática em ambientes de memória de tradução**: um estudo comparativo de dois métodos de trabalho. 2021. 119 p. Dissertação (Mestrado em Estudos da Linguagem) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2021.

Esta dissertação discute a utilização de sistemas de tradução automática em ambientes de memória de tradução (CAT), uma modalidade de trabalho cada vez mais presente no mercado de tradução especializada atual. Foi realizado um estudo experimental envolvendo quatro tradutores profissionais especializados na área de informática. Cada um dos profissionais traduziu dois textos, um deles de marketing de tecnologia e o outro altamente técnico, utilizando diferentes modalidades de trabalho. O objetivo do estudo foi verificar a existência de diferenças entre o uso de tradução automática com pós-edição no nível de segmento e o uso de tradução automática como sugestão no nível de subsegmento. As traduções foram analisadas utilizando recursos de linguística computacional por meio das seguintes métricas: variedade lexical, densidade lexical, distância de edição, considerando sequências de classes gramaticais, e produtividade. Para efeitos comparativos, foram incluídas no estudo experimental traduções 100% humanas e traduções automáticas sem pós-edição. As métricas utilizadas permitiram observar diferenças nos resultados atribuíveis às modalidades de trabalho, bem como comparar os efeitos nos diferentes tipos de textos traduzidos. Finalmente, as diversas traduções de um dos textos foram submetidas à avaliação de leitores para determinar as preferências.

Palavras-chave

Tradução; tradução automática; pós-edição; CAT Tools; avaliação de tradução.

Abstract

Davidson, Jorge Mario; Freitas, Maria Cláudia de. **Automatic translation in translation memory systems:** a comparative study of two work methods. 2021. 119 p. Dissertação (Mestrado em Estudos da Linguagem) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2021.

This dissertation addresses the use of automatic translation in translation memory systems (CAT), a fast-growing modality of work in today's specialized translation market. An experimental study was conducted with four professional translators specializing in the field of computing. Each professional translated two texts, one about technology marketing and the other, a highly technical document, using different modalities of work. The purpose of the study was to identify any differences resulting from the use of automatic translation, with segment-based post-editing, and the use of automatic translation as sub-segment translation suggestions. The resources of computational linguistics were employed to analyze the translations, considering the following metrics: lexical diversity, lexical density, edit distance, taking into account grammatical sequences, and productivity. For comparative purposes, the experimental study included 100% human translations and automatic translations that were not submitted to post-editing. The metrics employed turned out differing results attributable to the modalities of work, and allowed for the comparison of the effects on the different types of texts translated. Finally, the multiple translations of one of the texts were submitted to the evaluation of the readers, to determine their preferences.

Keywords

Translation; automatic translation; post-editing; CAT Tools; translation evaluations.

Sumário

1. Introdução	13
1.1. Objetivos da pesquisa, perguntas orientadoras e hipóteses	17
1.2. Estrutura da dissertação	18
2. O momento atual da indústria de serviços linguísticos	19
2.1. O espaço da tradução automática e da pós-edição	22
3. Os caminhos da MT e as ferramentas CAT: encontros e desencontros	27
3.1. Os primeiros passos da tradução por máquina	28
3.1.1. Mapeamento dos principais problemas	30
3.2. Avanços na MT: na procura da FAHQT	31
3.3. Avanços em MT e interatividade	35
3.3.1. A MT interativa no século XXI	37
3.4. As ferramentas CAT como um caso de interação bem-sucedido	40
4. Considerações sobre as “viradas” nos estudos da tradução e o papel da tecnologia	46
4.1. As viradas nos Estudos da Tradução e a tecnologia	46
4.2. É possível falar em uma “virada tecnológica”?	51
5. Pesquisas relacionadas	59
5.1. Pesquisas centradas na qualidade e produtividade	59
5.2. Utilização de MT no nível de subsegmento e tradução interativa	61
5.3. Características da TH e da PEMT e preferências dos leitores	63
6. O estudo experimental	69
6.1. Apresentação	69
6.2. Objetivos	70
6.3. Metodologia	71
6.3.1. Os tradutores e tradutoras participantes	71
6.3.2. O ambiente de TM + MT	72
6.3.3. O material enviado para os tradutores e tradutoras participantes	73
6.3.4. Os textos escolhidos	75
6.3.5. Diversidade lexical	77

6.3.6. Densidade lexical	79
6.3.7. Distância de edição	80
6.3.8. Produtividade	81
6.3.9. Preferências dos leitores	81
7. Resultados	83
7.1. Diversidade lexical (TTR)	83
7.2. Densidade lexical	86
7.3. Distância de edição	88
7.4. Produtividade	90
7.5. Preferências dos leitores	92
8. Conclusões	94
9. Referências bibliográficas	99
Apêndice A. Instruções para a gravação da tela do PC	107
Apêndice B. Ficha do projeto T1 para pós-edição de Machine Translation (PEMT)	108
Apêndice C. Ficha do projeto T1 para tradução humana (TH)	109
Apêndice D. Ficha do projeto T2 para pós-edição de Machine Translation (PEMT)	110
Apêndice E. Ficha do projeto T2 para tradução humana (TH)	111
Apêndice F. Texto do e-mail para os avaliadores	112
Apêndice G. Exemplos de simplificação e interferência T1	113
Apêndice H. Exemplos de simplificação e interferência T2	116

Lista de abreviaturas e siglas

MT	<i>Machine Translation</i> (Tradução Automática)
TM	<i>Translation Memory</i> (Memória de Tradução)
TH	Tradução Humana
PEMT	<i>Post-Editing of Machine Translation</i> (Pós-edição de Tradução Automática)
MTNSS	Tradução Automática no Nível de Subsegmento
CAT	<i>Computer Assisted Translation</i> (Tradução Assistida por Computador)
T1	Texto 1
T2	Texto 2
PEMT(1)	Pós-edição do Tradutor 1
PEMT(2)	Pós-edição do Tradutor 2
IA	Inteligência Artificial
HCI	<i>Human-Computer Interaction</i> (Interação Humano-Máquina)
TTR	<i>Type-Token Ratio</i> (Razão entre palavras diferentes/total de palavras)

Lista de figuras

Figura 1 –	Evolução do mercado de tradução entre 2009 e 2021	20
Figura 2 –	Principais serviços oferecidos pelas agências	23
Figura 3 –	Crescimento dos projetos que envolvem MT – 2019-2020	24
Figura 4 –	Triângulo de Vauquois	33
Figura 5 –	Interface do PTM	38
Figura 6 –	Produção realista por hora como PEMT	39
Figura 7 –	Augmented Translator	40
Figura 8 –	Interface do SDL Trados Studio	42
Figura 9 –	Recursos interativos nas ferramentas CAT (SDL Trados Studio 2019)	43
Figura 10 –	Recursos mais utilizados pelos tradutores	44
Figura 11 –	Mapa dos Estudos da Tradução baseado em Holmes-Toury	47
Figura 12 –	Esquema dos Estudos do Tradutor	50
Figura 13 –	Adaptação do mapa de Holmes-Toury	56
Figura 14 –	Possíveis superposições entre a “virada tecnológica” e outras “viradas”, temáticas e metodológicas	57
Figura 15 –	Captura de tela do SDL Trados Studio 2019 - MTNSS	73
Figura 16 –	Captura de tela do SDL Trados Studio 2019 - sdlxliff	74
Figura 17 –	Arquivo de análise do SDL Trados Studio (T1)	76
Figura 18 –	Arquivo de análise do SDL Trados Studio (T2)	76
Figura 19 –	Tela da aba Word List do concordanciador AntConc	78
Figura 20 –	Fragmento de um arquivo Excel	80
Figura 21 –	Diversidade lexical (T1)	84
Figura 22 –	Diversidade lexical (T2)	85
Figura 23 –	Densidade lexical (T1)	86
Figura 24 –	Densidade lexical (T2)	87
Figura 25 –	Distância de edição (T1)	88
Figura 26 –	Distância de edição (T2)	89
Figura 27 –	Produtividade em palavras/hora (T1)	91
Figura 28 –	Produtividade em palavras/hora (T2)	91

Lista de tabelas

Tabela 1 –	Relação das publicações sobre tradução	54
Tabela 2 –	Análise temática dos artigos publicados	55
Tabela 3 –	Quadro comparativo de tipos de erros e nível de gravidade de diferentes modelos de garantia de qualidade	60
Tabela 4 –	Diversidade lexical T1	84
Tabela 5 –	Diversidade lexical T2	85
Tabela 6 –	Densidade lexical T1	87
Tabela 7 –	Densidade lexical T2	87
Tabela 8 –	Distância entre o original e as diferentes modalidades de tradução de T1	89
Tabela 9 –	Distância entre a MT e as diferentes modalidades de tradução de T1	89
Tabela 10 –	Distância entre o original e as diferentes modalidades de tradução de T2	90
Tabela 11 –	Distância entre a MT e as diferentes modalidades de tradução de T2	90
Tabela 12 –	Tempo consumido pela tradução + revisão da T1 em cada modalidade	91
Tabela 13 –	Tempo consumido pela tradução + revisão da T2 em cada modalidade	92
Tabela 14 –	Resultado das avaliações	92

Introdução

Cada vez mais, o trabalho dos tradutores profissionais é mediado pela tecnologia, que deixou de ser um simples recurso de apoio para ocupar um lugar de centralidade crescente. A explosão de conteúdo digital observada nas últimas décadas, o aumento da demanda de serviços linguísticos, a redução dos prazos e a incapacidade da tradução humana tradicional para satisfazer essa demanda fizeram com que ferramentas tecnológicas passassem a ser usadas para aumentar a produtividade. E a penetração foi tão forte que diversos pesquisadores consideram que a atividade tradutória hoje é inseparável das tecnologias que a sustentam e, inclusive, referem-se a ela como uma forma de interação humano-computador (O'BRIEN, 2012).

Nesse panorama, duas ferramentas ganharam destaque na área de tradução especializada: as ferramentas de memória de tradução (*Translation Memory*, TM)¹, mais conhecidas como ferramentas CAT (*Computer Assisted Translation*) e, mais recentemente, os sistemas de tradução automática (*Machine Translation*, MT). Se bem inicialmente essas duas tecnologias tomaram caminhos diferentes, nas últimas duas décadas houve uma reconvergência entre ambas, que passaram a ter um nível de integração cada vez maior. Assim, no que Bundgaard (2017) denomina ambiente de TM assistida por MT, os tradutores recebem sugestões que podem vir da TM, que permite a reciclagem de traduções previamente produzidas, ou de um sistema de MT, para o caso de não existir nenhum conteúdo aproveitável na TM.

Considerando esse contexto, a presente dissertação pretende analisar alguns aspectos da interação entre os tradutores profissionais e essas duas tecnologias no ambiente de TM + MT. Especificamente, a pesquisa pretende comparar duas

¹ Optei por manter em inglês as siglas e acrônimos de uso consagrado na área de tradução. O resto das siglas e acrônimos foram traduzidos.

maneiras de utilizar os resultados da MT em um ambiente de ferramenta CAT: a pós-edição tradicional (*Post-edition of Machine Translation, PEMT*), onde o tradutor recebe um arquivo previamente traduzido por MT e o corrige e/ou edita, e a utilização de fragmentos provenientes da MT que o tradutor vai inserindo na medida em que traduz, quando e se achar conveniente, no nível de subsegmento.²

O interesse pela comparação entre essas duas maneiras de utilizar a MT + TM surgiu da minha própria experiência como tradutor especializado, profissão na qual atuo há mais de 15 anos. A fim de atender às demandas de agências de tradução brasileiras e estrangeiras, comecei a executar trabalho de PEMT dentro de ferramentas de TM há aproximadamente três anos, utilizando sistemas de MT diversos, desde aqueles baseados em modelos estatísticos (*Statistical Machine Translation, SMT*) até tradução automática neural (*Neural Machine Translation, NMT*), e desde sistemas de código aberto adaptados até sistemas desenvolvidos especificamente para empresas. Também comecei a utilizar MT + TM para trabalhos *free-lance* que se enquadravam em uma série de características, com o objetivo de aumentar a minha produtividade. Nesses casos, testei diversos sistemas de tradução automática públicos – como Google Translate, DeepL e, mais recentemente, ModernMT – que através de uma interface de programação de aplicações (*Application Program Interface, API*), isto é, um conjunto de normas que permite a comunicação entre esses sistemas, podem ser incorporados dentro do ambiente das ferramentas CAT.

Assim, movido pela curiosidade, comecei a testar o uso de MT no nível de subsegmento para reduzir o trabalho de pós-edição intensivo, por meio de uma intervenção inicial que me permitisse organizar a frase, em lugar de me limitar a corrigir a organização da frase proporcionada pela MT. Além disso, a minha percepção era que as soluções propostas pela MT em certas ocasiões “engessavam” a minha tradução de modo tal que era difícil de reverter. Observei que isso acontecia, principalmente, com textos que exigiam traduções mais criativas, como os textos de marketing e postagens de blogs de tecnologia que fazem parte do meu fluxo de trabalho cotidiano. Assim, impulsionado pela potencialidade de aumento da produtividade que oferece a MT, mas tentando

² Segmento é a unidade mínima em que as CAT Tools dividem um texto e na qual está baseada a sua lógica. De maneira predeterminada, qualquer texto processado por meio de uma CAT Tool é segmentado (dividido) considerando os seguintes separadores: ponto final, dois pontos, signo de exclamação/interrogação (fechamento) ou quebra de parágrafo.

manter o meu estilo de escrita, passei a utilizar cada vez mais as propostas da MT em forma de fragmentos, em certa forma me apropriando da MT de uma maneira diferente do habitual. Segundo Zetzsche (2020), embora muitos tradutores tenham adotado o uso de MT dessa maneira e essa modalidade de trabalho tenha demonstrado ser especialmente útil com MT estatística (SMT), considerando-se que os fragmentos oferecidos habitualmente foram traduzidos por outros tradutores profissionais, no momento não existem muitas pesquisas sobre esse processo utilizando MT neural (NMT), como o estudo realizado na presente dissertação se propõe.

A presente dissertação de mestrado pretende, então, comparar esses dois métodos de utilizar a MT em um ambiente de TM + MT. O primeiro é a PEMT tradicional, que consiste em editar, alterar ou corrigir um texto pré-traduzido por um sistema de MT. Nesse caso, a intervenção do tradutor acontece uma vez que o texto completo tenha sido traduzido. O segundo método consiste na seleção e inserção, pelo tradutor, de fragmentos de diversas extensões fornecidos por um sistema de MT dentro de um segmento. Nesse caso, podemos dizer que o uso da MT realiza-se no nível de subsegmento (MTNSS).

Para comparar ambos os métodos, foi feito um estudo experimental envolvendo quatro tradutores profissionais especializados que traduziram textos da área de tecnologia em um ambiente de TM + MT. Para o estudo foram selecionados dois textos, um deles com conteúdo totalmente técnico (características de um sistema de monitoramento de sistemas informáticos) e outro que aborda um assunto também técnico, mas com uma linguagem que poderíamos considerar mais jornalística (uma postagem de blog sobre a reforma de um estádio de futebol). Os textos foram traduzidos/pós-editados do inglês para o espanhol por tradutores nativos na língua de destino.

Cada texto foi traduzido usando quatro métodos:

- a) Tradução por máquina sem pós-edição (MT);
- b) Pós-edição tradicional (PEMT) no nível de segmento;
- c) Utilização de MT no nível de subsegmento (MTNSS) durante a tradução;
- d) Tradução humana (TH).

A relevância do assunto vai muito além do meu interesse pessoal. Primeiramente, a pesquisa pretende fazer mais um aporte à construção de uma ponte entre os tradutores profissionais especializados (não literários) e a pesquisa

acadêmica, um processo que, felizmente, começou a se afiançar há alguns anos, mas que precisa ser reforçado e, devido à rápida mudança da tecnologia e da indústria de serviços linguísticos, precisa ser atualizado de maneira contínua. Em tal sentido, a realização de pesquisas que partam da perspectiva dos tradutores e que sejam capazes de fornecer insumos para pensar no seu trabalho enriquecem tanto o âmbito acadêmico como o profissional.

Em segundo lugar, uma série de fatores tem feito com que a PEMT seja uma modalidade de trabalho bastante utilizada pelos tradutores profissionais especializados. Mesmo que ainda exista uma certa rejeição da MT por parte de muitos profissionais por diversos motivos (principalmente por considerá-la como uma ameaça à sua fonte de renda e ao futuro da profissão), o número de tradutores que passaram a incluir a MT dentro dos seus fluxos de trabalho, seja por imposição das agências e clientes ou por sua própria vontade, vem aumentando. Uma pesquisa do projeto SCATE (*Smart Computer-Aided Translation Environment*) de 2015 concluiu que 27% dos tradutores utilizavam MT dentro das suas ferramentas CAT³, e podemos arriscar a afirmar que esse percentual deve ter aumentado significativamente nos últimos anos, considerando o crescimento do número de sistemas de MT para os quais as ferramentas CAT oferecem compatibilidade.

Em terceiro lugar, podemos mencionar a falta de pesquisas que analisam a interação com NMT durante a tradução no nível de subsegmento, como a proposta nesta pesquisa.

Finalmente, não podemos deixar de mencionar que a pesquisa foi feita pensando na tradução especializada e por um tradutor especializado, categoria que responde por uma parte muito significativa do volume total de conteúdo traduzido na atualidade. De fato, um estudo feito pela consultora Nimdzi em 2021 sobre os maiores fornecedores de serviços linguísticos determinou que as verticais de Tecnologia, TI e software, Ciências da Vida e Financeira e Legal são os segmentos prevalentes em termos de participação de mercado (HICKEY; AGULLÓ GARCÍA, 2021).

³ Disponível em: <https://www.arts.kuleuven.be/ling/ccl/projects/scate/deliverables>.

1.1

Objetivos da pesquisa, perguntas orientadoras e hipóteses

Considerando o que foi até aqui exposto, foram estabelecidos os seguintes objetivos para a pesquisa:

- a) avaliar comparativamente os resultados obtidos utilizando duas modalidades de trabalho nas quais se integra a tecnologia de TM com MT, 1) a pós-edição tradicional (PEMT); e 2) o uso de MT no nível de subsegmento (MTNSS). Observar se as duas maneiras de trabalhar causam impacto de formas diferentes em gêneros textuais diversos.
- b) comparar os resultados obtidos por meio de PEMT e MTNSS com a tradução humana (TH) por meio de uma série de análises computacionais.
- c) avaliar as preferências de um grupo de leitores com relação às traduções obtidas por meio das modalidades de trabalho propostas.

Para atingir esses objetivos, foram formuladas as seguintes perguntas orientadoras da pesquisa:

- a) Em que medida o trabalho em um ambiente de TM + MT oferece resultados diferentes usando a MT no nível de segmento e no nível de subsegmento?
- b) Em que consistem essas diferenças, caso existam?
- c) O uso de MT no nível de subsegmento é mais adequado para certos gêneros de texto do que para outros? A hipótese foi que o uso de MTNSS seria mais adequado para textos de tipo jornalístico e que haveria menos diferenças em textos técnicos com maior nível de especialização.
- d) As duas maneiras de trabalhar com a MT, PEMT e MTNSS, permitem observar diferenças em termos de produtividade (palavras/hora)?
- e) Os leitores avaliadores têm preferência pelos textos traduzidos com alguma das modalidades mencionadas (PEMT, MTNSS e TH)?

1.2

Estrutura da dissertação

A presente dissertação está organizada em 8 capítulos, sendo o primeiro este capítulo introdutório. O capítulo 2 apresenta um breve panorama da indústria de serviços linguísticos na atualidade, onde se presta especial atenção ao advento da tradução automática e o espaço que a pós-edição ocupa no mercado. O capítulo 3 explora a evolução da MT a partir de duas perspectivas, a da inteligência artificial (IA) e a da interface humano-computador (HCI), que partiram de premissas diferentes e cujas ideias e resultados estão presentes na atualidade. O capítulo 4 faz algumas considerações sobre as “viradas” da tradução e o momento atual da pesquisa relacionada com a tradução profissional e a tecnologia. O capítulo 5 apresenta uma resenha de diversas pesquisas relacionadas com o estudo experimental. O capítulo 6 descreve o estudo experimental realizado. O capítulo 7 apresenta os resultados e, finalmente, o capítulo 8 discute as conclusões e as contribuições do presente trabalho para os tradutores profissionais e os pesquisadores da área.

O momento atual da indústria de serviços linguísticos

Neste capítulo abordamos brevemente algumas mudanças ocorridas nas últimas décadas no setor de serviços linguísticos. O objetivo não é oferecer uma descrição detalhada, mas explorar algumas questões específicas, com especial ênfase na incorporação da tecnologia e, em particular, da MT e a PEMT. Veremos como a tecnologia de MT está desafiando algumas ideias bem estabelecidas no mercado de serviços linguísticos e que até pouco tempo atrás ninguém questionava, como possibilidade de gerar traduções com qualidades diferentes, dependendo da finalidade à qual estejam destinadas.

Como afirma Estill (2017, p. 9), a tradução como atividade “trocou suas práticas artesanais por procedimentos de alta tecnologia, até se tornar uma indústria bilionária no intervalo de algumas décadas”. De fato, o setor de serviços linguísticos, no qual a tradução encontra-se inserida, tem atravessado fortes mudanças nos últimos anos que o levaram a crescer de forma acelerada, duplicando seu tamanho em dez anos até atingir um volume de negócios de 55 bilhões de dólares em 2020, e com uma estimativa de continuar crescendo até alcançar entre 56 e 58 bilhões de dólares em 2021 (ver Figura 1) (HICKEY; AGULLÓ GARCÍA, 2021).

As mudanças no setor de tradução estão inseridas em um movimento maior acontecido no mundo do trabalho. Negri e Hardt (2000) denominaram essa mudança como o processo de pós-modernização ou informatização, que gerou uma migração da indústria para os serviços, alterando a qualidade e a natureza do trabalho. Dentro dessa nova conformação, a tradução pode ser classificada como trabalho imaterial e sofreu mutações similares a outras indústrias, como a descentralização e a dispersão global dos processos e lugares de produção.

As mudanças acontecidas nas últimas décadas na indústria de serviços linguísticos foram impulsionadas por uma série de fatores, como a explosão do

volume de informação gerada, a internacionalização de empresas em diversos setores, a digitalização e o colapso do ciclo de vida dos produtos, entre outros.

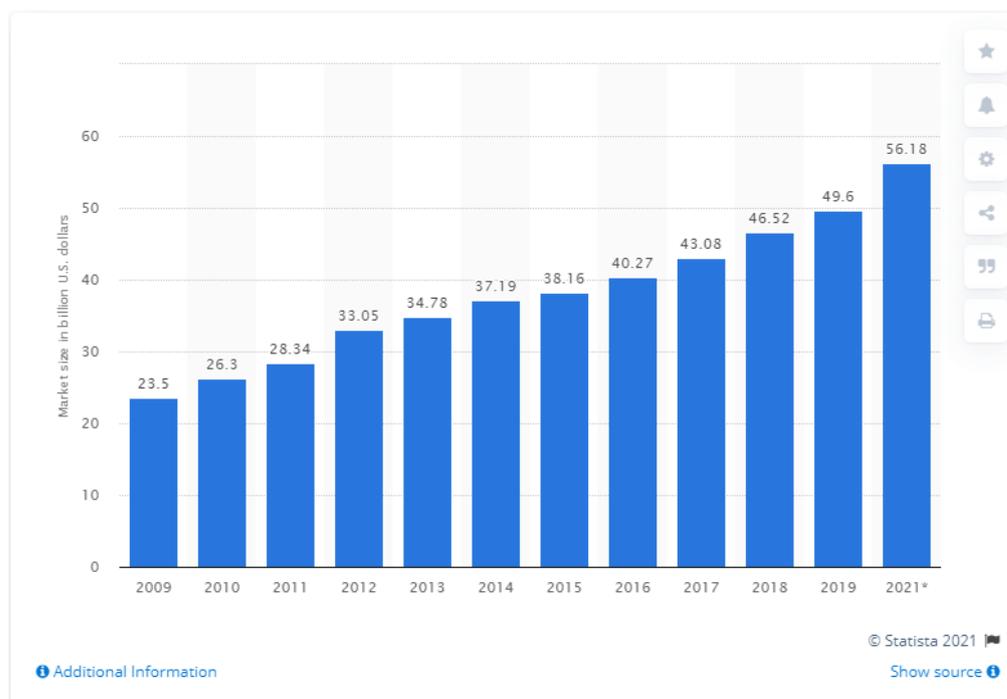


Figura 1 – Evolução do mercado de tradução entre 2009 e 2021.

Fonte: STATISTA RESEARCH DEPARTMENT (2021).

Como resposta a essas mudanças, a tecnologia vem ocupando um papel cada vez mais central e pervasivo na indústria. Em tal sentido, ela foi e continua sendo vista como a maneira de superar os principais obstáculos que os fornecedores de serviços de tradução enfrentam em decorrência do crescimento acelerado do setor: reduzir os prazos e os custos por meio do aumento da produtividade dos tradutores. A propagação do uso das ferramentas CAT a partir da década de 1990 e da MT a partir dos anos 2000, bem como a automatização de diversos processos, como a gestão de terminologia e de projetos, entre outros, podem ser enquadradas como tentativas de superar os obstáculos antes citados. Apesar da introdução das ferramentas CAT – hoje ferramentas quase obrigatórias para os tradutores especializados – ter significado um aumento significativo da produtividade, isso não foi suficiente, e o problema gerado pela explosão da produção de conteúdo e a redução dos prazos exigiu outras estratégias. Foi justamente em resposta a isso que se introduziu a MT e o processo de PEMT a ela associado, como veremos na continuação.

Embora o tema da qualidade na tradução tenha sido e continue sendo um tema controverso, nos últimos tempos a qualidade das traduções passou a ter exigências diferentes dependendo da finalidade para a qual será utilizada. Sem ser um fenômeno totalmente novo, o fato é que começou a ser assumido tanto pelos clientes como pelos prestadores de serviços linguísticos. Desse modo, empresas e agências começaram a oferecer serviços diferenciados em termos de qualidade, com modalidades de trabalho também diferenciadas (tradução 100% humana, PEMT light e full) e uma oferta de qualidade, preços e prazos também diferentes atrelados a cada uma delas. Além disso, alguns clientes passaram a utilizar MT sem nenhuma intervenção humana para diversas finalidades. Assim, uma mesma empresa, como a Microsoft, por exemplo, pode usar tradução por máquina “crua”, isto é, sem pós-edição, para seu conteúdo de suporte técnico, e tradução 100% humana para seu material de marketing. Outras grandes empresas, como o Volkswagen Group (PORSIEL, 2008), entre outras, têm sistemas de tradução automática próprios ou customizados (alimentados e mantidos de acordo com a terminologia e o tipo de comunicações da empresa) que são utilizados para diferentes finalidades, que vão de e-mails, minutas e comunicações internas, até relatórios com diversos graus de complexidade.

Como afirma Cronin:

A tradução como uma prática artesanal de alto custo está confinada aos folhetos de estratégia corporativa, vendas e marketing, e à informação de serviço e suporte de alto nível. Na medida em que as operações de suporte dos produtos acumulam informação e conteúdo, o volume de material a traduzir pode se tornar muito dispendioso, motivo pelo qual surgiu o que a SDL chama de “fábrica de idioma local”, com suas “economias de escala” correspondentes. (CRONIN, 2013, p. 93).

De maneira similar, Bédard (*apud* BOWKER, 2020) classifica as traduções demandadas pelo mercado atual em três tipos: a) traduções de prestígio de alta qualidade, limitadas a um pequeno número de documentos, como publicações prestigiosas; b) traduções de qualidade artesanal; e c) traduções de qualidade industrial, a mais comum e aplicada a grandes volumes de texto. Em tal sentido, Koby *et al.* (2014, p. 416) oferecem uma definição ampla de qualidade de tradução:

a qualidade de uma tradução demonstra a precisão e fluência requerida pelo público e a finalidade e cumpre com todas as outras especificações negociadas entre o solicitante e o fornecedor, considerando as necessidades do usuário final.

Essas necessidades, agora assumidamente diferentes, converteram o mercado em um continuum que tem lugar tanto para MT “crua” como para tradução humana de alta qualidade, e diversas opções entre elas dependendo da finalidade à qual estão destinadas (BOWKER, 2020).

2.1

O espaço da tradução automática e da pós-edição

É justamente nesse espaço de procura de redução dos tempos e custos, e em consonância com a classificação das traduções em categorias que demandariam qualidades diferentes, onde entram a tradução por máquina (MT) e a pós-edição (PEMT). Elas vêm sendo crescentemente implementadas pelas agências de tradução para diversos cenários e encontram-se entre os serviços mais oferecidos pelos fornecedores de serviços de tradução (71,5%), apenas atrás de tradução e localização, e apresentando um crescimento de 22% em dois anos, segundo uma pesquisa recente da consultora Nimdzi (Figura 2).

Apesar dos sistemas de MT terem uma trajetória que se inicia na segunda metade do século XX, só começaram a ganhar importância no mercado nos últimos anos, devido a uma melhora na sua qualidade e um aumento da sua disponibilidade (KOPONEN, 2016).

No caso da MT “crua”, ela se aplica a diversos cenários, como volumes gigantescos ou com prazos muito curtos, onde a tradução humana seria impossível ou inviável do ponto de vista econômico ou operacional, ou conteúdo que não precisa de uma tradução de alta qualidade, mas simplesmente compreensível (*good enough*).

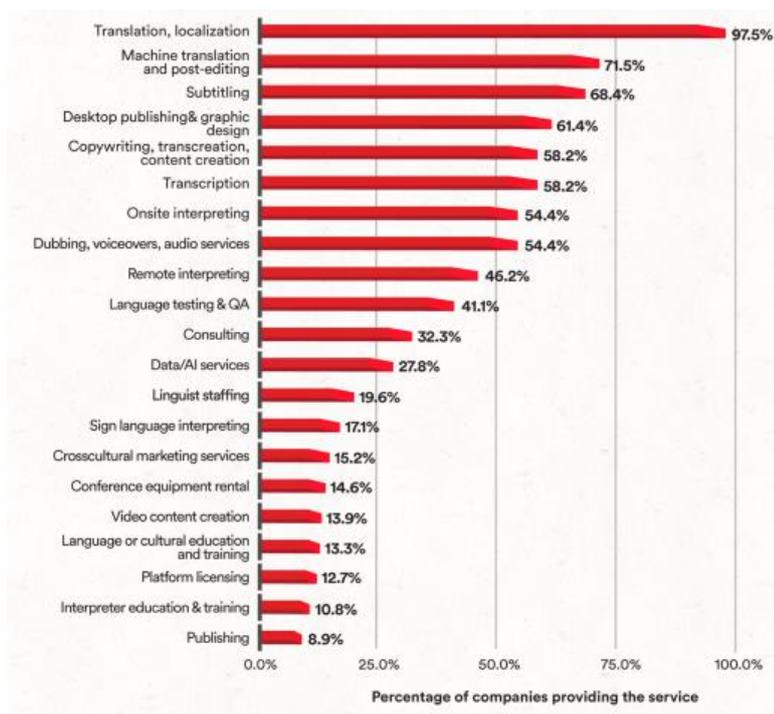


Figura 2 – Principais serviços oferecidos pelas agências.

Fonte: HICKEY; AGULLÓ GARCÍA (2021).

Um exemplo do primeiro caso seria o do eBay, um marketplace global que, em um momento determinado no tempo, pode ter até 800 milhões de anúncios de produtos ativos, cada um deles com aproximadamente 300 palavras. Como explica Rowda (2016), a tradução humana seria inviável uma vez que seriam necessários 1.000 tradutores trabalhando durante 5 anos só para traduzir os 60 milhões de anúncios para serem publicados em um único país. E isso sem se esquecer de que um anúncio pode ter uma vida útil de até uma semana, ou incluso de um dia só. Então, nesse caso, bem como em muitos outros, uma tradução com uma precisão razoável e entregue dentro do prazo é mais valiosa que uma gramática e fluência perfeitas, com suas implicações de tempo e custo (VASHEE, 2021). Outras vezes, o conteúdo traduzido não está destinado a um público externo ou não será lido por nenhuma pessoa, mas alimentará um sistema. Koby *et al.* (2014) oferecem um exemplo disso no caso do uso de MT para tentar estabelecer coincidências entre as buscas dos usuários e os elementos existentes em uma base de dados. No caso de que a MT não consiga identificar uma coincidência, pode ser necessária a intervenção de um tradutor humano (denominado “operador”). A função do tradutor seria corrigir o sistema ou então simplesmente retraduzir para obter essa coincidência nessa busca e em buscas

futuras: “Como essa tradução não será nunca usada com um cliente externo, a aparência do texto não tem importância nenhuma. Portanto, o operador NÃO deve realizar nenhum tipo de limpeza nem aprimoramento” (KOBY et al., 2014, p. 419). Enquanto muitos tradutores se perguntam se alguma vez serão substituídos pela tradução automática, considerando isso como um evento catastrófico que aconteceria em algum momento no futuro, a realidade é que bilhões de palavras já são traduzidas todos os dias por máquinas para diversas finalidades e sem nenhum tipo de pós-edição.

Um segundo caso de uso de MT é com intervenção de um tradutor a posteriori, isto é, com pós-edição (PEMT). Considerando que a MT não consegue oferecer uma qualidade apropriada para muitos pares de línguas, tipos de textos e áreas da tradução, a PEMT foi adotada pela indústria basicamente para reduzir custos e aumentar a produtividade dos tradutores, conforme ficou demonstrado em diversas pesquisas (ARENAS, 2010; FEDERICO *et al.*, 2012; ZECHEV, 2014), embora seja importante assinalar que as pesquisas sobre PEMT e produtividade apresentam resultados parcialmente conflitantes (KOPONEN, 2016).

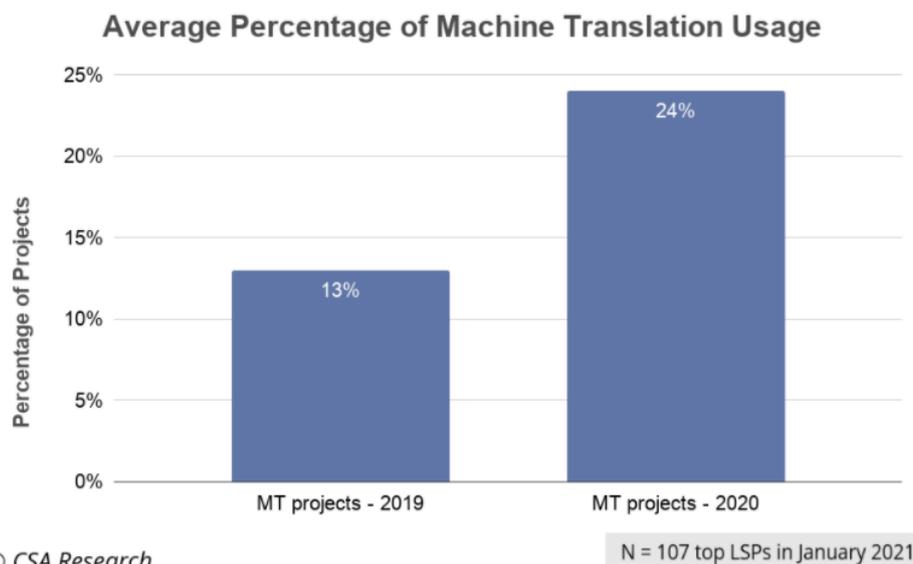


Figura 3 – Crescimento dos projetos que envolvem MT – 2019-2020.

Entre os 107 maiores fornecedores de serviços linguísticos.

Fonte: CSA RESEARCH (2021).

Segundo a norma ISO 18587:2017, a pós-edição consiste em editar e corrigir os resultados fornecidos por um sistema de MT. A norma diferencia dois tipos de pós-edição: *full post-editing* e *light post-editing*. A *full post-editing* é

descrita pela norma como o processo de pós-editar um texto pré-traduzido por um sistema de MT para obter um produto comparável à tradução humana. A *light post-editing* é definida como o processo de pós-editar um texto pré-traduzido por um sistema de MT para obter um texto meramente compreensível, mas que não tem a pretensão de ser comparável à tradução humana. Em ambos os casos, a norma indica que as agências de tradução devem garantir que os pós-editores atinjam os seguintes objetivos ao longo do processo: a) que o texto pós-editado seja compreensível; b) que exista uma correspondência entre o conteúdo do idioma fonte e o conteúdo do idioma objetivo; c) que se cumpram os requisitos e especificações de pós-edição definidos pelo fornecedor de serviços de tradução. A norma também estabelece os requisitos para os profissionais envolvidos na tarefa, que são idênticos a aqueles dos profissionais envolvidos na tradução (POPIOLEK, 2021).

Por sua vez, a *Automated Translation Society* (TAUS), um *think tank* que agrupa organizações importantes do universo da tradução automática, como fornecedores e compradores de serviços de tradução automática, mas que também inclui universidades e uns poucos governos e ONGs, define a pós-edição como o processo de melhorar traduções geradas por MT com o mínimo trabalho manual possível.⁴ A TAUS também diferencia duas qualidades de PEMT: qualidade de tradução humana e “boa o suficiente” (*good enough*), que se corresponderiam a grandes traços com as definições de *full* e *light post-editing* da norma ISO. O requisito da primeira é que seja comparável à tradução humana, e o da segunda é que a tradução seja compreensível e precisa, mas sem necessidade de ser estilisticamente atraente. Cabe destacar que diversas organizações e fornecedores de serviços linguísticos tem concepções diferentes sobre o que implica a *full* e a *light post-editing*, especialmente com relação à última modalidade.⁵

Assim, a pós-edição, tanto *full* como *light*, são tarefas cada vez mais demandadas pelo mercado. E é importante destacar que, embora exista a possibilidade de pós-editar um texto monolíngue em qualquer ferramenta de processamento de texto, como o Word, e que isso poderia ser feito por outros profissionais, na atualidade a pós-edição é feita majoritariamente por tradutores

⁴TAUS (2010). Post-editing in Practice. A TAUS Report. Disponível em: <http://www.translationautomation.com/articles/>. Acesso em: 5 mar. 2013.

⁵Para ver um estudo comparativo de pautas de pós-edição, cf. Hu e Cadwell, 2016.

profissionais que trabalham com textos bilíngues, em ambientes de ferramenta CAT.

Em resumo, o crescimento do uso de MT e PEMT responde a uma necessidade de negócios. Como afirma Vashee:

O valor de negócio da tradução muitas vezes é definido por sua utilidade para o consumidor em uma viagem digital: sua compreensibilidade básica, uma precisão utilizável e uma entrega oportuna são mais importantes que uma gramática e uma fluência perfeitas. A frase ‘boa o suficiente’ é usada ao mesmo tempo em termos depreciativos e como um atributo positivo para a tradução que é útil para um cliente incluso quando for menos que ‘perfeita’. (VASHEE, 2021)⁶.

Considerando o exposto, tudo parece indicar que a MT, a PEMT e outros serviços combinando tradução humana e inteligência artificial, já existentes ou que serão desenvolvidos no futuro, continuarão ocupando um lugar de importância crescente no setor.

⁶ Todas as citações incluídas na presente dissertação foram traduzidas quando o original não era em português.

3

Os caminhos da MT e as ferramentas CAT: encontros e desencontros

No presente capítulo será feito um breve resumo da evolução da MT e o desenvolvimento das CAT Tools para tentar entender o momento atual de convergência de ambas as tecnologias.

Green *et al.* (2015) afirmam que, a fim de analisar a evolução da MT e das CAT Tools, é importante olhar para duas disciplinas: a interação humano-máquina (*Human-Computer Interaction*, HCI) e a inteligência artificial (IA). Eles destacam o fato de que hoje esses dois campos são mais mutuamente influentes que nunca, com diversos sistemas amplamente utilizados, como o Google Tradutor, que utilizam recursos de IA ocultos atrás de interfaces simples e intuitivas desenvolvidas segundo os princípios da HCI. Porém, os autores destacam ambos os campos surgiram em momentos diferentes da história e tiveram objetivos e abordagens divergentes com relação à MT.

A IA, a partir da qual foi possível o desenvolvimento da MT, teve suas origens a inícios da década de 1950 a partir de pesquisas feitas fundamentalmente por matemáticos que procuravam desenvolver algo capaz de rivalizar ou inclusive superar a inteligência humana. Um dos seus pioneiros, o professor John McCarthy, da Universidade de Standford, a definiu, em 1955, como “a ciência e engenharia de fazer máquinas inteligentes”. Por sua vez, Martin Minsky a definiu como “a ciência de fazer com que máquinas façam coisas que exigiriam inteligência se fossem feitas por pessoas”.

A diferença da IA, na década de 1950 o campo da HCI se debruçava sobre problemas mais empíricos relacionados com a usabilidade e os fatores humanos, com o objetivo de fazer com que as máquinas fossem mais úteis para os seres humanos. Na área de computação, mais especificamente, a ideia era facilitar as interações entre os programadores e os computadores. Uma das pioneiras, Grace

Hopper, descreveu sua meta como sendo liberar os matemáticos de fazerem matemática, o que poderia ser traduzido nos nossos dias como a meta de usabilidade de liberar os usuários de fazer seu trabalho (GRUDIN, 2012, p. 31).

3.1

Os primeiros passos da tradução por máquina

Sem dúvida, a massificação do uso da MT na atualidade, principalmente mediante o uso de sistemas gratuitos não destinados ao uso profissional, como o Google Translate (GT), é impressionante. Basta dizer que em maio de 2016, o GT traduzia 143 bilhões de palavras por dia, oferecendo 100 combinações de idiomas e com um ritmo de crescimento que duplicou o volume em 4 anos (WAY, 2020). A MT também vem sendo cada vez mais utilizada pelos tradutores profissionais nos últimos anos, o que pode ser identificado como uma tendência com um crescimento sólido.

A MT é mais antiga que a IA e que a linguística computacional, campos com os quais hoje está intimamente relacionada. Ela foi a primeira aplicação não numérica dos computadores e durante mais de uma década, desde inícios da Guerra Fria, no pós-guerra, até a metade da década de 1960, foi uma área de pesquisa que recebeu uma atenção e um fluxo de fundos significativos e que, inclusive, gerou muitas expectativas e atraiu a atenção do público. Seu desenvolvimento também deu o impulso inicial a um novo campo: o processamento de linguagem natural (PLN), que concentrou sua atenção na MT durante duas décadas (1940-1960), despertando grande entusiasmo e otimismo nos pesquisadores, que abordavam uma tarefa muito complexa, ao tempo em que utilizavam uma nova ferramenta, o computador. Em qualquer caso, é importante lembrar que os recursos informáticos da época eram bastante precários, com muitas restrições para acessar as máquinas e baixas capacidades de armazenamento e processamento (SPARCK JONES, 2001).⁷

⁷ Citando Plath, Sparck Jones lembra que as velocidades de processamento eram muito limitadas. Por exemplo, o processamento de uma frase longa podia demorar uns 7 minutos.

Em 1949, depois de conversações com Andrew Booth e uma troca epistolar sobre o assunto com Norbert Wiener, especialista em cibernética, um memorando elaborado por Warren Weaver propôs a possibilidade de utilizar computadores para traduzir, inspirado no uso da computação para decodificar mensagens cifradas durante a Segunda Guerra Mundial, tarefa em que se destacou o matemático e cientista da computação Alan Turing. Embora o memorando só se concentrasse em questões estratégicas gerais e objetivos de longo prazo, gerou interesse, inclusive por parte da imprensa (HUTCHINS, 1986).

Um dos primeiros grupos de pesquisa criados para desenvolver a MT nasceu no Laboratório de Pesquisas em Eletrônica do Massachusetts Institute of Technology (MIT), onde o filósofo israelense Yehoshua Bar-Hillel foi nomeado assistente de pesquisa e organizou a primeira conferência da área, em 1952. A partir da conferência foi fundado um novo grupo de pesquisa da Universidade de Georgetown, liderado por Leon Dostert, que começou a trabalhar em um experimento piloto para demonstrar a viabilidade da tradução de máquina. O grupo estabeleceu uma parceria com a IBM, cujo pessoal, dirigido por Peter Sheridan, implementou os procedimentos necessários para utilizar o computador IBM 701. Em 7 de janeiro de 1954 foi realizada a primeira demonstração do sistema no escritório da IBM em Nova York. Apesar das limitações do experimento (foi utilizado um vocabulário de apenas 250 palavras em russo, umas poucas regras e frases em russo cuidadosamente selecionadas), ele foi fundamental para demonstrar a viabilidade do objetivo e ajudou a obter financiamento das agências governamentais americanas (HUTCHINS, 1995).

A partir desse momento, outros grupos de pesquisa foram criados em Cambridge, na Universidade de Harvard e no Instituto de Mecânica de Precisão e Tecnologia de Computação em Moscou, na então União Soviética. Além disso, foi criada a primeira publicação dedicada inteiramente à MT, a *Mechanical Translation*, financiada pela Fundação Rockefeller.

3.1.1

Mapeamento dos principais problemas

Na primeira conferência sobre MT acima mencionada, Bar-Hillel apresentou o documento *The present state of research on mechanical translation*, um trabalho onde fez uma avaliação do estado da arte da MT e os principais problemas que enfrentava. É interessante constatar que muitos dos pontos citados por Bar-Hillel ainda continuam vigentes.

Em primeiro lugar, ele antecipou dois usos para a MT. O primeiro, que seria chamado de *disseminação*, responderia a “urgência de ter publicações em idioma estrangeiro, principalmente nos campos da ciência, as finanças e a diplomacia, traduzidas com alta precisão e um velocidade razoável”. O segundo, chamado de *assimilação*, seria uma resposta à necessidade de obter informação traduzida de grandes volumes de material impresso em jornais, folhetos etc. (de inimigos reais ou potenciais), que exigiria alta velocidade, mas que isso talvez redundaria em um grau de precisão menor (GREEN *et al.*, 2015).

Perante a constatação de que a tradução totalmente automática de alta qualidade (*Fully Automatic High Quality Translation*, FAHQT) seria inatingível em curto prazo, ou então atingível só aceitando o custo da imprecisão, Bar-Hillel advogou pela chamada “MT mista”, isto é, um processo de tradução em que um ser humano interviria em um momento anterior ao processamento pelo computador (pré-edição) para preparar o texto, ou em um momento posterior ao processamento pelo computador (pós-edição), para corrigir os eventuais erros e imprecisões da MT. Mais uma vez, o filósofo tocou em um assunto que hoje continua sendo estudado no âmbito da tradução profissional e que tem gerado muitas pesquisas na área dos Estudos da Tradução orientados à tecnologia. No primeiro caso, o da pré-edição, a resposta mais atual tem sido o uso de linguagem controlada na redação do conteúdo a ser traduzido, isto é, a aplicação de limitações no léxico, gramaticais e de estilo para melhorar o resultado da MT e/ou reduzir o esforço de pós-edição. No segundo caso, Bar-Hillel falava diretamente da pós-edição como se realiza na atualidade, segundo se explorou no capítulo anterior e se utilizou no experimento da presente dissertação.

A visão do Bar-Hillel se opunha ao otimismo generalizado do campo de IA que sustentava a ideia de que as máquinas seriam capazes de fazer o trabalho tradutório, e qualquer outro de maneira similar à humana, em pouco tempo. Por sua vez, os pesquisadores da HCI que atuavam na área da tradução estavam mais interessados na interação humano-máquina e debruçaram-se ao longo dos anos seguintes sobre assuntos mais práticos, como a necessidade de reduzir a importante demanda de tempo que exigiam as tarefas de pós-edição, que em alguns casos era comparável ao tempo que demandaria a tradução do zero, melhorando a interface para facilitar o trabalho do pós-editor.

O ceticismo inicial de Bar-Hillel com relação a alcançar a FAHQT em um curto prazo foi corroborado pelos resultados do estudo plasmado no relatório do ALPAC (Comité Assessor de Processamento de Linguagem Automático), de 1966, que levou a uma redução substancial dos fundos investidos pelo governo dos EUA uma vez que determinou que a MT era mais lenta, menos precisa e mais cara que a tradução humana. Segundo Melby, o relatório foi um “anúncio funerário para o financiamento significativo da tradução por máquina” (MELBY, *apud* BUNDGAARD, 2017, p. 10).

Apesar da decepção inicial, é importante destacar alguns resultados alcançados nessa etapa de pesquisa incipiente. Como destaca Sparck Jones (2001), os primeiros pesquisadores estabeleceram bases sólidas ao identificar e tentar responder às necessidades do processamento de linguagem computacional, além de ter pesquisado diversos assuntos da linguagem, como a polissemia, o processamento e a geração. Também foram capazes de identificar arquiteturas de sistemas e estratégias de processamento e trouxeram ideias que depois se mostrariam influentes, como o uso de representações lógicas.

3.2

Avanços na MT: na procura da FAHQT

Apesar dos efeitos do relatório da ALPAC, as pesquisas de MT recuperaram impulso e financiamento a partir de finais da década de 1960 e inícios da década de 1970 e atravessaram diversas etapas, com melhoras significativas.

No referente às estratégias, foram definidas três abordagens que já tinham sido delineadas no primeiro momento e que se tornariam clássicas: a direta, a transferência e a interlíngua. A abordagem direta é baseada principalmente em um léxico (um dicionário que define, não o conteúdo, mas a variabilidade lexical das palavras) e consiste na realização de uma substituição do tipo palavra por palavra, com certos ajustes, usando um grande dicionário bilíngue. Na segunda abordagem, a transferência, o texto fonte converte-se em uma representação a partir de uma estrutura sintática intermediária específica da língua fonte, que então é transformada por regras específicas do par linguístico em uma representação equivalente na língua de destino, para, finalmente, gerar o texto na língua de destino. A terceira, a interlíngua (ver Figura 4), considerada por Hutchins (2015), a mais interessante e ambiciosa daquele momento, se baseia na ideia de que a MT deve ir além da informação puramente linguística (sintaxe e semântica) e envolver a compreensão do conteúdo do texto. A partir dos progressos no desenvolvimento da linguística, e com uma influência marcante das pesquisas de Chomsky, esses sistemas consideravam que, apesar das diferenças significativas entre as estruturas sintáticas da superfície das línguas, elas compartilhavam uma estrutura profunda com representações semânticas “universais”. Então, a abordagem interlíngua consiste em analisar o texto na língua de origem para gerar uma representação de significado abstrata e universal independente (a interlíngua) e, a partir dessa representação, gerar o significado usando unidades lexicais e as construções sintáticas da língua de destino.

Até finais dos anos 1980, a MT baseou-se essencialmente em regras linguísticas de vários tipos: regras para análise sintática, regras lexicais, regras para geração sintática, regras para morfologia etc. (HUTCHINS, 1995). Essa abordagem foi conhecida como tradução por máquina baseada em regras (*Rule Based Machine Translation*, RBMT).

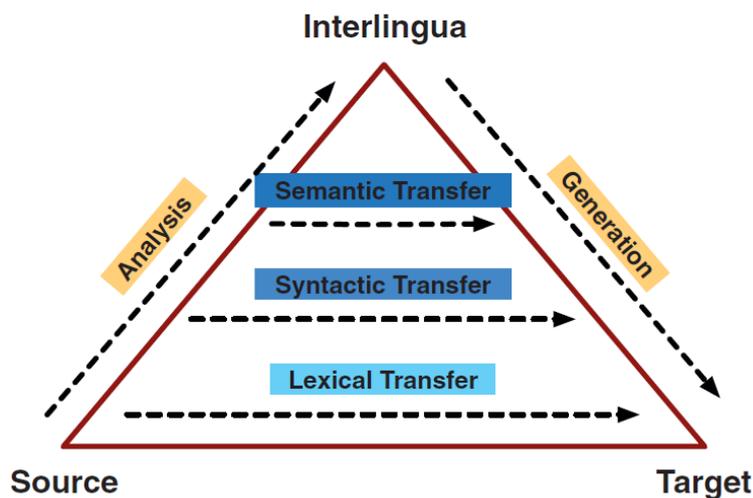


Figura 4 – O triângulo de Vauquois.

Representa o processo de compreender o significado de uma frase na língua fonte, converter isso em uma representação semântica independente da língua e gerar uma frase na língua de destino.

Fonte: KOEHN (2020).

Essa tendência foi quebrada a partir de 1989, quando começou a dominar a abordagem dos métodos denominados de maneira geral como baseados em corpus, que se apoiavam em estatísticas de correspondências de frases, grupos de palavras e palavras individuais em grandes corpora bilíngues. Assim, o treinamento dos sistemas passou a ser realizado a partir de corpora paralelos, prescindindo do uso de regras. Dois sistemas fruto dessa abordagem são a tradução por máquina estatística (*Statistical Machine Translation*, SMT) e a tradução por máquina baseada em exemplos (*Example Based Machine Translation*, EBMT). Os sistemas estatísticos podem utilizar modelos baseados em palavras, mas os que atingiram melhores resultados foram aqueles baseados em frases (*Phrase Based Machine Translation*, PBMT). Nos sistemas PBMT, a frase (como uma série de palavras contíguas) é a unidade atômica em que se subdividem as orações no texto de entrada. Então, essas frases são mapeadas a frases no texto de saída, e podem ser reordenadas (KOEHN, 2010).

O último avanço significativo em tradução por máquina é a MT neural (*Neural Machine Translation*, NMT) que, como seu nome indica, utiliza redes neurais artificiais para o processamento. O elemento principal desta nova abordagem é o uso da tecnologia de aprendizagem profunda, que se inspirou no

conhecimento sobre o cérebro humano e está conduzindo à MT a novos níveis de precisão e fluência (VAN DER MEER, 2020).

Inicialmente, ela combinou-se com a MT estatística, mas em pouco tempo se iniciaram as pesquisas para gerar uma MT neural pura. Para se ter uma ideia do avanço que implicou a NMT, Koehn (2010) destaca o rápido aumento da presença da NMT na conferência de referência mundial da área, o *Workshop on Machine Translation* (WMT). Esse evento desenvolve-se a partir de tarefas nas quais uma série de grupos de pesquisa competem, e onde o desempenho é avaliado a partir de métricas de qualidade padronizadas que calculam a similitude entre a MT e a tradução humana dos mesmos textos. Em 2015, apenas um sistema de NMT pura foi apresentado, que apesar de ser competitivo não conseguiu superar a MT estatística. Em 2016, um sistema de NMT pura venceu em quase todas as línguas, e em 2017 quase todos os sistemas concorrentes eram de NMT.

A estrutura dos modelos de NMT é mais simples que a dos modelos estatísticos anteriores baseados em frases, já que não exige o uso de um modelo de linguagem, um modelo de tradução e um de reordenamento separados, mas “um único modelo de sequência que prediz uma palavra de cada vez” e onde a predição da sequência está condicionada pela frase completa na língua de origem, isto é, todas as palavras e seu contexto são codificadas em uma única representação numérica, através da qual o decodificador gera a frase na língua de destino (WAY, 2020, p. 317). A NMT utiliza modelos de treinamento obtidos a partir de corpora paralelos, do mesmo modo que nos modelos estatísticos baseado em frases.⁸

Apesar de apresentar vantagens significativas, os sistemas de NMT também têm pontos desafiantes. Koehn e Knowles (2017) destacam alguns deles, como o fato de apresentarem menor qualidade fora da área de treinamento do modelo, ao ponto de sacrificar totalmente a adequação em benefício da fluência; que eles têm uma curva de aprendizagem mais pronunciada com relação à quantidade de dados de treinamento; a menor qualidade de tradução que oferecem em frases longas, entre outros. Mesmo assim, existe consenso entre muitos dos pesquisados de que a NMT apresenta vantagens significativas em diversos cenários em comparação

⁸ Koehn, P. The State of Neural Machine Translation (NMT). Disponível em: <https://omniscien.com/blog/state-neural-machine-translation-nmt/>

com a SMT, e que já é o sistema de MT mais avançado, embora ainda subsistam problemas na avaliação das melhoras com os modelos existentes, que foram pensados principalmente para a PBMT (WAY, 2020). A adoção ampla por parte da indústria também é um assunto complexo, já que o treinamento dos modelos de NMT é bem mais lento que os de PBMT, os dados disponíveis são insuficientes e o hardware bem mais caro.

3.3

Avanços em MT e interatividade

Paralelamente à procura da FAHQT, houve desenvolvimentos que apontaram no sentido da interação da MT com o tradutor mais centrados na interação humano-máquina (HCI). Esses trabalhos apoiaram-se nas ideias de pesquisadores que sugeriam que possibilitar que os seres humanos fossem potenciados (*augmented*) pelas máquinas era uma abordagem mais realista e razoável que considerar a simples substituição do ser humano pela máquina. Um deles, J. C. R. Licklider, em seu artigo *Man-Computer Symbiosis*, falava em uma aliança simbiótica entre o ser humano e o computador. Talvez prevendo que a substituição do ser humano pelas máquinas não era tão previsível no curto prazo, ele afirmou:

Um grupo de estudo multidisciplinar, examinando os problemas de investigação e desenvolvimento futuros da Força Aérea estimou que para 1980 o desenvolvimento da inteligência artificial faria com que as máquinas sozinhas fossem capazes de pensar e resolver em grande medida os problemas significativos em termos militares. Isso deixaria, digamos cinco anos para desenvolver a simbiose ser humano-computador e 15 anos para usá-la. Esses 15 anos podem ser 10 ou 500, mas esses anos deverão ser os mais criativos e entusiastas de história. (LICKLIDER, 1960, p. 4).

O racoconto que Green *et al.* (2015) fazem da evolução dos sistemas interativos inclui os seguintes projetos, que principalmente procuravam reduzir o tempo e o esforço de pós-edição:

- a) Projeto MIND, desenvolvido por Martin Kay e sua equipe da RAND Corporation como o primeiro do tipo, na década de 1960. Apesar de

que nunca chegou a ser construído, o MIND considerava a necessidade de pré-edição e pós-edição para complementar a tradução por máquina.

- b) *Interactive Translation System (ITS)*, uma iniciativa conduzida por Alan Melby na *Bringham Young University*, em 1973, como parte do projeto *Automated Language Learning (ALP)*, para a qual os computadores estariam a serviço dos tradutores humanos, ao invés de simplesmente substituí-los. Melby concluiu que a interação humano-máquina exigiria operadores treinados, o que aumentaria ainda mais os custos, além do fato de que a maioria dos tradutores não gostavam de pós-editar. Por esses motivos, o projeto nunca chegou na etapa de produção.
- c) Projetos *TransType* (primeiro sistema de MT estatístico interativo) e *TransAhead*, que oferecem, através de texto preditivo (recurso que estamos acostumados a utilizar ao escrever em programas de mensagens nos smartphones ou ao fazer buscas no Google, por exemplo), opções geradas por MT adaptável que o tradutor pode aceitar ou rejeitar e cujo objetivo é aumentar a produtividade.

Pelo peso conferido à pós-edição e pela facilitação do trabalho dos tradutores, os projetos antes descritos têm mais proximidade com a HCI, uma vez que a edição de texto é um exemplo paradigmático da própria interação humano-máquina pelos seguintes motivos: “(1) a interação é veloz; (2) a interação se transforma em uma extensão inconsciente do usuário; (3) os editores de texto são provavelmente os programas informáticos mais utilizados; e (4) os editores de texto são representativos de outros sistemas interativos” (GREEN *et al.*, 2015, p. 8).

Apesar da ideia dos sistemas de MT interativos estarem presentes desde a década de 1960, as pesquisas envolvendo HCI foram deixadas de lado em benefício da procura da FAHQT. O motivo dessa marginalização no momento inicial das pesquisas, segundo Green *et al.* (2015), deveu-se ao enviesamento dos pesquisadores de IA, que mantiveram como objetivo principal a substituição total do ser humano no processo de tradução, o que fez com que áreas que compartilhavam objetivos estratégicos similares tivessem uma interação muito limitada durante vários anos.

3.3.1

A MT interativa no século XXI

A melhora dos resultados da MT estatística baseia-se no treinamento, isto é, na adição de traduções de boa qualidade para garantir que, no futuro, os melhores candidatos sejam os escolhidos. Esse processo é demorado e caro, especialmente quando se trata de volumes importantes de informação. A MT adaptável (Adaptive MT) veio para reduzir esse tempo e esforço aproveitando de imediato as correções feitas pelos tradutores no processo de pós-edição.

Em 2014, Green, Heer e Manning desenvolveram o projeto *Predictive Translation Memory* (PTM), uma iniciativa híbrida que utiliza MT e TM na qual o ser humano e a máquina interagem para refinar as traduções e que se enquadra no objetivo acima mencionado. Os desenvolvedores definiram o PTM como um novo sistema de tradução interativa de iniciativa mista (GREEN *et al.*, 2014). A diferença da pós-edição normal, onde o tradutor interage com segmentos previamente traduzidos, o sistema de MT gera sugestões que o tradutor vai selecionando e incorporando ao texto na medida em que traduz. O tradutor também pode escolher, a qualquer momento, incorporar a frase completa oferecida pelo sistema de MT, caso a considere correta ou se avaliar que exigirá poucas alterações. O sistema aproveita avanços no PLN, na busca rápida e no aprendizado de parâmetros online, e o resultado é uma interação que vai aumentando a qualidade do que a MT oferece progressivamente, na medida em que o tradutor avança no projeto e alimenta o sistema com suas escolhas.

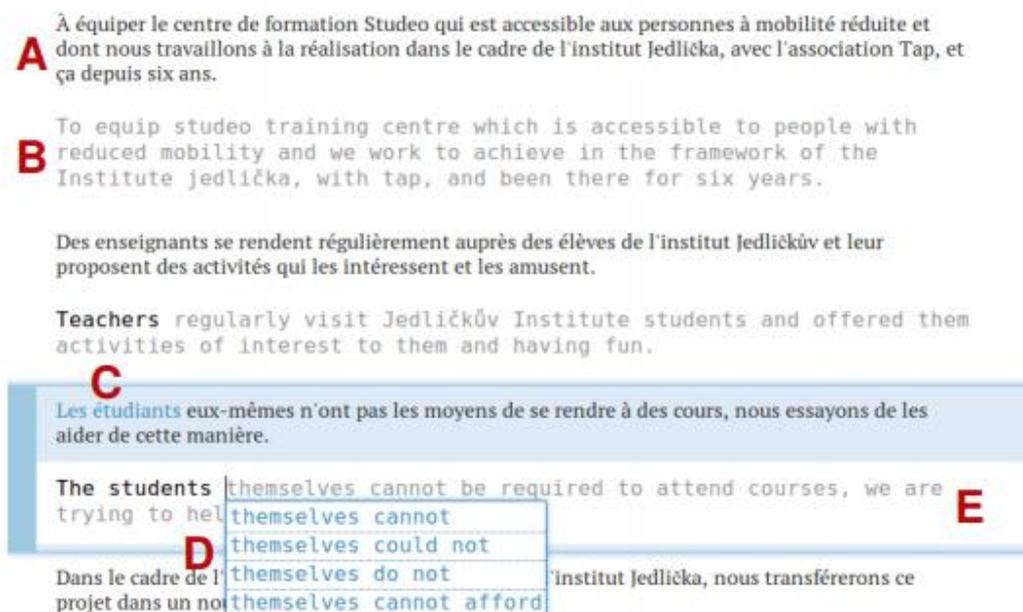


Figura 5 – Interface do PTM.

Onde é possível ver o contexto do documento completo, com a tradução completa da frase e uma caixa onde são oferecidos fragmentos.

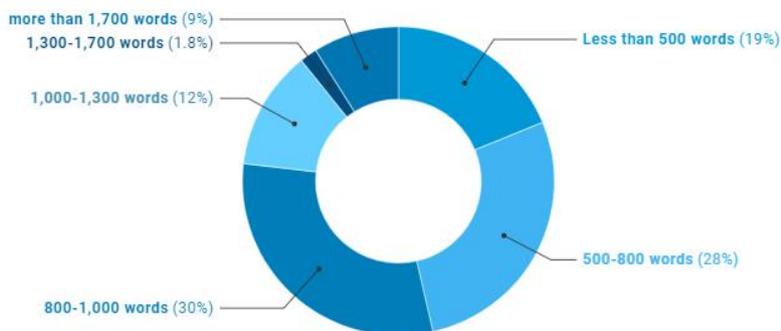
Fonte: GREEN *et al.* (2014).

Em 2015, John DeNero (professor de Ciências da Computação em UC Berkeley) e Spence Green (PhD em Ciências da Computação de Stanford) fundaram Lilt, uma empresa pioneira que oferece um ambiente de TM com Adaptive MT. Inspirada claramente no projeto PTM, a principal inovação da Lilt é o chamado “circuito de feedback neural”, primeiro sistema do seu tipo em aumentar a produtividade do tradutor por meio de IA e um ciclo de *feedback* em tempo real. O sistema de NMT desenvolvido pela empresa observa as preferências do tradutor para gerar sugestões cada vez melhores, na medida em que são dirigidas pelas escolhas anteriores, o que se alinha diretamente com as ideias que enfatizam a interação humano-máquina. O primeiro estudo de caso de aplicação da MT adaptável da *Lilt* em larga escala (1,77 milhões de palavras traduzidas para seis idiomas em dez dias por 100 tradutores) indicou uma alta produtividade: 820 palavras/hora para inglês > português, e 701 palavras/hora para inglês > espanhol. Mesmo sendo uma produtividade elevada, cabe destacar que não está muito longe da média obtida com a PEMT normal, com traduções providenciadas por um sistema de alta qualidade, como indica uma pesquisa feita pelo blog Slator entre

tradutores profissionais, onde mais da metade dos participantes (58%) afirmaram que conseguiam produzir de 500 a 1000 palavras/hora (Figura 3).

What's a realistic PEMT per-hour output to produce high quality translation assuming good quality MT as the input?

Poll on May 3, 2019. Vote Count: 112



Source: Slator Weekly Email Newsletter • Get the data • Created with Datawrapper

Figura 6 – Produção realista por hora como PEMT.

Qual é uma produção realista por hora com PEMT para produzir uma tradução de alta qualidade assumindo uma MT de boa qualidade como ponto de partida?

Fonte: BLOG SLATOR (2019).

Outros sistemas de Adaptive MT funcionam de maneira integrada em ambientes de CAT Tool. É o caso da AdaptiveMT lançada pela SDL em 2017, e a ModernMT lançada em 2021, que podem ser utilizadas de maneira integrada (através de uma API) com ferramentas como SDL Trados Studio e o MateCat.

Hoje, o uso da MT parece apontar em duas direções diferentes. Por um lado, continua a procura de atingir o objetivo almejado pelos pioneiros da MT, a FAHQT, isto é, a tradução totalmente automática de alta qualidade, onde se poderia prescindir totalmente da intervenção humana. Para isso, a indústria aposta no aperfeiçoamento dos sistemas de MT, na utilização de linguagem controlada para a elaboração dos documentos a serem traduzidos e na retroalimentação dos sistemas de MT a partir da pós-edição automática (APE – *Automatic Post-editing*), “uma área que explora métodos de aprendizagem a partir de dados pós-editados por tradutores para obter melhores resultados de MT” (CARMO *et al.*, 2020). Assim, um sistema de APE utiliza o texto fonte, o texto traduzido e as correções realizadas pelo pós-editor para aprender o necessário para melhorar a MT.

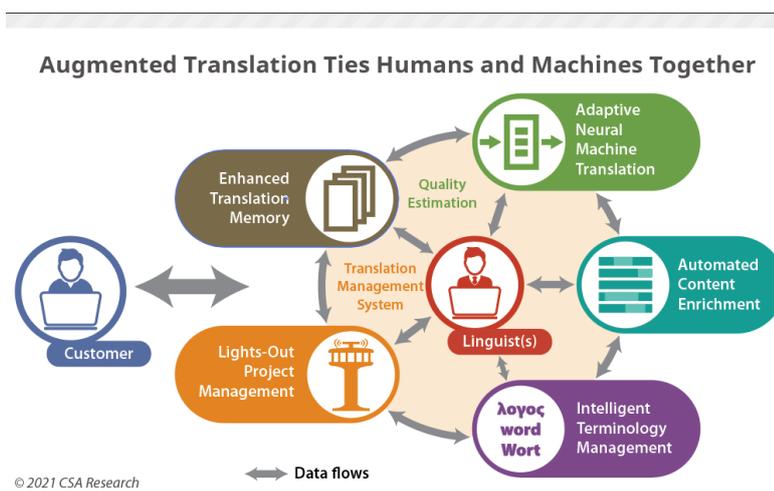


Figura 7 – *Augmented Translator*.
Fonte: CSA (2021).

Por outro lado, é possível observar o surgimento de um modelo em que o tradutor continua tendo um papel central em estreita colaboração com a MT, em uma relação simbiótica humano-máquina, o que a consultora Common Sense Advisory denomina *Augmented Translator*.⁹ Nesse modelo, o tradutor não estaria numa posição reativa de limpar as traduções obtidas por MT, mas estaria posicionado no meio do processo, participando do treinamento dos sistemas de tradução neural em tempo real (DEPALMA, 2021). A ferramenta Lilt, antes descrita, seria um exemplo dessa abordagem.

As duas tendências acima descritas podem ser mapeadas aos caminhos que a MT seguiu a partir das suas etapas de desenvolvimento iniciais, analisadas no capítulo anterior.

3.4

As ferramentas CAT como um caso de interação bem-sucedido

As chamadas ferramentas CAT, também conhecidas como Ferramentas de Ambiente de Tradução (*Translation Environment Tools*, TEnT), constituem um caso de sucesso orientado pela HCI. Elas foram desenvolvidas com o objetivo de

⁹ A ideia do “*augmented*” no sentido de incrementar a capacidade do ser humano de abordar problemas complexos foi desenvolvida inicialmente por Engelbart, da década de 1960. Cf. <https://www.dougenelbart.org/content/view/138>.

umentar a produtividade dos tradutores, ao mesmo tempo em que melhoram a qualidade e uniformidade das traduções. Apesar de terem surgido nas décadas de 1970 e 1980, seu uso popularizou-se na década de 1990 em resposta ao aumento da demanda por parte de empresas e instituições e a expansão do uso de computadores em empresas e lares.

Em poucas palavras, uma ferramenta CAT é um software desenvolvido para reutilizar traduções anteriores armazenadas em um repositório. Elas oferecem um ambiente de edição que permite interagir com dois elementos básicos: uma memória de tradução (*Translation Memory*, TM) e uma base terminológica (*Termbase*, TB).

Uma TM é uma base de dados que armazena segmentos na língua de origem e os segmentos correspondentes na língua de destino, o que se denomina unidade de tradução (*Translation Unit*, TU). Através das TMs, as ferramentas CAT permitem o reaproveitamento de textos traduzidos, considerando-se que em tradução especializada existe uma grande taxa de repetição, isto é, de frases que se repetem de forma exata ou parcial entre documentos ou, inclusive, dentro de um mesmo documento. As TMs são alimentadas pelo tradutor durante seu trabalho e ao mesmo tempo oferecem ao tradutor segmentos iguais ou similares (até em 75% de coincidência, chamada de coincidência parcial) traduzidos previamente para seu reaproveitamento. No caso de coincidências parciais, a janela da memória mostra exatamente qual é a diferença entre o segmento previamente armazenado e o novo segmento a traduzir, para facilitar a intervenção do tradutor.

As TBs permitem armazenar e recuperar termos durante o processo de tradução. Ao serem usadas de maneira integrada nas ferramentas CAT, a tecnologia chamada *push* faz com que tradutor não precise procurar os termos que constam na TB, mas eles são reconhecidos automaticamente pelo programa e destacados de alguma maneira. No caso dos termos incluídos na TB, eles são oferecidos quando um termo que consta no glossário aparece no texto na língua de origem. Nesses casos, o termo é destacado de alguma maneira no texto fonte (por exemplo, com uma linha vermelha na parte superior), é mostrado na janela de terminologia e, quando o tradutor começa a digitá-lo, ele aparece em uma pequena janela onde o tradutor só precisa selecioná-lo e pressionar a tecla *Enter* para introduzi-lo no texto.

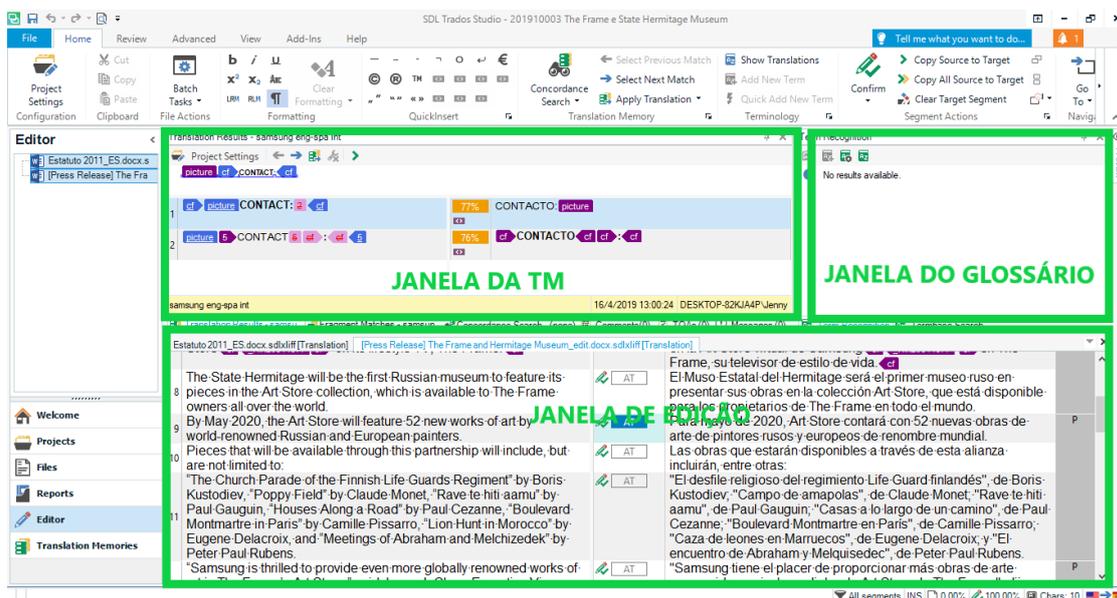


Figura 8 – Interface do SDL Trados Studio.

Captura de tela mostrando a interface da vista de edição do SDL Trados Studio com a descrição das suas janelas principais.

Fonte: Elaboração própria (2021).

Embora as primeiras ferramentas CAT só oferecessem a interação com a TM e a TB, elas foram evoluindo e incluindo aos poucos a possibilidade do tradutor utilizar outros recursos para aumentar a produtividade, como dicionários de texto preditivo gerados a partir de uma TM específica, fragmentos de segmentos, segmentos reparados automaticamente e MT oferecida por fornecedores como Google, Microsoft, SDL/RWS, DeepL ou ModernaMT, entre outros. Como veremos a seguir, o uso desses recursos dentro do ambiente de edição das ferramentas CAT produz-se em grande velocidade e eles funcionam como uma extensão quase inconsciente do tradutor, elementos considerados como paradigmáticos da HCI, segundo Green *et al.* (2014), como vimos no ponto anterior.

Os dicionários de texto preditivo são complementos que oferecem sugestões automáticas durante a tradução no nível de subsegmento. Eles são criados a partir das próprias TMs, portanto utilizam as traduções prévias como fonte, em lugar de um dicionário ou qualquer outra base de dados genérica. Ao fornecer termos de uso habitual na área do tradutor, evita que ele tenha que digitar palavras ou conjuntos de palavras, o que permite aumentar a produtividade e evitar erros de digitação. A maneira de incorporar os termos ao texto é igual à dos termos das TBs (ver Figura 9).

Em algumas ferramentas CAT o tradutor também pode receber sugestões de fragmentos de frases armazenadas (como resultado de um processo chamado de *upLiFT* no *SDL Trados Studio*) e inclusive fragmentos “reparados” (isto é, alterados considerando as diferenças com relação ao novo segmento) automaticamente (processo denominado *upLiFT Fuzzy Repair* na mesma ferramenta), que se apresentam em uma pequena janela e podem ser incorporados da mesma maneira que nas opções anteriores (ver Figura 9).

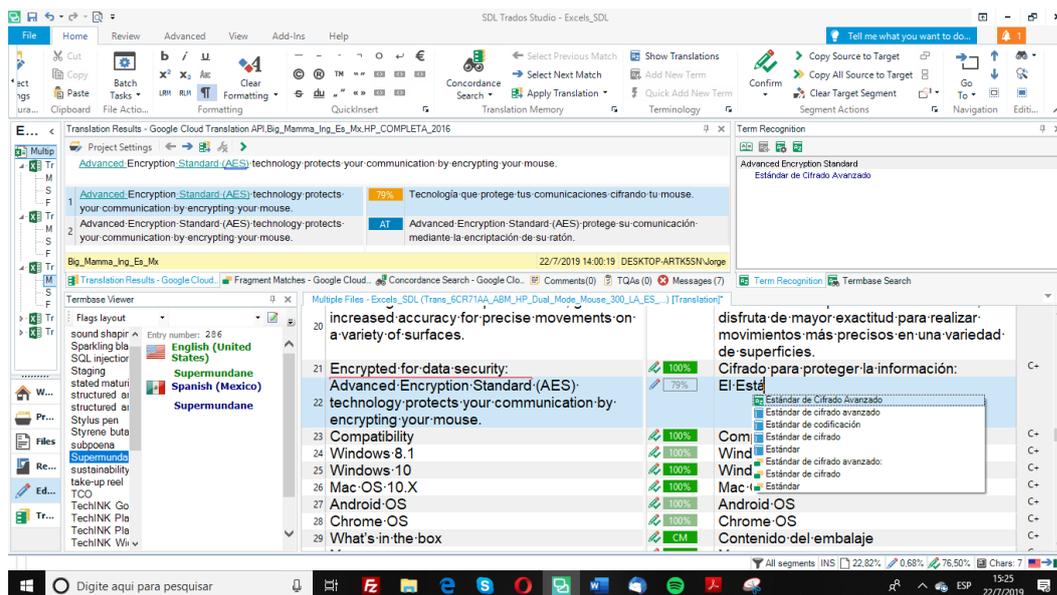


Figura 9 – Recursos interativos nas ferramentas CAT (SDL Trados Studio 2019).

As sugestões estão identificadas pela sua origem (ícone de livro verde: TB; ícone de livro azul: AutoSuggest Dictionary; ícone verde-amarelo (fragmento de TU). A linha vermelha acima do termo indica que ele consta do glossário.

Fonte: Elaboração própria (2021).

A partir de 2010, diversas ferramentas CAT começaram a oferecer integração com sistemas de MT, como Google Translate, DeepL, ModernMT. Essa integração é feita através de uma API, uma espécie de programa “ponte” que permite conectar aplicativos ou plataformas. Assim, o *input* proporcionado pelos sistemas de MT pode ser utilizado dentro do ambiente das ferramentas CAT basicamente de duas maneiras. A primeira é no nível de segmento, isto é, o texto é pré-traduzido (de maneira total ou segmento por segmento, na medida em que o tradutor vai avançando) e o tradutor faz a pós-edição, aplicando as correções necessárias. Esse procedimento é conhecido como pós-edição ou PEMT (*Post-*

Editing Machine Translation), descrito com maior detalhe no Capítulo 2. Essa é a maneira mais habitual de utilização de MT. Porém,

muitos tradutores perceberam que a MT pode ser usada de outra maneira que simplesmente corrigindo erros. Em vez disso, pode ser usada como um recurso adicional com fragmentos úteis que os tradutores podem escolher usar ou ignorar (ZETZSCHE, 2020, p. 177).

Assim, cada vez mais ferramentas permitem também o uso de MT no nível de subsegmento, isto é, oferecendo ao tradutor fragmentos de segmentos de diversas extensões provenientes da MT que o tradutor pode incorporar no lugar que ele achar mais conveniente simplesmente pressionando a tecla Enter. Os fragmentos, que podem ir desde algumas poucas palavras até a frase quase completa, e mudam de maneira dinâmica na medida em que o tradutor avança na tradução do segmento.

Essa modalidade de uso oferece ao tradutor o controle da organização da frase, da ordem das palavras e inclusive da vantagem ou não de incorporar (ou incorporar só parcialmente) o que a TM está propondo. Dependendo do sistema de MT integrado na ferramenta CAT, essa interação pode ser de uma via, isto é, o tradutor só recebe *input* do sistema de MT, mas não proporciona nenhum *output* para ele, ou de via dupla, quando o sistema de tradução utilizado é do tipo MT adaptável, onde a tradução é enviada para o sistema para ser aproveitada para melhorar futuras.

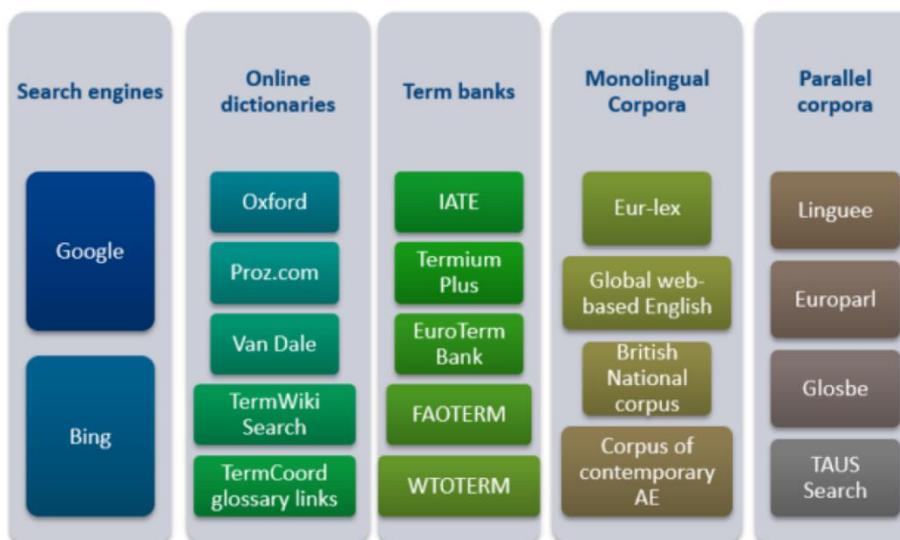


Figura 10 – Recursos mais utilizados pelos tradutores.
Fonte: VANDEGHINSTE *et al.* (2019).

Apesar das diversas opções até aqui mencionadas, muitas vezes os tradutores esgotam os recursos dentro das ferramentas CAT e devem fazer buscas em recursos locais ou *online*. Entre esse tipo de recursos, Vandeghinste *et al.* (2019) identificara os seguintes: buscadores (Google, Bing etc.); dicionários online; bancos terminológicos; corpora monolíngues; e corpora paralelos (ver Figura 10). Também existem *plug-ins* que permitem fazer pesquisas terminológicas *online* desde o interior das ferramentas CAT, como o que possibilita o uso dentro do SDL Trados Studio do *IntelliWebSearch*, um programa que permite fazer buscas terminológicas *online* a partir de diversos sites pré-configurados. Cabe destacar que esse tipo de busca implica uma interrupção do fluxo para o tradutor, que deve sair da interface da CAT e trabalhar em outra janela (no caso de não usar dois monitores, algo bastante habitual entre tradutores profissionais).

Finalmente, é importante assinalar que, se em um primeiro momento as ferramentas CAT se diferenciaram claramente da MT no sentido que na primeira que fazia a tradução era o ser humano e na segunda, um sistema informático, a integração crescente entre ferramentas CAT e MT mencionada neste ponto de alguma maneira fez com que essa divisão fique menos nítida e o trabalho humano e a MT tendam a se imbricarem cada vez mais profundamente.

4

Considerações sobre as “viradas” nos estudos da tradução e o papel da tecnologia

No presente capítulo se apresentarão algumas considerações teóricas sobre a tecnologia e os Estudos da Tradução e se discutirá a adequação de considerar a existência de uma “virada tecnológica”.

4.1

As viradas nos Estudos da Tradução e a tecnologia

Desde que Holmes publicou seu artigo seminal *The Name and Nature of Translation Studies* (1972/1988), que ele mesmo considerou como um mapa de uma utopia disciplinar, e que hoje é reconhecido por muitos autores como o documento fundador dos Estudos da Tradução (ET) modernos, houve muitas mudanças significativas no campo. O uso crescente da tecnologia por parte dos tradutores de todas as áreas talvez seja uma das mudanças mais pervasivas e marcantes, especialmente considerando a forte penetração das ferramentas CAT a partir da década de 1990 e o aumento da presença da MT no fluxo de trabalho dos profissionais, especialmente dos tradutores especializados, a partir dos anos 2000. É possível observar um exemplo disso no espaço que a tecnologia ocupa, por exemplo, na obra *Translation Studies*, da Susan Bassnett, publicada em 1980, onde a tecnologia aparece com uma única menção solitária à pesquisa de tradução baseada em corpus da Mona Baker, embora seja reconhecida como uma área com um futuro promissor. No início do século XX, porém, é notável a publicação de obras que colocam a tecnologia em um lugar de maior centralidade nos ET, como se verá no próximo ponto onde se discute a existência ou não de uma “virada tecnológica”.

Ao analisar a evolução dos ET desde suas origens, Mary Snell-Hornby (2006) destaca algumas mudanças significativas de ponto de vista no desenvolvimento da disciplina, que ela denomina *turns* (viradas). A autora destaca que não se tratam de etapas claramente definidas, mas que é possível observar limites difusos, sobreposições e brechas, e onde sempre é possível achar pioneiros em períodos prévios. Depois de explorar as raízes dos ET, que se situam no romantismo alemão, e os primeiros anos, Snell-Hornby assinala como a primeira mudança significativa a “virada pragmática”, na década de 1970, que impulsionaria a libertação dos estudos da tradução, tanto da linguística como da literatura comparada. A tradução deixou de ser vista como uma simples troca de elementos da linguagem para começar a ser considerada como “um ato concreto que envolve participantes e uma situação extralinguística, o que nesse momento era uma novidade, pelo menos em tradução não literária” (SNELL-HORNBY, 2006, p. 27). A “virada pragmática” receberia um impulso definitivo nos trabalhos posteriores na área da linguística de Austin e Searle (na década de 1960) e a inclusão dos aspectos sociais e comunicacionais e o surgimento da linguística textual. Finalmente, Snell-Hornby destaca que a “virada pragmática” permitiu assentar as bases dos Estudos da Tradução com o trabalho integrador de Holmes, que permitiu que a tradução se emancipasse da teoria literária e da linguística e se estabelecesse como uma disciplina independente (SNELL-HORNBY, 2006, p. 3).

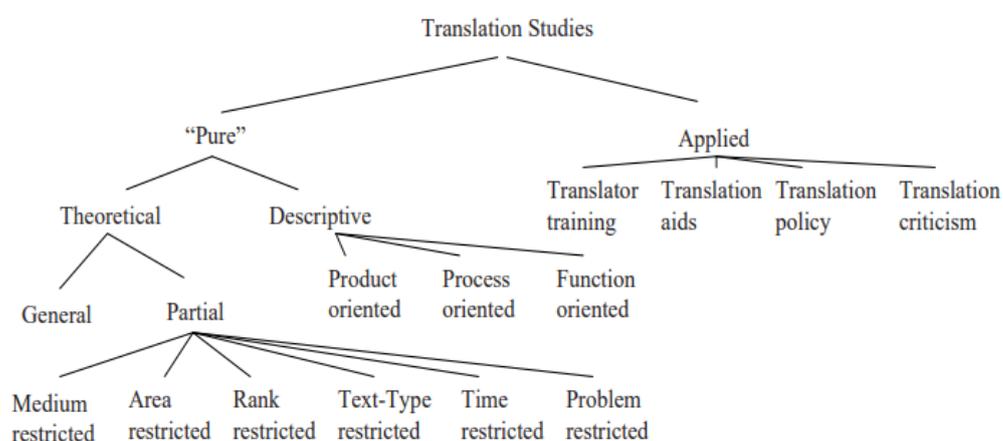


Figura 11 – Mapa dos Estudos da Tradução baseado em Holmes-Toury.
Fonte: CHESTERMAN (2009).

A segunda virada identificada por Snell-Hornby é a “virada cultural”, que colocou ênfase na função das traduções nas culturas de destino, em uma abordagem funcional que fez abandonar as perspectivas prescritivas, fundamentalmente orientadas ao texto de origem. O trabalho de Toury e outros pesquisadores em Israel e nos Países Baixos completaria o olhar de Holmes e permitiria a consolidação dos Estudos Descritivos da Tradução (*Descriptive Translation Studies*, DTS), surgidos na década de 1970. Segundo Martins:

O modelo dos DTS fundamenta-se na suposição de que traduzir é uma atividade orientada por normas culturais e históricas; a própria escolha dos textos a serem traduzidos, as decisões interpretativas tomadas durante o processo tradutório e a divulgação, a recepção e a avaliação das traduções são fatores consideravelmente influenciados pelos distintos contextos socioculturais observados em momentos específicos da história. (MARTINS, 2002, p. 34).

Assim, a “virada cultural” significou o abandono definitivo da abordagem linguística e trouxe uma mudança do foco de interesse, que passou do “texto” à “cultura” (SNELL-HORNBY, 2006).

Dentro das mudanças relacionadas com a “virada cultural” têm especial destaque os estudos do tipo funcionalista, onde o conceito de cultura ganhou um rol central e nos quais os textos foram considerados como partes verbalizadas de uma cultura, que fazem parte de uma situação determinada e, portanto, que sofrem um condicionamento devido a essas circunstâncias. Uma das abordagens mais relevantes nessa linha foi a *Skopostheorie*, com Hans J. Vermeer como um dos seus principais expoentes. Nas suas palavras:

Qualquer forma de ação tradutória, inclusive a própria tradução, pode ser concebida como uma ação, como seu nome indica. Qualquer ação tem um objetivo o propósito para a tradução. [...] A palavra *skopos*, portanto, é um termo técnico para o objetivo ou o propósito da tradução (VERMEER, 1989, p. 191).

Na década de 1990, Snell-Hornby observa uma “virada empírica” nos ET, com trabalhos que tentavam estudar os processos mentais que acontecem na tradução, com a utilização do método de “pensar em voz alta” (*think aloud*), emprestado da psicologia, que implicava a verbalização por parte do tradutor dos seus pensamentos durante a tradução. Essa verbalização era gravada, transcrita e analisada com posterioridade.

Outra vertente da “virada empírica” foi o trabalho com corpora de grande escala informatizados, cuja potencialidade foi resumida por Mona Baker no livro *Text and Technology: In Honour of John Sinclair* (1993), quando afirmou que o “efeito profundo que os corpora terão nos estudos da tradução [...] serão uma consequência de capacidade que nos fornecerão de identificar características dos textos traduzidos que ajudarão a compreender que é a tradução e como funciona” (BAKER, 1993, p. 243).

A última virada mencionada por Snell-Hornby é a “virada da globalização”, considerando o forte impacto em todos os aspectos culturais resultantes da aceleração desse processo. Nesse sentido, ela destaca a noção da “identidade cultural” e o fenômeno da linguagem como expressão de comunidades culturais autoconscientes, bem como a tensão que existe entre a globalização e os nacionalismos, e até levanta a discussão do papel do setor de serviços linguísticos no ressurgimento do nacionalismo, a identidade cultural e a globalização (SNELL-HORNBY, 2006, p. 130). Nessa linha também entrariam outras mudanças de abordagem, como a perspectiva de gênero, a pós-colonial, a relacionada com as identidades etc., que Munday (2012) denomina “viradas ideológicas”.

Dentro da “virada da globalização”, Snell-Hornby destaca três assuntos: a tecnologia e o tradutor, a tradução e publicidade e a prevalência do inglês. Com relação ao primeiro ponto, a autora destaca que a relação tradutor-tecnologia vinha sendo um tema recorrente nas pesquisas profissionais, impulsionada pela constatação do fortíssimo impacto que a tecnologia teve na atividade e no perfil dos tradutores profissionais. Também ressalta o surgimento de novas áreas, como a terminologia, a tecnologia da linguagem (que inclui os sistemas de memória de tradução) e a tradução automática, embora ela assinale que, em grande medida, essas pesquisas tiveram rápidos desenvolvimentos mais relacionados com disciplinas como as ciências da computação e a linguística que com os Estudos da Tradução.

Alguns autores defendem que a mudança de rumo a partir dos anos 2000 poderia ser considerada como uma “virada sociológica”. Essa “virada” indicaria a mudança paradigmática que implica a centralidade da reflexão sobre os condicionamentos do processo de tradução, uma vez que a tradução é compreendida como uma prática social. Segundo Wolf (2007), a emergência da

“virada sociológica” teria a vantagem de incluir os dois níveis que condicionam o processo de tradução: o cultural e o social. O primeiro nível seria estrutural e incluiria fatores como poder, religião e economia, e o segundo se debruçaria nos agentes envolvidos no processo de tradução, que atuam a partir da internalização das estruturas antes mencionadas (WOLF, 2007, p. 4).

Assim, essa mudança de perspectiva seria mais abrangente que a da “virada cultural” identificada por Snell-Hornby. Sobre o assunto, Snell-Hornby sustenta que devido ao fato das implicações sociais nas traduções estarem presentes há muito tempo nos ET, isso não constituiria uma “virada” propriamente, embora considere a sociologia da tradução como uma abordagem muito bem-vinda e promissora como alternativa às abordagens puramente linguísticas (SNELL-HORNBY, 2006, p. 172).

Outra “virada” foi assinalada por Chesterman (2009), considerando o crescente número de pesquisas centradas nos tradutores. Segundo ele, a visão de Holmes dos ET dava um peso desproporcionado aos textos, em lugar das pessoas que os produzem. Para compensar isso, ele propôs o modelo dos agentes, isto é, um modelo de ET centrado nos agentes envolvidos no processo de tradução. Partindo dessa base, Chesterman dividiu o modelo de Estudos do Tradutor em três ramos: cultural, cognitivo e sociológico (Figura 8). O ramo cultural inclui os valores, ética, ideologias, tradições e história que influenciaram os tradutores e intérpretes, considerando-os como agentes de evolução cultural; o ramo cognitivo aborda os processos mentais, a tomada de decisões, a personalidade, as atitudes etc.; e o ramo sociológico lida com as condutas observáveis dos indivíduos, grupos e instituições, bem como as redes sociais e assuntos como o status, as relações com outros grupos, as tecnologias utilizadas etc. (CHESTERMAN, 2009, p. 19).

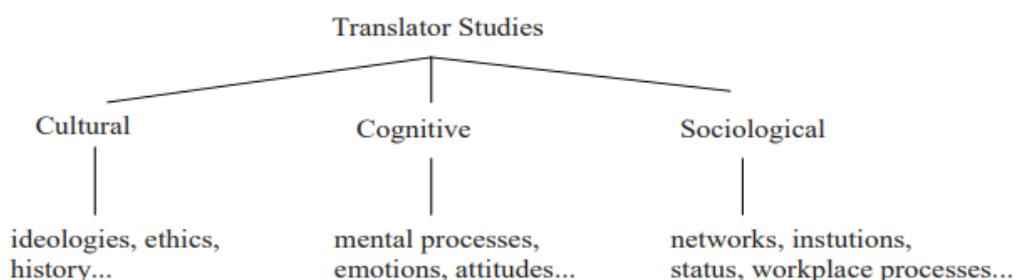


Figura 12 – Esquema dos Estudos do Tradutor.
Fonte: CHESTERMAN (2009).

Finalmente, entre outras “viradas” possíveis propostas por diversos pesquisadores, podemos destacar a “virada da tradução audiovisual”, que cobre “todo tipo de transferência de textos audiovisuais entre duas línguas e culturas (interlíngua) ou dentro da mesma língua e cultura (intraíngua)” (CHAUME, 2013, p. 4), como as que acontecem na dublagem, legendagem, audiodescrição, legendagem para pessoas com deficiências auditivas, dublagem e legendagem feitas por fãs (*fandubbing* e *fansubbing*).

4.2

É possível falar em uma “virada tecnológica”?

A partir de década de 1990 houve um crescimento importante das pesquisas empíricas que involucram assuntos relacionados com a tecnologia, com um número significativo de publicações abrangendo aspectos diversos. Considerando isso, alguns autores identificam nisso uma “virada tecnológica” que, de maneira similar às viradas até aqui mencionadas, representaria uma mudança de foco nos ET. Michael Cronin (2010), por exemplo, sustenta a ideia de que as novas maneiras em que processo de tradução se desenvolve exigem a necessidade de reexaminar o que constitui uma tradução e qual seria o papel do tradutor. Cronin parte do pensamento do Marshall McLuhan, filósofo e teórico da comunicação canadense pioneiro nos estudos das transformações sociais provocadas pela tecnologia, especialmente na ideia de que os processos político-culturais estão mais atrelados à tecnologia dos meios de comunicação que ao conteúdo das mensagens transmitidos através deles. Nessa premissa, Cronin afirma que a “virada tecnológica” seria “resultado de mudanças significativas na maneira como se realiza a tradução no mundo contemporâneo [...] que exige a reavaliação sistemática da compreensão de que constitui uma tradução e da posição do tradutor” (CRONIN, 2010, p. 1). Para apoiar sua ideia, o autor apresenta como exemplos as mudanças na maneira como as traduções se produzem, são utilizadas e compreendidas pelos públicos-alvo. No primeiro caso, o autor menciona como exemplo a tradução colaborativa (*crowdsourcing*), onde muitos voluntários

traduzem conteúdo. Talvez um dos casos mais conhecidos dessa modalidade seja o TED Open Translation Project, uma comunidade de voluntários criada para traduzir e legendar as apresentações do TED que, na data em que esta dissertação foi redigida, tinha 38.934 tradutores que já haviam traduzido 182.545 TEDs para 115 idiomas, segundo seu sítio web.¹⁰ Cronin sustenta a ideia de que isso representa uma quebra fundamental nos Estudos Descritivos da Tradução, que se apoiam fortemente na suposição de que as traduções são produzidas por um agente ativo para serem consumidas por um público passivo. No caso da tradução colaborativa, os papéis do produtor e do consumidor da tradução se confundem no “prosumidor” (*prosumer*), isto é, um consumidor que ao mesmo tempo faz parte da produção daquilo que consome.

Um segundo fenômeno mencionado por Cronin são as mudanças na maneira de consumir o conteúdo escrito nos meios digitais (fundamentalmente, a internet). Ele cita diversos estudos que determinaram que o tempo que as pessoas ficam em páginas web, inclusive em páginas com muito conteúdo, é de apenas uns segundos (entre 19 e 27 mais precisamente) antes de passar para a página seguinte, segundo uma pesquisa feita em vários países por uma empresa israelense chamada Clicktale, em 2008. Assim, Cronin sustenta que, devido a um poder de navegação acelerado, a internet impulsiona o abandono da leitura fixa, cumulativa e linear, com a qual a tradição da tradução tem fortes vínculos. Autores como Biau Gil e Pym (2006) destacam, também, que a natureza dos textos que o tradutor traduz, que incluem referências cruzadas muitas vezes em forma de enlaces (*links*), também colabora para a substituição dessa leitura linear que tem um princípio e um fim ao permitir pular de texto em texto em um encadeamento virtualmente infinito, o que, segundo os autores, faz com que o termo “leitores” tenha sido substituído pelo de “usuários” (BIAU GIL; PYM, 2006, p. 11).

Além disso, Cronin (2010) cita o uso cada vez mais frequente de tradução automática (por exemplo, utilizando o Google Tradutor) por parte dos usuários de internet para compreender a ideia geral (*gist translation*) de um texto, que leva à aceitação de uma tradução de qualidade inferior, o que cria uma nova maneira do leitor aproximar-se do conteúdo web traduzido, mais instrumental e acelerado.

Para Cronin:

¹⁰ Cf.: <https://www.ted.com/participate/translate>.

A virada tecnológica não é impulsionada por desenvolvimentos teóricos em áreas de pesquisa afins, apesar de que pode ser informada por elas, mas se trata de uma propriedade emergente de novas formas de práticas de tradução. Isto é, a virada em questão é o resultado de mudanças significativas na forma em que a tradução se realiza no mundo contemporâneo. Estas mudanças demandam que a compreensão convencional do que constitui a tradução e a posição do tradutor seja sistematicamente reexaminada. (CRONIN, 2010, p. 1).

Alinhados com Cronin, outros autores sustentam que a ideia de uma “virada tecnológica” também se justificaria a partir da mudança da relação entre o tradutor e a tecnologia. Muitas das pesquisas de finais de século XX sobre a relação entre os tradutores e a tecnologia eram claras em um aspecto central: a subordinação da tecnologia às escolhas e necessidades do tradutor. Em tal sentido, O’Hagan (2021) destaca que, ao especular sobre o futuro da tecnologia da tradução, autores como Kay e Melby, entre outros, imaginaram o tradutor como alguém que estava no controle das tecnologias, que eram utilizadas no seu fluxo da maneira que ele considerasse mais eficiente. Porém, a autora destaca que hoje essa centralidade e controle do tradutor estão sendo questionados devido a transformação que os fluxos de trabalho sofreram. Inclusive, O’Brien (2012) chega a caracterizar a tradução especializada hoje (e talvez poderiam ser incluídos outros tipos de tradução) como uma forma de interação ser humano-computador e Risku (2010) afirma que as ferramentas de tradução não podem ser consideradas simplesmente como artefatos auxiliares, mas devem ser vistas como partes de uma rede complexa, onde os computadores funcionam como extensões da memória e o conhecimento humanos. Nessa linha de pensamento, Alonso e Calvo cunharam o termo “tradutores trans-humanos” para descrever os profissionais que “assumem um rol diferente no processo de tradução, interagindo com a tecnologia como se fosse uma extensão das suas capacidades e estabelecendo um processo com uma dimensão social, criativa e de aprendizagem” (ALONSO; CALVO, 2015).

Mais um argumento em favor da existência de uma “virada tecnológica” estaria na constatação retrospectiva da publicação de numerosas obras que abordam o tema e de um número crescente de pesquisas e matérias publicadas em revistas especializadas nos últimos anos. Entre as primeiras podemos mencionar a *Routledge Encyclopedia of Translation Technology*, editada por Chan Sin-wai (2015); *Translation and Technology*, de Chiew Kin Quah (2006); *Translation and Globalization*, de Michael Cronin (2003); e *The Future of Translation*

Technology: Towards a World without Babel, de Chan Sin-wai (2016); entre muitas outras.

No referente às pesquisas e publicações em revistas especializadas, Christensen, Flanagan e Schjoldager, três pesquisadoras da universidade de Aarhus, Dinamarca, fizeram, em 2017, um mapeamento da pesquisa de Tecnologia da Tradução nos ES na introdução de uma seção temática da revista *Hermes*. Elas pesquisaram 9 publicações online da área que não publicavam exclusivamente artigos relacionados com ferramentas CAT e MT: 1) *JosTrans – Journal of Specialised Translation*, 2) *Hermes*, 3) *Perspectives – Studies in Translatology*, 4) *Meta: Translators’ Journal*, 5) *Linguistica Antverpiensia – New Series – Themes in Translation Studies*, 6) *Target*, 7) *The Translator*, 8) *Fachsprache*, e 9) *Trans-kom*. A partir de uma seleção de termos, fizeram uma triagem de artigos publicados entre 2006 e 2016, e os classificaram em três categorias: 1) pesquisas orientadas para a tecnologia; 2) pesquisas orientadas para o fluxo de trabalho o e setor de serviços linguísticos; e 3) pesquisas em tradução teórica.

	Number of issues from 2006-2016	Total number of papers analysed	Number of papers dealing with translation technology	Percentage of papers dealing with translation technology
JosTrans	22	230	67	29.1
Hermes	20	168	6	3.6
Perspectives	43	284	24	8.5
Meta	40	490	37	7.6
Linguistica Antverpiensia	11	146	27	18.5
Target	26	172	11	6.4
The Translator	25	152	6	3.9
Fachsprache	22	85	1	1.2
Trans-kom	17	113	13	11.5

Tabela 1 – Relação das publicações sobre tradução.

Analisadas por Christensen *et al.*, com o número de artigos dedicados à tecnologia.

Fonte: CHRISTENSEN *et. al.* (2017).

No levantamento, Paulsen *et al.* (2017) concluíram que os artigos relacionados com tecnologia da tradução corresponderam a 10,4% do total dos artigos publicados, e que a maioria das pesquisas foram as classificadas como pesquisas em tradução teórica, seguidas de pesquisas orientadas para tecnologia e pesquisas orientadas para o fluxo de trabalho. Entre as pesquisas em tradução teórica, Paulsen *et al.* (2017) descobriram que a maioria abordava temas como o impacto da tecnologia nos tradutores no seu treinamento e na profissão; nas

pesquisas orientadas para aspectos técnicos, a maioria se debruçava nas funções das ferramentas e na avaliação da qualidade; e entre as pesquisas orientadas para o processo e o lugar de trabalho, a maioria abordava a implementação da tecnologia no setor de serviços linguísticos, e poucas pesquisavam a tecnologia no lugar de trabalho e a interação entre os tradutores e as ferramentas e como isso afetava os processos e os profissionais.

A conclusão de Christensen *et al.* (2017) é que, apesar da presença ainda insuficiente, a pesquisa relacionada com tecnologia e tradução já faz parte dos ET devido ao forte impacto que a introdução da tecnologia provocou na profissão e à constatação de que estamos perante um processo de automatização acelerado.

Journal/Total numbers	1) Technology-oriented research papers	2) Workflow-oriented and industrial research papers	3) Translation-theoretical research papers
JosTrans	19	15	33
Hermes	5	0	1
Perspectives	7	7	10
Meta	14	6	17
Linguistica Antverpiensia	17	3	7
Target	2	2	7
The Translator	2	0	4
Fachsprache	1	0	0
Trans-kom	9	1	3
Total numbers	66	34	92

Tabela 2 – Análise temática dos artigos publicados.
Segundo a classificação de Christensen *et al.*
Fonte: CHRISTENSEN *et al.* (2017).

Ao analisar a ausência da tecnologia nos principais debates da ET, Quah (2006) sustenta que o mapa da disciplina de Holmes-Toury (Figura 7) confinou a tecnologia nos Estudos Aplicados, mais especificamente nas subáreas de Treinamento de Tradutores e, principalmente, em Ferramentas Auxiliares da Tradução. Partindo dessa premissa e considerando o peso que a tecnologia tem hoje na área, ele propõe a substituição do ramo de Assistência à Tradução (*Translation Aids*) por Tecnologia de Tradução (*Translation Technology*), que seria dividido por sua vez em Ferramentas de Tradução Automática e Ferramentas de Tradução Assistida por Computador. No ramo das Ferramentas de Tradução Assistidas por Computador, Quah inclui ferramentas linguísticas (dependentes e não dependentes da língua), ferramentas de tradução e ferramentas de localização. Além dessa reformulação, Quah propõe a inclusão, dentro dos estudos descritivos

do processo de tradução, da pré-edição e da pós-edição, ambas já descritas com mais detalhe na presente dissertação.

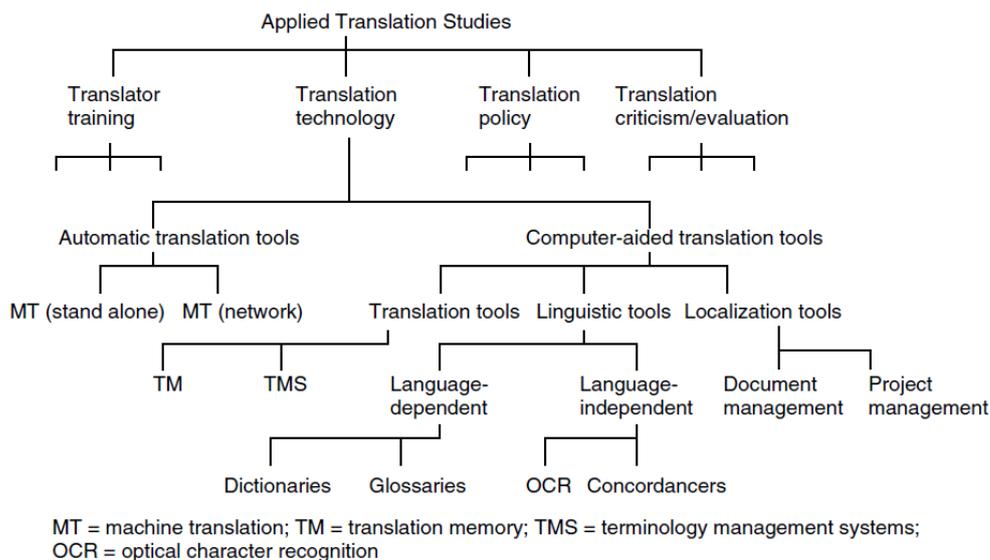


Figura 13 – Adaptação do mapa de Holmes-Toury.
Fonte: QUAH (2006).

A partir de uma perspectiva mais instrumental, Alcina (2008) enumera a seguinte classificação de subcategorias de tecnologia da tradução:

- a) equipamento informático (elementos para o funcionamento geral dos computadores, como sistemas operativos, software antivírus etc.);
- b) ferramentas e recursos de comunicação e documentação (correio eletrônico, redes virtuais e ambientes de colaboração, recursos para interagir com os clientes e colegas, para obter informação, como bases terminológicas online, websites etc.);
- c) edição de textos e editoração eletrônica (ferramentas para escrever, editar e revisar textos, especialmente processadores de texto);
- d) ferramentas e recursos linguísticos (para recopilar e organizar dados linguísticos, como corpora, bases de dados e dicionários online);
- e) ferramentas de tradução (projetadas especificamente para fazer parte do processo de tradução, estabelecendo relações entre um texto na língua de origem e o mesmo texto na língua de destino, geralmente no nível de segmento, e pode incluir outros elementos como bases terminológicas, como CAT Tools e sistemas de MT).

Com uma visão alinhada com Christensen *et al.* (2017), O'Hagan (2013) observa um elemento fundamental na relação entre tecnologia e os ET: a falta de

conexão entre as pesquisas que involucram o uso de tecnologia na tradução, como a utilização de ferramentas de tradução e corpus para o treinamento de tradutores e os estudos envolvendo PEMT, por exemplo, e a área de pesquisa pura, segundo o mapa de Holmes-Toury. Sobre as pesquisas motivadas pelo aumento da importância do alfabetismo informacional e da tecnologia na Era da Informação, O'Hagan afirma:

Enquanto sua contribuição é inegável para destacar características particulares de certa tecnologia ou de ambientes tecnológicos, este conhecimento geralmente não conduziu a teorizações ou modelado da tradução no 'ramo de pesquisa pura' dos estudos da tradução. Isto indica uma desconexão entre os ramos aplicado e puro no que respeita os temas orientados à tecnologia (O'HAGAN, 2013, p. 508).

Enquanto Holmes via a relação entre os estudos descritivos, teóricos e aplicados como uma relação dialética, com cada um dos ramos proporcionando material para os outros, isso não parece estar acontecendo especificamente com as pesquisas vinculadas a aspectos tecnológicos.

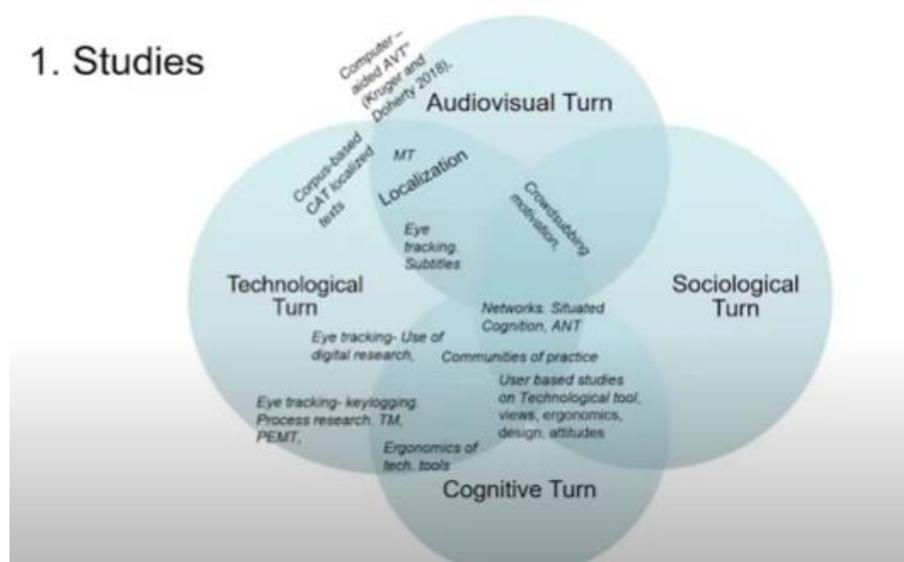


Figura 14 – Possíveis superposições entre a “virada tecnológica” e outras “viradas”, temáticas e metodológicas.

Fonte: TSC... (2019).

Em síntese, mesmo com as mudanças profundas que o uso intensivo da tecnologia está provocando no processo de tradução e sua forte influência, direta e indireta em muitas áreas dos ET, a discussão sobre a existência ou não de uma “virada tecnológica” não chegou a uma conclusão até o momento. Enquanto

alguns autores, como Alcina (2008), sustentam que a tecnologia da tradução seria uma nova disciplina acadêmica situada no meio do caminho entre as ciências da computação e os ET, outros defendem que deveria ser considerada um campo *per se* e sua presença crescente nos ET, tanto em desenvolvimentos teóricos como em pesquisas em diversas subdisciplinas deveria ser vista como uma “virada”. “Virada” ou simples mudança de perspectiva, tudo aponta a que a interrelação entre tradução e tecnologia e as interseções da tecnologia com outras “viradas” e metodologias de pesquisa (ver Figura 14) ocupará um lugar de importância crescente nos ET no futuro.

5

Pesquisas relacionadas

Considerando o crescimento da pós-edição e o uso cada vez mais habitual da MT por parte dos tradutores, a partir do ano 2000 houve um número importante de trabalhos relacionados com esses assuntos, embora os objetivos e abordagens apresentem diferenças substanciais. Levando em conta o objetivo desta dissertação, o presente capítulo faz um levantamento de trabalhos que envolvem PEMT com objetivos e metodologias diferentes. Os trabalhos foram classificados em três categorias: i) trabalhos que analisam a PEMT privilegiando aspectos referentes à qualidade e produtividade, fundamentalmente utilizando a metodologia de contagem, classificação e avaliação de erros; ii) trabalhos que analisam sistemas de TM + MT interativos; e, finalmente, iii) trabalhos que analisam e/ou comparam as características das traduções obtidas por PEMT e TH e as preferências dos usuários/leitores.

5.1

Pesquisas centradas na qualidade e produtividade

Uma parte importante das pesquisas sobre pós-edição debruçam-se sobre aspectos relacionados com qualidade e produtividade. Na maioria, a qualidade é medida por meio da detecção e contagem de erros, sua classificação segundo uma taxonomia determinada, e a atribuição de um peso a cada tipo de erro segundo a sua severidade. Essa abordagem apoia-se em modelos taxonômicos como o da extinta LISA (*Localization Industry Standards Association*), a TAUS (*Translation Automated Society*) e o Comitê Europeu de Normalização. Entre as categorias de erros mais utilizadas, podemos mencionar a precisão, a omissão, os termos incorretos, a pontuação etc. Por sua vez, a severidade costuma ser classificada

segundo grau, como menor, maior e crítica (ver Tabela 3). A produtividade é medida em quantidade de palavras pós-editadas ou traduzidas, segundo for o caso, em uma unidade de tempo determinada (habitualmente por hora).

Como exemplo desse tipo de trabalho podemos mencionar o estudo realizado por Arenas (2009), no qual nove tradutores profissionais trabalharam com um texto de 791 palavras, em que um conjunto de palavras eram segmentos com coincidências parciais (*fuzzy matches*) provenientes de uma TM, outro era produto de um sistema de MT que devia ser pós-editado e o restante devia ser traduzido com 100% de tradução humana. Em seguida, foi comparado o tempo que demandou cada grupo de segmentos e calculou-se a produtividade e a qualidade final de cada grupo. No que diz respeito à produtividade, a pesquisadora observou que o uso da TM e da MT proporcionou benefícios comparativamente com a tradução sem o uso desses recursos, sendo que a MT superou a TM. Para avaliar a qualidade, utilizou-se a taxonomia da LISA e a conclusão foi que o maior número de erros concentrou-se nos segmentos com coincidências com a TM, seguidos dos segmentos de MT pós-editados e dos segmentos novos (TH). Arenas (2009) também analisou os resultados considerando a experiência em pós-edição dos tradutores e concluiu que a experiência era um elemento importante com relação à produtividade.

METRICS	LISA QA Model	SAE J 2450	TAUS Benchmarking	QAT
ERROR TYPES		Miscellaneous		
	Accuracy	Omission	Accuracy	Omission
	Terminology	Wrong Term	Terminology	Terminology
	Language	Syntactic Punctuation Misspelling	Language	Grammar Punctuation Spelling
		Word structure or agreement error		
	Country		Country standards	
	Consistency		Consistency	Reference documents
	Style		Style	Clarity
	Mistranslation		Mistranslation	Sense
SEVERITY LEVELS	minor, major, critical	minor, major	minor, major, critical	minor, major

Tabela 3 – Quadro comparativo de tipos de erros e nível de gravidade de diferentes modelos de garantia de qualidade.

Fonte: MATEO (2014).

Em uma pesquisa posterior similar à anterior, Arenas (2014) fez um teste com 24 tradutores de inglês-espanhol e um sistema de tradução automática treinado (Moses). A conclusão foi que a produtividade aumentou com o uso dos segmentos proporcionados pela TM (*fuzzy matches*) e pela MT, e que a melhoria da produtividade variou entre os tradutores. Com relação à qualidade, a conclusão da Arenas (2014) foi que tanto a TM como a MT tiveram um impacto positivo, apresentando significativamente menos erros que a categoria de segmentos de TH.

Em outro trabalho com uma abordagem similar, Almeida (2013) analisou o trabalho de 18 tradutores dos pares inglês-francês e inglês-português que pós-editaram (foi solicitado uma pós-edição de nível intermediário) um total de aproximadamente 1000 palavras traduzidas por um sistema de MT para analisar o desempenho dos tradutores durante a pós-edição. O estudo mediu o número de palavras/hora pós-editadas além de se debruçar sobre as intervenções dos tradutores, que foram classificadas segundo uma taxonomia própria. No nível mais alto, o modelo classificou as intervenções como Essenciais, Preferenciais, Essenciais não implementadas e Erros introduzidos. A pesquisadora concluiu que, em ambos os pares de idiomas, o resultado da classificação nessas categorias principais foi, em ordem decrescente, as Alterações essenciais, seguidas das Preferenciais, as Essenciais não implementadas e, finalmente, os Erros introduzidos. Em um nível inferior, Almeida (2013) considerou oito categorias principais – onde se incluíam precisão, uniformidade e formato, entre outras – e 21 subcategorias – onde se incluíam texto não traduzido, informação faltante ou extra na MT, entre outras. No referente à relação entre experiência, tempo, qualidade e opinião sobre MT, embora o estudo indicou que as características pessoais dos tradutores tinham relevância, não foi possível obter correlações significativas entre esses fatores.

5.2

Utilização de MT no nível de subsegmento e tradução interativa

Nessa classificação incluímos trabalhos que avaliam o uso de MT no nível de subsegmento e tradução interativa, embora alguns deles se debruçem em

estabelecer comparações de qualidade e produtividade, como os estudos incluídos no ponto anterior.

Uma pesquisa realizada por Koehn (2009) explorou o uso de sugestões provenientes de um sistema de MT para o completamento de frases. Para isso, desenvolveu uma ferramenta própria chamada Caitra, similar às ferramentas TransType e Lilt, descritas em 3.3 Avanços em MT e interatividade. Um total de dez tradutores não profissionais nativos dos idiomas inglês ou francês traduziram 192 frases de notícias do francês para o inglês utilizando o sistema Caitra. A interação entre os tradutores e os diversos recursos que a ferramenta oferece foi registrada para analisar quais recursos foram utilizados e quanto tempo foi dedicado a atividades como clicar no mouse, digitar ou pausar o trabalho. O principal interesse do estudo experimental centrou-se na velocidade de tradução. A ferramenta Caitra oferece três tipos de assistência: predição, opções de uma tabela de traduções e pós-edição de MT no nível de segmento (frase completa). No primeiro caso, a MT é utilizada como um auxílio para o tradutor completar as frases. Como afirma Koehn, nessa modalidade de uso de MT o tradutor humano permanece no controle da tradução, mas recebe a ajuda de um sistema de MT que automaticamente faz sugestões para completar a frase. O sistema vai atualizando as sugestões de maneira dinâmica a partir das soluções preferidas pelo tradutor. O segundo tipo de assistência é uma tabela de traduções obtidas por meio de MT estatística baseada em frases na qual o tradutor pode escolher a opção que achar mais adequada e inseri-la com um clique. Finalmente, o terceiro tipo de assistência permite que o tradutor escolha introduzir a frase completa pré-traduzida por MT para, em seguida, pós-editá-la. Além de analisar o processo, as traduções, acompanhadas dos textos na língua original, foram submetidas à avaliação de um conjunto juízes para determinar qual representava melhor o original em termos de fluência y equivalência de significado. O estudo concluiu que os tradutores eram mais rápidos (alguns tradutores duplicaram a velocidade) e produziam melhores traduções quando recebiam assistência da ferramenta. No referente às preferências de trabalho dos tradutores, determinada a partir de uma enquete, a classificação foi a seguinte: sem auxílio da ferramenta (1 tradutor), pós-edição (1 tradutor), predições (1 tradutor), opções (2 tradutores) e predições + opções (3 tradutores).

Um estudo feito por Green *et al.* (2014) explorou o que chamaram de Memória de Tradução Preditiva (*Predictive Translation Memory, PTM*), um sistema misto que permite que os tradutores produzam as suas traduções usando sugestões proporcionadas por um sistema de MT, descrito com mais detalhe em 3.3 Avanços em MT e interatividade. O estudo comparou o uso de PEMT com MT interativa/PTM em três textos de gêneros diferentes, software, conteúdo médico e notícias informais, e em dois pares de idiomas, francês-inglês e inglês-alemão. Os tradutores participantes eram profissionais recrutados em uma conhecida plataforma para tradutores freelance (Proz) e foram pagos pela tarefa. Todos eles tinham experiência em pós-edição. A análise comparou ambas as formas de trabalho em termos de tempo e qualidade. A qualidade foi medida mediante o uso do modelo BLEU + 1, uma variante do método de avaliação de tradução automática BLEU (*Bilingual Evaluation Understudy*), que qualifica as traduções obtidas por MT em função da sua semelhança com traduções humanas de referência de boa qualidade. A conclusão do estudo foi que os tradutores participantes que utilizaram PTM ficaram próximos dos que fizeram PEMT em termos de velocidade e os superaram em termos de qualidade, avaliada automaticamente. Uma enquete demonstrou também que os tradutores gostaram de utilizar os recursos de ajuda interativos, e vários deles afirmaram que acreditavam que, com mais prática, conseguiriam aumentar a produtividade.

5.3

Características da TH e da PEMT e preferências dos leitores

Nesta classificação incluímos trabalhos que observaram aspectos diferentes, como as preferências dos leitores com relação aos textos traduzidos utilizando TH e PEMT. Fiederer e O'Brien (2009) fizeram um estudo com 30 frases obtidas por TH e PEMT para verificar as preferências de 15 avaliadores, todos eles falantes nativos da língua de destino e mestres ou estudantes de mestrado de Estudos da Tradução da Dublin City University (DCU). A conclusão das autoras foi que, em termos de clareza e precisão, o texto obtido por PEMT apresentou uma qualidade similar, senão superior, à tradução humana. Porém, os avaliadores indicaram uma

clara preferência pelas frases obtidas por TH sobre as obtidas por PEMT (63% a 37%), o que indicaria que a preferência estaria mais vinculada ao estilo do que com a clareza e a precisão.

Outra pesquisa no mesmo sentido foi feita por Bowker e Buitrago (2015) para verificar a utilidade que a MT teria para imigrantes falantes de espanhol no Canadá. Participaram de pesquisa 114 imigrantes recentes que compararam textos traduzidos por MT, TH e PEMT. O resultado foi que 49% dos participantes preferiram textos pós-editados (24% *full post-editing* e 25% *light post-editing*) contra 42% que preferiram TH e 9% que preferiam MT sem edição.

Daems *et al.* (2017) procuraram determinar se os seres humanos seriam capazes de identificar traduções humanas (TH) de traduções obtidas por PEMT (*full post-editing*), e se era possível identificar diferenças objetivas e quantificáveis entre textos obtidos por TH e por PEMT. A partir de traduções obtidas por meio de TH e PEMT, criaram modelos computacionais alimentados com as diferenças que conseguiram observar para verificar também se o computador era capaz de identificar a TH da PEMT. A primeira conclusão foi que os participantes não conseguiram distinguir entre a TH e a PEMT (embora os autores assinalaram que uma explicação possível poderia ser o par de línguas escolhido para o teste, no caso inglês e holandês). Inclusive, foi mais comum os leitores considerarem como TH os textos obtidos por PEMT que o contrário. Apesar de, inicialmente, a análise computacional permitir achar combinações de características que poderiam ser capazes de diferenciar a TH da PEMT, o resultado obtido aplicando os modelos computacionais desenvolvidos em outro conjunto de dados chegou à mesma conclusão anterior, isto é, que o computador também não era capaz de distinguir com precisão a TH da PEMT. Portanto, os autores determinaram que o “pós-editês” (*post-editese*), considerado como texto com características diferentes, não existia ou, em outras palavras, que os textos obtidos por TH são indistinguíveis dos textos obtidos por PEMT.

Farrell (2018) tentou verificar a existência de marcadores de MT nos textos obtidos por PEMT. Trabalhando com 50 alunos de pós-graduação de tradução (26 de fazendo PEMT e 24 TH), ele comparou textos da Wikipedia em inglês traduzidos para o italiano, obtidos por TH e PEMT no referente ao tempo, os erros e a existência de marcadores de MT. No seu estudo, o autor denomina marcadores

de MT os n-gramas¹¹ identificados em uma verificação manual simples por terem sido traduzidos de maneira mais diversa nos textos de TH que nos textos de PEMT (por exemplo, foram identificados como marcadores de MT os n-gramas *by some*, *numerous attractions*, *mainly* e *destination*, entre outros). Em termos de rapidez, o estudo constatou uma diferença favorável à PEMT, embora não particularmente significativa. Com relação aos erros, a diferença também não foi significativa, motivo pelo qual o estudo concluiu que as duas modalidades de trabalho eram comparáveis nesse quesito. Com relação aos marcadores de MT, Farrell observou que eles estavam mais presentes nos textos de PEMT que nos de TH, o que faria com que a PEMT fosse mais normalizada, homogênea e repetitiva. Mesmo assinalando que a variedade e a inventividade não são características desejáveis sempre (por exemplo, em textos extremamente técnicos, como manuais, pode ser até um defeito), em muitos tipos de textos (marketing, publicidade, jornalismo, entretenimento etc.) a uniformidade lexical, que foi constatada em todos os textos obtidos por PEMT, pode fazer com que o texto seja “menos interessante de ler e menos estimulante intelectualmente”. Em tal sentido, Farrell (2018) concluir que a contagem e classificação de erros e inclusive a medição de fluência e adequação podem não ser suficientes na hora de avaliar a qualidade de uma tradução, e isso explicaria o motivo de estudos como o de Fiederer e O’Brien (2009), acima mencionado terem observado que a TH é melhor em termos de estilo. Farrell também indica que o problema original não seria da PEMT, mas da MT que serve como base para a ela, e que esse efeito poderia ser minimizado programando o sistema de tradução automática para aumentar a aleatoriedade das escolhas.

O trabalho de Toral (2019), *Post-editeuse: na Exacerbated Translationese*, que forneceu insumos para o estudo experimental da presente dissertação, também teve como objetivo verificar a existência de características diferenciadoras nos textos obtidos por TH e PEMT. Para analisar o conteúdo traduzido por TH e PEMT ele recorreu aos “universais da tradução”, dos quais falaremos a seguir e retomaremos na seção de metodologia do estudo experimental.

Em um capítulo do livro *Text and Technology: In Honour of John Sinclair* (1993), Mona Baker sustentou a ideia de que vários pesquisadores perceberam

¹¹ Um n-grama é uma sequência contígua de n itens de uma amostra de texto ou fala.

intuitivamente algumas características dos textos traduzidos que seriam resultado do próprio processo de tradução, independentemente das línguas envolvidas. Essas características ficariam conhecidas dentro dos estudos descritivos como “universais da tradução”. Segundo Laviosa:

Os universais da tradução são características linguísticas que estão presentes tipicamente em textos traduzidos, em lugar de em textos originais, e se considera que são independentes da influência dos pares de línguas específicos envolvidos no processo de tradução. (LAVIOSA, 1998, p. 288).

Baker identificou cinco “universais”. O primeiro é uma elevação evidente no nível de explicitude com relação, tanto ao texto fonte como a textos elaborados na língua de destino. Baker definiu a explicitação como a tendência geral a explicar as coisas em detalhe em lugar de deixá-las implícitas na tradução (BAKER, 1993, p. 180). O segundo é uma tendência à desambiguação e a simplificação, isto é, a simplificar a linguagem utilizada na tradução. Por exemplo, simplificar a sintaxe mais complexa, colocar aspas e outros signos de pontuação quando o texto original está duvidoso etc. O terceiro é a normalização, que se manifesta como uma forte tendência a exagerar características típicas do idioma de destino, que se manifestaria na preferência pela gramaticalidade convencional, como finalizar frases incompletas, corrigir expressões que se desviam das regras da gramática, omitir autocorreções, mesmo que sejam propositais etc. O quarto universal proposto por Baker é uma tendência a evitar as repetições presentes nos textos originais, seja omitindo elas o parafraseando-as. E o quinto é uma tendência geral a exagerar as características das línguas de destino. Para, Baker essas características distintivas dos textos traduzidos seriam resultado da confrontação entre a língua de origem e a língua alvo.

Em uma obra mais recente, *Translation Universals: Do they exist?* (MAURANEN; KUJAMAKI, 2004), Toury destaca a tensão que existe entre a procura de regularidades e padrões recorrentes nas ciências e a falta de interesse – ou até mesmo a hostilidade – que esse tipo de abordagem gera em quem valoriza, antes de tudo, a diferença e exclusividade. Além disso, ele destaca que antes que nos perguntarmos a respeito da existência ou não dos “universais”, deveríamos pensar se esse recurso permitiria obter novo conhecimento, no sentido de aumentar a capacidade de a teoria da tradução descrever e explicar os fenômenos

individuais, o que estaria no cerne dos ET como disciplina empírica. Na mesma obra, Chesterman afirma que os “universais” podem ser classificados em dois grupos: os Universais T (de *target* ou língua destino), que marcariam diferenças entre traduções e textos similares que não tenham sido traduzidos; e os Universais S (de *source* ou língua fonte), que seriam características resultantes da maneira como os tradutores processam o texto original. Entre os potenciais *Universais S*, Chesterman menciona, entre outros, o tamanho das frases (tendência a que as traduções sejam mais extensas que seus textos originais; a lei da interferência e a lei da estandardização (ambas descritas por Toury); e a redução da repetição. Entre os potenciais *Universais T*, ele menciona a simplificação (menor densidade e variedade lexicais) e a normalização (descrita por Baker), entre outras. Com relação aos problemas, Chesterman cita, entre outros, o da representatividade das amostras analisadas e a própria ideia de universalidade, motivo pelo qual afirma que qualquer ideia sobre “universais” deve ser considerada como uma aproximação; a operacionalização do uso dessas noções abstratas; e a causalidade, isto é, a capacidade de atribuir essas características ao fato de os textos serem traduções e não a outros fatores.

Voltando ao trabalho de Toral (2019), utilizando o conceito de “universais da tradução”, o autor considerou que os textos traduzidos teriam características que os diferenciariam dos textos redigidos originalmente em uma língua determinada: daí o uso dos termos *Translationese* (“Tradutês”), para denominar os textos traduzidos, e *Post-editeese* (“Pós-editês”), para denominar os textos traduzidos usando PEMT. Toral aplicou uma série de análises computacionais a diversos conjuntos de dados que continham TH e PEMT, com diversos idiomas e direções (inglês < > alemão; inglês < > francês, espanhol < > alemão e chinês < > inglês) para tentar determinar a existência do “pós-editês”, que como é sugerido pelo título do artigo seria um “tradutês” com as suas características exacerbadas. Os textos obtidos por PEMT incluíam tanto *full* como *light post-editing*, e foram pós-editados tanto por tradutores profissionais como por amadores. Da mesma forma, os textos de TH também foram obtidos a partir do trabalho de tradutores profissionais e não profissionais, com diversos graus de competência. Os sistemas de MT utilizados também foram diversos: incluíram tradução automática baseada em frases (PBMT) e tradução automática neural (NMT). Por meio da análise da variedade lexical, a densidade lexical, a razão de comprimento das frases e as

sequências de PoS, Toral (2019) concluiu que a PEMT tem menor variedade e densidade lexicais que a TH, o que foi relacionado como o “universal” de simplificação. O “pos-editês” (PEMT), portanto, seria lexicalmente mais simples que o “tradutês” (TH). Com relação ao comprimento das frases, o estudo determinou que o comprimento na PEMT era mais próximo do comprimento do texto fonte que na TH, o que foi relacionado com a interferência e a normalização. No referente à PoS, o estudo de Toral (2019) determinou que as sequências de PoS eram mais similares às sequências de PoS do original do que as sequências da TH. Segundo Toral (2019), mesmo a PEMT sendo útil em termos de produtividade, a tendência da adoção ampla da PEMT poderia ter consequências negativas no longo prazo, com um empobrecimento da língua e uma interferência do idioma fonte. Da mesma forma que Farrel (2018), Toral (2019) também destacou que a responsabilidade não pode ser atribuída, na sua totalidade, à PEMT, mas que ela origina-se nos sistemas de MT utilizados como ponto de partida para a pós-edição.

6

O estudo experimental

Neste capítulo apresentaremos o estudo experimental realizado, descrevendo os objetivos e a metodologia utilizada.

6.1

Apresentação

Como foi explicado anteriormente, seja por opção pessoal ou por exigência do mercado, o uso combinado de tradução automática (MT) e memória de tradução (TM) dentro do ambiente de CAT Tool é uma modalidade de trabalho cada vez mais habitual entre os tradutores profissionais especializados. Porém, a pós-edição de textos traduzidos por MT não é a única maneira de utilizar os sistemas de tradução automática. Mesmo sem que seja uma exigência de agências de tradução ou de clientes finais, cada vez mais profissionais utilizam os sistemas de MT, já seja para pré-traduzir o conteúdo, para obter uma compreensão geral de uma frase (*gist translation*) ou para receber subsídios que os auxiliem na tradução. Cada vez mais CAT Tools permitem utilizar a MT durante o processo de tradução no nível de subsegmento, isto é, oferecendo fragmentos de diversos tamanhos traduzidos por MT que o tradutor pode ou não incorporar, para acelerar o trabalho, mas mantendo sua própria iniciativa na organização das frases e aumentando o controle do processo. Essa maneira de trabalhar é menos utilizada pelos tradutores e tem sido menos explorada nas diversas, apesar de oferecer algumas vantagens, especialmente quando observamos as reclamações mais habituais dos tradutores que trabalham na modalidade PEMT, como a afirmação de que as opções oferecidas pela MT são de qualidade insuficiente, são próximas demais (“calcadas”) do texto de origem, ou de que, em certa maneira, a MT limita as

opções do tradutor, que encontra dificuldades para fugir das soluções propostas por ela inicialmente (BUNDGAARD, 2017).

Para realizar a comparação entre os textos traduzidos por PEMT e MTNSS utilizamos algumas das características denominadas “universais da tradução” (BAKER, 1993), mencionadas no ponto anterior. Mais especificamente, simplificação (tendência à desambiguação e simplificação, o que implica menor variedade e densidade lexical), normalização ou homogeneização (tendência a se ajustar às características típicas da língua de destino) e interferência (influência da língua de origem na língua de destino em uma situação de contato que gera algum tipo de desvio). No caso, os “universais” foram utilizados exclusivamente de maneira instrumental e não como uma métrica de qualidade.

6.2

Objetivos

Os objetivos e as perguntas que orientaram o estudo experimental foram apresentados oportunamente no capítulo 1, mas serão retomados a seguir para facilitar a leitura.

Objetivos:

- a) avaliar comparativamente os resultados obtidos utilizando duas modalidades de trabalho nas quais se integra a tecnologia de TM com MT, a) a pós-edição tradicional (PEMT), em que o texto é pré-traduzido inteiramente por MT e depois pós-editado; e b) o uso de MT no nível de subsegmento (MTNSS), isto é, incorporando fragmentos fornecidos pela ferramenta CAT oriundos de um sistema de MT durante a tradução. Observar se as duas maneiras de trabalhar impactam de forma diferenciada em gêneros textuais diversos;
- b) comparar os resultados obtidos por meio de PEMT e MTNSS com a tradução humana (TH) por meio de uma série de análises computacionais;
- c) a avaliar as preferências de um grupo de leitores com relação às traduções obtidas por meio das modalidades de trabalho propostas.

Perguntas orientadoras:

- a) em que medida o trabalho em um ambiente de TM + MT oferece resultados diferentes usando a MT no nível de segmento (PEMT) e no nível de subsegmento (MTNSS)?
- b) em que consistem essas diferenças, caso existam?
- c) o uso de MT no nível de subsegmento é mais adequado para certos gêneros de texto do que para outros?
- d) as duas maneiras de trabalhar com a MT, PEMT e MTNSS, permitem observar diferenças em termos de produtividade (palavras/hora)?
- e) os leitores avaliadores têm preferência pelos textos traduzidos com alguma das modalidades mencionadas (PEMT, MTNSS e TH)?

6.3

Metodologia

6.3.1

Os tradutores e tradutoras participantes

No estudo experimental participaram dois tradutores e duas tradutoras profissionais especializados que atuam na área de tecnologia da informação (TI). Três deles trabalham em tempo integral para a mesma agência de tradução, de maneira remota, e o restante presta serviços para outras agências e clientes que trabalham com conteúdo similar.

Todos os tradutores e tradutoras participantes tinham como língua nativa o espanhol latino-americano. No momento do estúdio experimental eles tinham entre oito e 15 anos de experiência e, a exceção de um, todos trabalhavam de maneira habitual com PEMT e com textos similares aos incluídos no estudo experimental. O único profissional que não trabalhava habitualmente com PEMT foi aquele que fez a tradução humana (TH).

Antes da realização do experimento, houve um contato prévio com os profissionais para explicar que participariam de um estudo experimental no

contexto de uma dissertação de mestrado e para conversar brevemente sobre a tarefa. Também lhes foi comunicado que o trabalho seria pago e aceitaram o valor por palavra oferecido.

6.3.2

O ambiente de TM + MT

A ferramenta CAT escolhida para a realização do estudo foi o SDL Trados Studio (versões 2015 e 2019), que era a ferramenta principal de trabalho dos profissionais participantes, além de ser uma das ferramentas mais utilizadas na indústria.¹² A ideia foi tentar reproduzir condições similares às do trabalho cotidiano.

A MT utilizada foi o Google Translate (GT), que foi escolhida por ser um sistema muito conhecido e amplamente usado pelos tradutores autônomos e empresas.¹³ A versão utilizada foi *Cloud Translation Basic* (versão paga), com o modelo de NMT, e o acesso (tanto para a pré-tradução dos arquivos antes do envio como para a MTNSS) foi feito por meio de uma API (ver 3.4 As ferramentas CAT como um caso de interação bem-sucedido) incorporada de forma predeterminada no SDL Trados Studio.

¹² Segundo a enquete European Language Industry Survey, o SDL Trados Studio 2019 foi considerado como a ferramenta CAT de maior impacto nas empresas em 2019. Cf.: https://ec.europa.eu/info/sites/default/files/2020_language_industry_survey_report.pdf.

¹³ Segundo a pesquisa da indústria de serviços linguísticos 2020 da Association of Language Companies 2020, 56% das empresas usam Google e Microsoft como sistemas de MT. Cf.: <https://www.alcus.org/page/2020IndustrySurvey>.

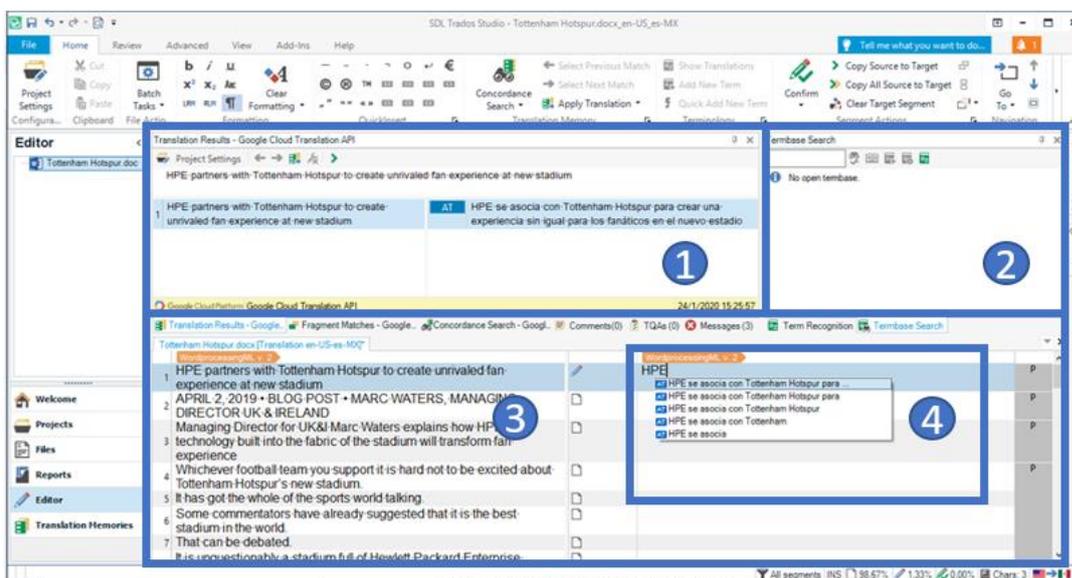


Figura 15 – Captura de tela do SDL Trados Studio 2019 - MTNSS.

Legendas: 1. Janela da memória de tradução (TM); 2. Janela do glossário; 3. Janela de edição; 4. Opções de fragmentos traduzidos por MT.

Fonte: Elaboração própria (2021).

No caso da PEMT, foi feita a pré-tradução dos arquivos no ambiente do SDL Trados Studio antes da inclusão dos arquivos bilíngues no pacote de trabalho. No caso da MTNSS, o sistema de MT foi usado como AutoSuggest, isto é, o tradutor foi escolhendo e incorporando fragmentos de soluções de MT oferecidas pelo GT ao longo da tradução. Na Figura 13 é possível ver a maneira como as opções de MT, desde frases completas a só um par de palavras, apresentam-se para o tradutor. Essas opções mudam de maneira dinâmica na medida em que o tradutor avança na frase. O tradutor pode escolher a opção que desejar usando a tecla de seta acima/abaixo e, uma vez selecionada, incorporar o fragmento preferido simplesmente pressionando a tecla Enter.

6.3.3

O material enviado para os tradutores e tradutoras participantes

Para cada um dos textos traduzidos, foi enviado um pacote preparado para ser aberto na própria ferramenta CAT (sdlppx). O pacote, que é um arquivo

comprimido com um formato específico para ser aberto no SDL Trados Studio, continha o seguinte material:

- O texto a ser traduzido (formato .sdlxliff¹⁴, pré-traduzido para os profissionais que fizeram PEMT e sem traduzir para os que fizeram MTNSS e TH);
- Uma memória de tradução (formato .sdltm) com 132.762 unidades de tradução (TU's), alimentada com textos da área de tecnologia;
- Um glossário (formato .sdltb) com 777 termos da área de tecnologia.

Além do pacote de tradução, foi enviado um pequeno tutorial sobre a instalação e o uso do programa VLC para gravar a tela durante a realização da tarefa encomendada (ver Apêndice 1). A gravação foi usada exclusivamente para controle do tempo dispendido com a finalidade de calcular a produtividade.

Junto com esse material foi enviada uma ficha que explicava brevemente o projeto, similar às ordens de compra (PO) geradas por algumas agências de tradução (ver Apêndices 2, 3, 4 e 5). A ficha indicava o tipo de trabalho esperado: pós-edição com qualidade de tradução humana e tradução só com uso de TM (sem MT). Para a modalidade de MTNSS, onde a tradução foi feita pelo autor da dissertação, não houve necessidade de ficha.

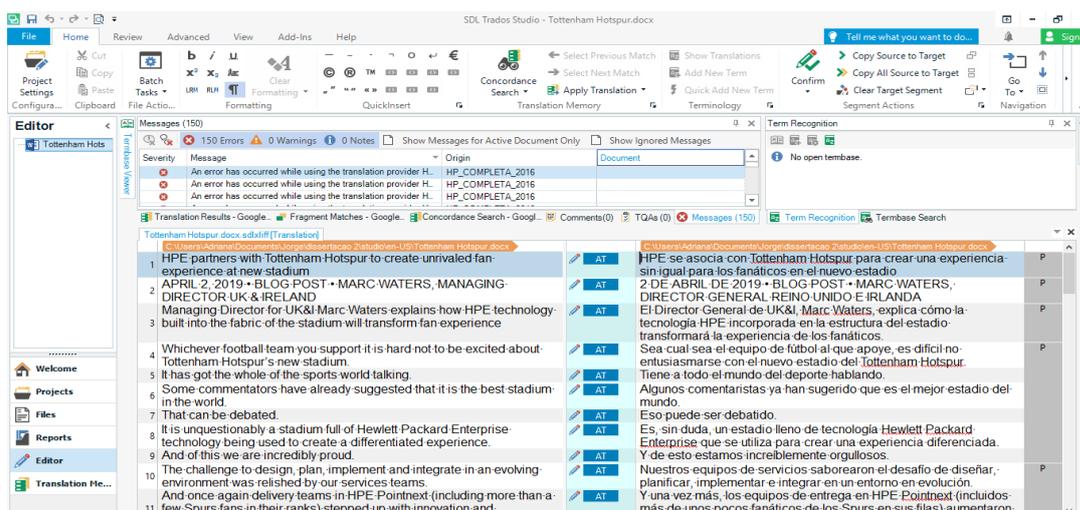


Figura 16 – Captura de tela do SDL Trados Studio 2019 - sdlxliff. Arquivo sdlxliff pré-traduzido por MT recebido para fazer a PEMT. Fonte: captura de tela própria.

¹⁴ Sdlxliff (SDL XML-based Localization Interchange File Format) é o formato dos arquivos bilíngues do SDL Trados Studio. São arquivos bicolunados onde na coluna da esquerda se apresenta o texto para traduzir e na coluna da direita o tradutor introduz a sua tradução. No caso dos tradutores e tradutoras que trabalharam com PEMT no estudo experimental, eles receberam os arquivos pré-traduzidos pelo GT.

6.3.4

Os textos escolhidos

Os dois textos originais selecionados para o estudo experimental foram obtidos da internet e não estavam traduzidos para a língua de destino até o momento da realização do teste. Eles tinham um total de 897 e 976 palavras. Um dos arquivos tinha um aproveitamento de memória de 37 palavras 100% *match*, isto é, cuja tradução já estava armazenada na TM, e 24 palavras de *fuzzy match* (correspondência parcial com a TM). O outro não tinha nenhum aproveitamento. Essa informação pode ser observada com mais detalhe nos arquivos de registro resultantes da análise realizada utilizando o SDL Trados Studio (ver Figuras 15 e 16).

Os textos foram redigidos originalmente em inglês dos EUA e o idioma de destino solicitado foi o espanhol latino-americano¹⁵, línguas com as quais os tradutores e tradutoras participantes do estudo trabalham com mais frequência.

O texto T1 é uma postagem sobre marketing de tecnologia em um blog de uma grande empresa de tecnologia, que descreve uma reforma feita para a inclusão de diversos recursos tecnológicos no estádio de um conhecido time europeu de futebol. O público-alvo eram clientes atuais e potenciais da empresa, mas também o público em geral interessado em tecnologia. O estilo de redação tem pontos de contato com um texto jornalístico, com uma estrutura dinâmica e de leitura fácil. No caso desse tipo de textos, a expectativa dos clientes costuma ser que a tradução seja mais criativa e “solta”, com um grau maior de adaptação e menos sujeita formalmente ao texto original.

¹⁵ Se bem o espanhol latino-americano não existe como variante linguística, a denominação é utilizada pelas agências de tradução para se referir a um espanhol o mais neutro possível destinado ao público de América latina.

Totals											
Total	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags	
Files:1	PerfectMatch	0	0	0	0.00%	0	0	0		0	
Chars/Word:5.13	Context Match	0	0	0	0.00%	0	0	0		0	
	Repetitions	0	0	0	0.00%	0	0	0		0	
	Cross-file Repetitions	0	0	0	0.00%	0	0	0		0	
	100%	0	0	0	0.00%	0	0	0		0	
	95% - 99%	0	0	0	0.00%	0	0	0		0	
	85% - 94%	0	0	0	0.00%	0	0	0		0	
	75% - 84%	0	0	0	0.00%	0	0	0		0	
	50% - 74%	0	0	0	0.00%	0	0	0		0	
	New/AT	54	976	5011	100.00%	16	34	0		0	
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0		0	
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0	
	Total		54	976	5011	100%	16	34	0	0.00%	0

Figura 17 – Arquivo de análise do SDL Trados Studio (T1).

Fonte: Captura de tela própria (2021).

File Details											
File	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags	
Insight_Online_FAQ.docx.sdlxliff	PerfectMatch	0	0	0	0.00%	0	0	0		0	
Chars/Word:5.69	Context Match	0	0	0	0.00%	0	0	0		0	
	Repetitions	0	0	0	0.00%	0	0	0		0	
	Cross-file Repetitions	0	0	0	0.00%	0	0	0		0	
	100%	15	37	217	4.12%	0	0	0		0	
	95% - 99%	8	8	128	0.89%	11	0	0		8	
	85% - 94%	1	1	3	0.11%	1	0	0		0	
	75% - 84%	2	5	32	0.56%	0	2	0		0	
	50% - 74%	3	10	58	1.11%	6	2	0		5	
	New/AT	69	836	4666	93.20%	98	142	0		43	
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0		0	
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0	
	Total		98	897	5104	100%	116	146	0	0.00%	56

Figura 18 – Arquivo de análise do SDL Trados Studio (T2).

Fonte: Captura de tela própria (2021).

O texto T2 é um texto altamente técnico de caráter estritamente informativo. Ele fornece aos clientes de uma grande empresa prestadora de serviços respostas às perguntas mais frequentes sobre uma solução automatizada para supervisionar o funcionamento de sistemas informáticos e dispositivos empresariais, de maneira remota. A estrutura é de perguntas e respostas, redigidas em frases curtas e objetivas. Na área técnica, esse tipo de texto é conhecido como FAQ (*Frequently Asked Questions*), que costuma ser traduzido como perguntas mais frequentes).

A ideia de utilizar dois textos de tipologias diferentes, embora ambos poderiam considerar-se textos técnicos, foi tentar verificar se o impacto dos métodos de tradução comparados, a PEMT e a MTNSS, seria diferente em função do tipo de texto traduzido. Em tal sentido, a hipótese inicial era que o texto mais técnico apresentaria menos diferenças que o texto mais jornalístico, uma vez que suas características limitam tanto as soluções mais criativas como o uso de terminologia mais diversa. De fato, uma das características necessárias de um texto técnico é a uniformidade terminológica.

6.3.5

Diversidade lexical

A diversidade lexical é uma das medidas geralmente utilizadas para avaliar a riqueza léxica de um texto. Diversos autores consideram a diversidade lexical como sinônimo de riqueza de vocabulário (MALVERN *et al.*, 2004). A diversidade lexical refere-se ao número de palavras diferentes utilizadas em um texto com relação ao número total de palavras. A maneira habitual de medir a diversidade lexical é a razão *type-token* (*type-token ratio*, ou TTR), onde *token* significa o número total de palavras que contém o texto e *tipo* (*type*) identifica as palavras diferentes. Isso pode ser expresso na seguinte fórmula:

$$TTR = \frac{\text{Número de tipos de palavras}}{\text{Número de tokens}}$$

Em poucas palavras, a TTR é uma quantificação da diversidade de palavras em um texto, isto é, quanto maior a variação, maior a diversidade. Assim, no nosso caso, podemos dizer que uma TTR elevada significa que o tradutor usa um vocabulário mais amplo, e uma TTR baixa implica que o tradutor utiliza um vocabulário mais limitado. Ao trabalhar com grandes corpora, se considera mais confiável fazer um cálculo da média da TTR dividindo o texto em conjuntos consecutivos de 1000 palavras, calculando a TTR de cada um de eles, e obtendo a média (BAKER, 2000). No caso dos textos aqui analisados, essa operação não foi necessária considerando que os textos têm um total aproximado de 1000 palavras.

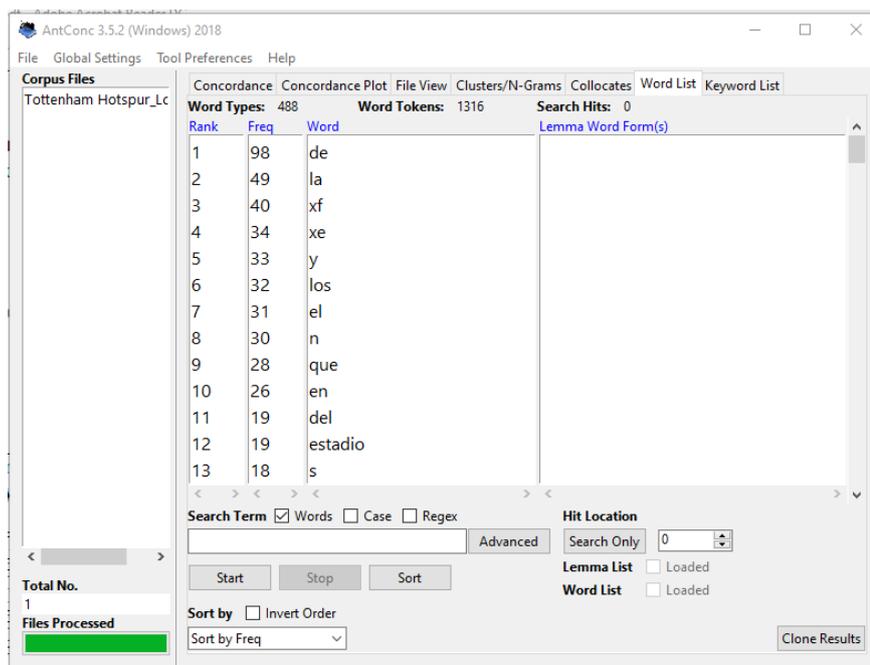


Figura 19 – Tela da aba Word List do concordanciador AntConc. Onde aparecem o número de tipos de palavras e o número de tokens. Fonte: captura de tela própria.

A determinação da TTR foi feita utilizando o AntConc, um software free-ware para análise de corpus desenvolvido por Laurence Anthony que permite aplicar as principais técnicas de linguística de corpus.¹⁶ Ao carregar os arquivos em formato de bloco de notas (txt), na aba *Word list* (lista de palavras) do programa é possível obter estatísticas de tipos de palavras (*Word Types*) e número total de palavras (*Word tokens*), como é possível observar na Figura 11.

Considerando que a MT tende a proporcionar um subconjunto de todas as traduções possíveis, o que se deve aos dados usados com mais frequência no treinamento (FARREL, 2018; TORAL, 2019), a nossa hipótese foi que quanto maior fosse a influência da MT, menor deveria ser a TTR, isto é, que a tradução utilizando PEMT deveria ter uma TTR menor que a da MTNSS, onde o uso da MT é menos intensivo.

¹⁶ Cf.: <https://www.laurenceanthony.net/software/antconc/>.

6.3.6

Densidade lexical

A densidade lexical é outro indicador relacionado à riqueza lexical. Ela assinala a quantidade de informação presente em um determinado texto ou a maneira como a informação é incorporada na estrutura gramatical. Assim, um texto mais denso do ponto de vista lexical seria um texto mais informativo (RAVID, 2006). A sua determinação pode ser realizada através de diversos métodos, um dos quais consiste em determinar a razão entre os itens lexicais das chamadas classes abertas (nomes, advérbios, adjetivos e verbos), também denominadas palavras de conteúdo, e o total de palavras do texto (tokens). Isso é feito por meio da seguinte fórmula:

$$\text{Densidade lexical} = \frac{\text{Número de palavras de conteúdo}}{\text{Número total de palavras}}$$

A classificação das palavras nos arquivos traduzidos que permitiu a contagem para efeitos de comparação foi feita utilizando UDPipe¹⁷, uma ferramenta para análise, lematização e etiquetagem gramatical (*Part of Speech –PoS– tagging*) do conteúdo de um texto em formato CoNLL-U, que usa um conjunto de etiquetas de classes de palavras que seriam aplicáveis a várias línguas. A ferramenta oferece modelos pré-treinados para 65 idiomas, inclusive espanhol e inglês, as línguas envolvidas no presente estudo. A anotação dos modelos pré-treinados é feita conforme a abordagem do projeto *Universal Dependencies* (UD)¹⁸, que tem como objetivo a criação de anotações linguísticas consistentes entre idiomas diferentes. Em tal sentido, vale a pena frisar que a taxa de sucesso do analisador morfológico para todas as Dependências Universais é de 99,7 para inglês e 99,7 para espanhol (STRAKA *et al.*, 2017).

¹⁷ Cf.: <https://cran.r-project.org/web/packages/udpipe/vignettes/udpipe-annotation.html#udpipe---general>.

¹⁸ Cf.: <https://universaldependencies.org/introduction.html>.

	Las alianzas crean innovación.								
	Las	el	DET	DET	Definite=Def	2 det	-	-	
1	alianzas	alianzas	NOUN	NOUN	Gender=Fem	3 nsubj	-	-	
2	crean	crear	VERB	VERB	Mood=Ind N	0 root	-	-	
3	innovación	innovación	NOUN	NOUN	Gender=Fem	3 obj	-	-	SpaceAfter=No
4	.	.	PUNCT	PUNCT	PunctType=F	3 punct	-	-	
5									

Figura 20 – Fragmento de um arquivo Excel.

Gerado a partir da conversão do arquivo ConNLL-U obtido por meio do processamento pelo UDPipe.

Fonte: Captura de tela própria (2021).

Os arquivos em formato CoNLL-U obtidos por meio do UDPipe foram convertidos a formato Excel e, em seguida, foi feita a classificação e contagem do número de palavras abertas (palavras de conteúdo).

Pelos mesmos motivos explicados no ponto anterior, a nossa hipótese foi que a PEMT, por estar mais influenciada pela MT, teria menor diversidade que a MTNSS, que se aproximaria mais da TH nesse quesito.

6.3.7

Distância de edição

De forma muito simplificada, a lei da interferência enunciada por Toury (2012), e identificada por autores como Laviosa-Braithwaite (2001) como um “universal” da tradução, afirma que os fenômenos pertencentes ao arranjo do texto-fonte tendem a serem transferidos ao texto-alvo. Isso pode ser positivo, quando aumenta a frequência de uso de recursos existentes na língua alvo, ou negativo, quando se produzem desvios da norma padrão. Os sistemas de MT são conhecidos por realizar menos reordenamento frasal que os tradutores humanos (TORAL, 2019), portanto, podemos inferir que conteriam maior interferência da língua-fonte em termos do ordenamento das frases.

Para observar esse fenômeno, comparamos as sequências de palavras segundo suas classes gramaticais do texto fonte com as sequências das traduções obtidas por meio da MT, a PEMT, a MTNSS e a TH. As sequências de PoS (*Parts of Speech*) foram obtidas utilizando UDPipe. Para determinar a distância entre o original e as diferentes traduções foi utilizado o chamado “cálculo de distância de

Levenshtein”, mais conhecido como cálculo de distância de edição, que considera o número mínimo de inserções, remoções e substituições necessárias para que uma determinada cadeia de caracteres seja transformada em outra. Quanto maior a distância, mais inserções, remoções e substituições foram feitas.

Embora a NMT tenda a interferir menos em termos de PoS que outros sistemas de MT, como a MT estatística (BENTIVOGLI *et al.*, 2016), é esperado que a interferência da estrutura do texto original seja maior na PEMT que na MTNSS devido a maior influência da MT (TORAL, 2019). Assim, a nossa hipótese foi que a distância de edição com relação à MT seria menor na PEMT que na MTNSS.

6.3.8

Produtividade

Seguindo as instruções enviadas, os tradutores e tradutoras participantes gravaram a atividade nas suas telas durante a tarefa utilizando o programa VLC, que além de permitir observar os procedimentos também permite o controle de tempo.

Para efeitos do presente estudo, a produtividade foi considerada de maneira simples como o número de palavras que um tradutor consegue produzir em uma hora.

A nossa hipótese foi que a PEMT teria a maior produtividade, seguida pela MTNSS, e que não haveria uma diferença muito significativa entre elas.

6.3.9

Preferências dos leitores

As traduções do T1 por PEMT, MTNSS e TH foram avaliadas por leitores críticos voluntários para observar as preferências. Participaram da avaliação dez leitores, quatro do sexo masculino e seis do sexo feminino, todos eles argentinos,

falantes nativos de espanhol, com idades entre 24 e 60 anos. Desse total, quatro eram tradutores/linguistas. Foi pedido para cada um dos leitores que lessem duas versões de um mesmo texto e respondessem qual das duas preferiam. Considerando que a tradução utilizando PEMT foram realizadas por duas tradutoras diferentes, denominamos elas PEMT(1) e PEMT(2). As comparações foram feitas entre as diferentes modalidades de trabalho da seguinte maneira:

PEMT(1) X TH;

PEMT(2) X TH;

PEMT(1) X MTNSS;

PEMT(2) X MTNSS; e

MTNSS X TH.

A comparação de cada combinação foi feita por dois leitores (um total de dez comparações). A comparação só foi realizada só com o texto T1, o texto com características mais próximas a uma matéria jornalística. O pedido enviado para os avaliadores se inclui como Apêndice 6.

7

Resultados

Este capítulo apresenta os resultados obtidos. Para cada parâmetro analisado, foram geradas tabelas e/ou gráficos, que também são descritos brevemente levando em conta o que era esperado. Além dos resultados da análise das traduções obtidas por PEMT e MTNSS, que era o foco do estudo, também foram incluídos como referência os resultados do texto original, da TH e da MT “crua” (sem edição). Nos apêndices 7 e 8 encontram-se alguns exemplos do que foi observado.

7.1

Diversidade lexical (TTR)

Em ambos os arquivos é possível constatar uma diminuição da TTR com relação ao original em todas as modalidades de tradução. Isso estaria alinhado com o “universal” que afirma que as traduções tenderiam a ter uma diversidade lexical menor que os textos redigidos em uma língua determinada. Também se observa em ambos os textos uma redução marcada da TTR nas traduções obtidas por MT com relação ao texto na língua original. Isso estaria de acordo com a afirmação de Farrel (2018) de que a MT tende a oferecer só um subconjunto de soluções possíveis, que estaria determinado pelos dados utilizados para o treinamento dos sistemas. Essa redução é muito mais significativa no arquivo T1 (-16,9%) que no arquivo T2 (-4,92%), o que poderia ser explicado pelo fato do T2 ser um texto mais simples, repetitivo e esquemático. De fato, a diferença pode ser observada também nas TTR's dos arquivos originais, onde a TTR do T2 é 16% menor que a de T1. Por outro lado, observamos que os resultados obtidos com

PEMT, MTNSS diferem para ambos os tipos de texto no referente à diversidade lexical.

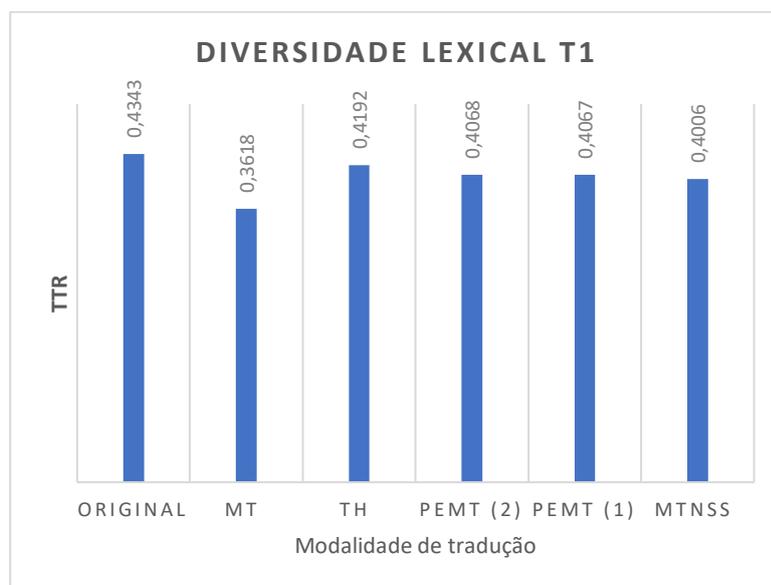


Figura 21 – Diversidade lexical (T1).
Fonte: Elaboração própria (2021).

Modalidade	TTR	Dif. c/rel. ao ORI	Dif. c/rel. à MT	Dif. c/rel. à TH
ORIGINAL	0,4343		20,04%	3,60%
MT	0,3618	-16,69%		-13,69%
PEMT (1)	0,4067	-6,36%	12,41%	-2,98%
PEMT (2)	0,4068	-6,33%	12,44%	-2,96%
MTNSS	0,4006	-7,76%	10,72%	-4,44%
TH	0,4192	-3,48%	15,87%	

Tabela 4 – Diversidade lexical T1.

A menor e a maior TTR (sem considerar a MT) se destacam em negrito.

Fonte: Elaboração própria (2021).

No arquivo T1, a redução significativa da TTR na MT (-16,69%) foi parcialmente compensada com um aumento na PEMT(1), PEMT(2) e MTNSS. Isso coincide com a observação de Toral (2019) de que o pós-editor adiciona algum grau de variedade lexical, mas que o fato de ter como ponto de partida a MT condiciona seu trabalho de modo tal que suas traduções não chegam a alcançar a diversidade da TH. Em tal sentido, embora todas as modalidades de tradução comparadas tenham apresentado uma TTR menor que a do original no T1, a tradução humana (TH) foi a que apresentou o valor mais alto entre as traduções e, portanto, mais próximo do texto original (-3,48%). Por sua vez, a

MTNSS é a que apresentou o valor mais distante (-7,76%). Isso confirmaria a hipótese de que a tradução humana é mais diversa que as modalidades de tradução com auxílio de MT. No entanto, enquanto a hipótese inicial era a de que a MTNSS fosse mais próxima da tradução humana do que a PEMT; no referente à diversidade lexical, isso não foi comprovado no T1.

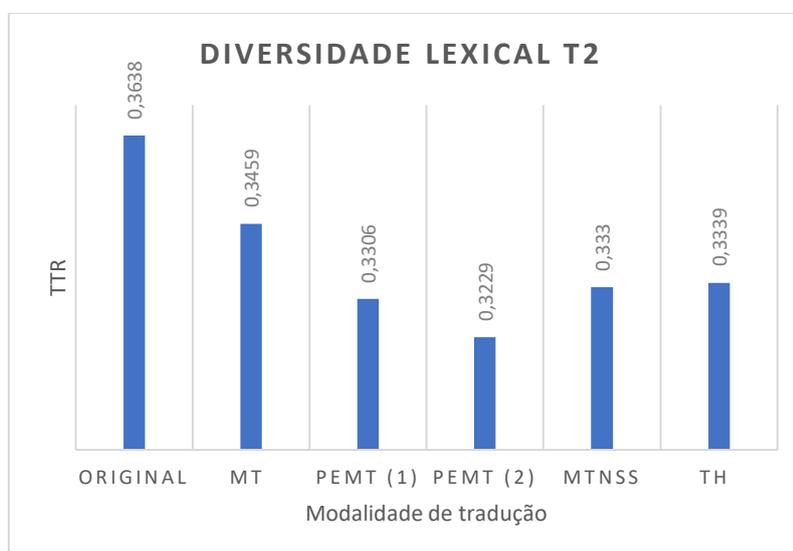


Figura 22 – Diversidade lexical (T2).
Fonte: Elaboração própria (2021).

Modalidade	TTR	Dif. c/rel. ao ORI	Dif. c/rel. à MT	Dif. c/rel. à TH
ORIGINAL	0,3638		5,17%	8,95%
MT	0,3459	-4,92%		3,59%
PEMT (1)	0,3306	-9,13%	-4,42%	-0,99%
PEMT (2)	0,3229	-11,24%	-6,65%	-3,29%
MTNSS	0,3330	-8,47%	-3,73%	-0,27%
TH	0,3339	-8,22%	-3,47%	

Tabela 5 – Diversidade lexical T2.

A menor e a maior TTR (sem considerar a MT) se destacam em negrito.

Fonte: Elaboração própria (2021).

No arquivo T2, todas as traduções apresentaram TTR's menores que o original, de acordo com o esperado segundo os "universais". A MT também apresentou uma redução da TTR com relação ao original. Aqui o resultado inesperado foi o de que tanto as PEMT como a MTNSS apresentaram TTR's menores ainda que a MT. Inclusive, a TH também apresentou uma TTR menor

que a MT. Fora essa consideração, no resto aconteceu algo similar ao arquivo T1, isto é, a TH ofereceu a maior diversidade e uma das PEMT [PEMT(1)] ofereceu a menor. A MTNSS ficou em segundo lugar como a mais diversa, depois da TH, o que comprovaria a hipótese inicial. De qualquer maneira, a diferença entre as modalidades de tradução com uso de MT e a TH para o T2 foi pouco significativa.

7.2

Densidade lexical

Considerando os “universais”, era esperado que as traduções tivessem menos palavras de conteúdo (SCARPA, 2006 *apud* TORAL, 2019) que os textos originais, o que também indicaria maior simplicidade do ponto de vista lexical. Isso só foi comprovado em um único caso de PEMT, e em um único arquivo. No restante das traduções, os textos tiveram mais palavras de conteúdo que o arquivo original. Assim como foi observado em relação à variedade lexical, no T1 a MT apresentou uma queda no número de palavras de conteúdo, que foi compensada por um aumento na MTNSS e nas PEMT. Na classificação, o maior número de palavras de conteúdo foi constatado no arquivo traduzido com MTNSS, seguido da TH e dos dois arquivos PEMT. Aqui as variações percentuais foram bem menores que na variedade lexical.

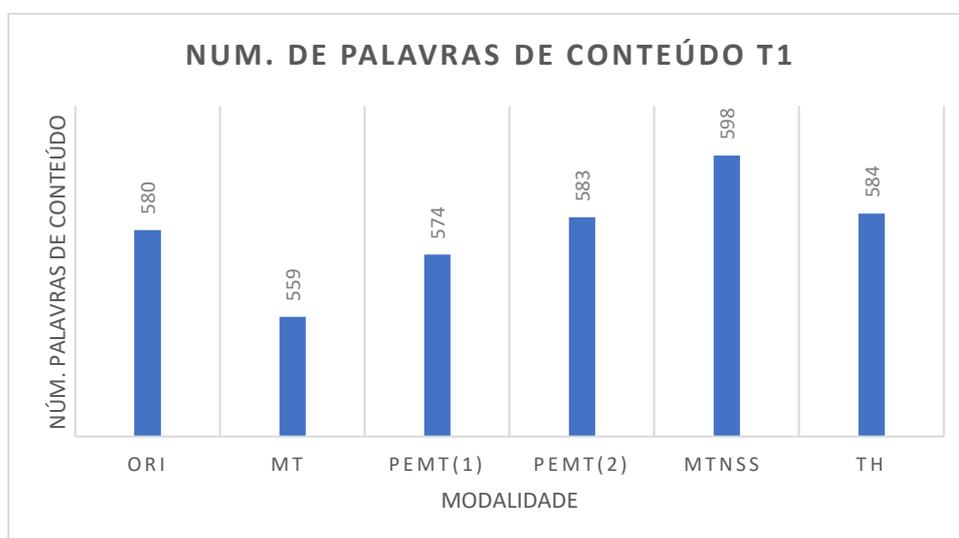


Figura 23 – Densidade lexical (T1).
Fonte: Elaboração própria (2021).

Modalidade	Num. pal. de conteúdo	Dif. rel. c/rel. a ORI	Dif. c/rel. a MT	Dif. c/rel. a TH
ORI	580		3,76%	-0,68%
MT (GT)	559	-3,62%		-4,28%
PEMT (T1)	574	-1,03%	2,68%	-1,71%
PEMT (T2)	583	0,52%	4,29%	-0,17%
MTNSS (T3)	598	3,1%	6,98%	2,40%
TH (T4)	584	0,69%	4,47%	

Tabela 6 – Densidade lexical T1.

A menor e a maior densidade lexical se destacam em negrito.

Fonte: Elaboração própria (2021).

No T1 foi verificada uma redução da densidade lexical na MT, que também foi compensada na pós-edição com um aumento em todas as modalidades (PEMT, MTNSS e TH). A maior densidade verificada foi na MTNSS, seguida da TH, e a menor densidade foi observada em uma das traduções obtidas por PEMT. A MTNSS demonstrou ser mais densa que as duas PEMT e, inclusive, que a TH.

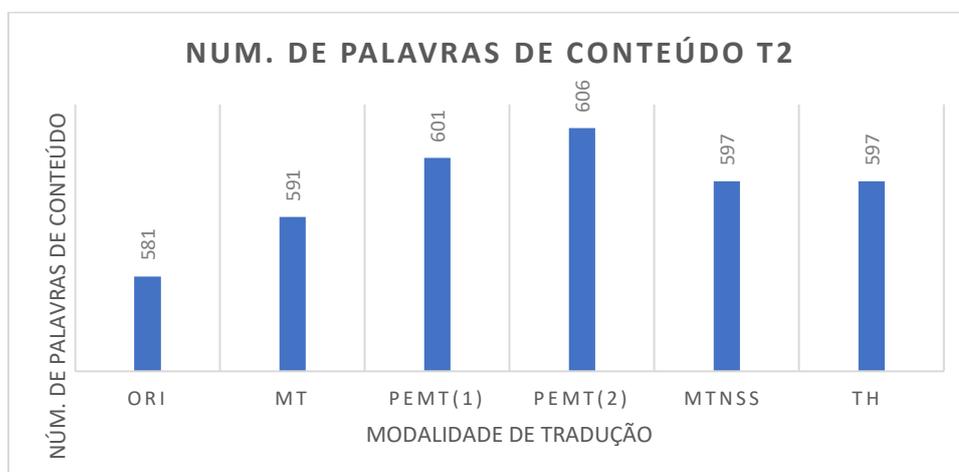


Figura 24 – Densidade lexical (T2).

Fonte: Elaboração própria (2021).

Modalidade	Num. pal. de conteúdo	Dif. c/rel. ao ORI	Dif. c/rel. a MT	Dif. c/rel. a TH
ORI	581		-1,69%	-2,68%
MT (GT)	591	1,72%		-1,01%
PEMT (T1)	601	3,44%	1,69%	0,67%
PEMT (T2)	606	4,3%	2,54%	1,51%
MTNSS (T3)	597	2,75%	1,02%	0,00%
TH (T4)	597	2,75%	1,02%	

Tabela 7 – Densidade lexical T2.

A menor e a maior densidade lexical se destacam em negrito.

Fonte: Elaboração própria (2021).

No arquivo T2 o resultado foi diferente. No texto obtido com MT, observou-se um aumento da densidade lexical, que foi ampliado ainda mais em todas as modalidades de tradução. A maior densidade lexical foi verificada em uma das traduções obtidas por PEMT, e a TH e a MTNSS ficaram iguais no último lugar (menor densidade).

7.3

Distância de edição

Considerando que os sistemas de MT realizam menos reordenamento de frases que os tradutores humanos, embora os sistemas de NMT introduzam mais alterações desse tipo que os sistemas baseados em frases (PBMT) (TORAL; SÁNCHEZ-CARTAGENA, 2017), poderíamos esperar que a distância de edição entre a MT e as PEMT fosse maior que entre a MT e a MTNSS.

No arquivo T1, a maior distância com respeito ao original corresponde à TH, como era esperado. A TH apresenta também a maior diferença com relação à MT, seguida da MTNSS, que apresenta uma distância significativamente maior que as outras duas modalidades de PEMT. Isso comprovaria a hipótese inicial de que nessa modalidade o tradutor trabalha mais livremente e é menos limitado pela MT.

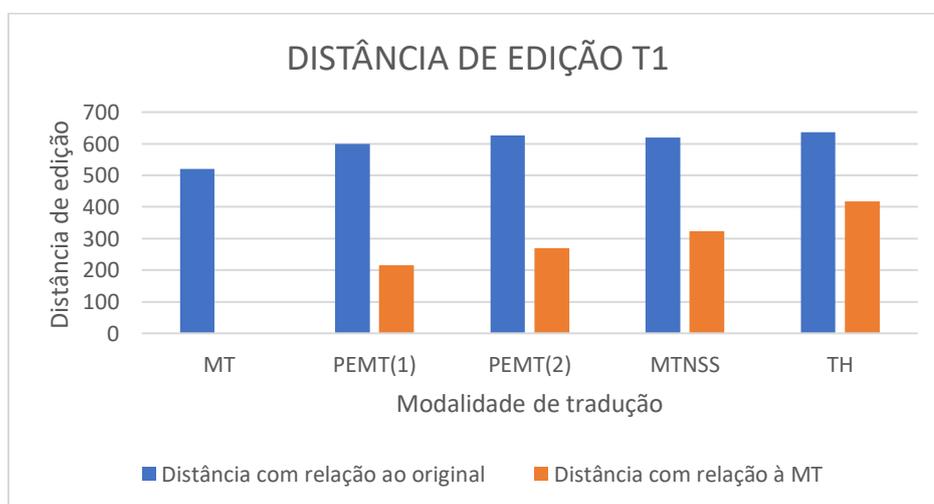


Figura 25 – Distância de edição (T1).
Fonte: Elaboração própria (2021).

Modalidade de tradução	Distância com relação ao original
MT	521
PEMT (T1)	600
PEMT (T2)	626
MTNSS (T3)	620
TH (T4)	636

Tabela 8 – Distância entre o original e as diferentes modalidades de tradução de T1.

A maior e a menor distância se destacam em negrito.

Fonte: Elaboração própria (2021).

Modalidade de tradução	Distância com relação à MT
PEMT(1)	215
PEMT(2)	270
MTNSS	324
TH	417

Tabela 9 – Distância entre a MT e as diferentes modalidades de tradução de T1.

A maior e a menor distância se destacam em negrito.

Fonte: Elaboração própria (2021).

No caso do arquivo T2, a maior distância de edição com respeito ao original foi apresentada pela PEMT(2) e a menor distância se observou na TH. A MTNSS apresenta a distância maior com relação à MT, e a TH fica em segundo lugar. As duas versões de PEMT estão mais próximas da MT, o que indicaria uma maior interferência.

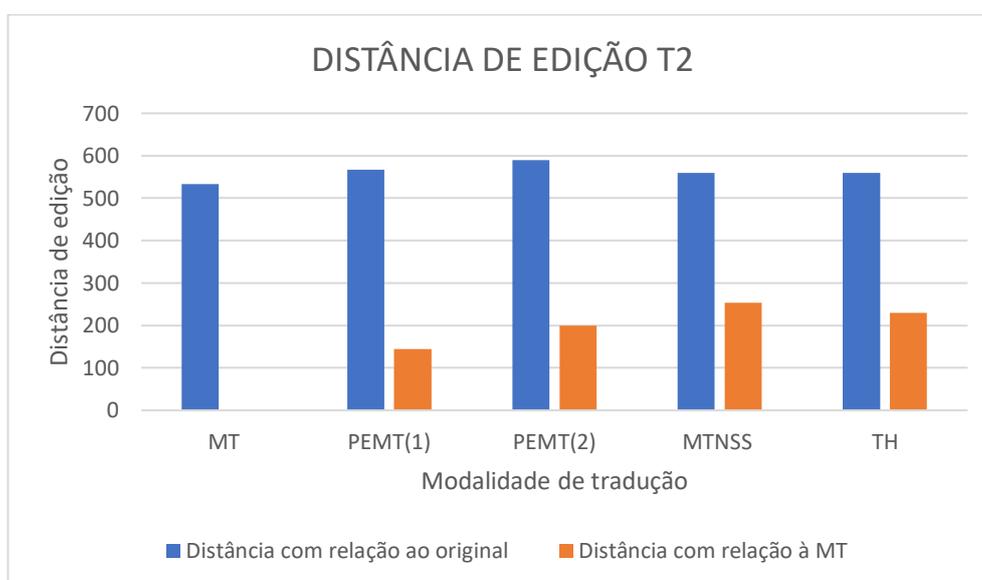


Figura 26 – Distância de edição (T2).

Fonte: Elaboração própria (2021).

Modalidade de tradução	Distância com relação ao original
MT	533
PEMT(1)	567
PEMT(2)	590
MTNSS	560
TH	559

Tabela 10 – Distância entre o original e as diferentes modalidades de tradução de T2.

A maior e a menor distância se destacam em negrito.

Fonte: Elaboração própria (2021).

Modalidade de tradução	Distância com relação à MT
PEMT (T1)	144
PEMT (T2)	200
MTNSS (T3)	253
TH (T4)	230

Tabela 11 – Distância entre a MT e as diferentes modalidades de tradução de T2.

A maior e a menor distância se destacam em negrito.

Fonte: Elaboração própria (2021).

A diferença entre T1 e T2 em termos da distância de edição parecer ser motivada pela diferença entre os tipos de texto, como será ampliado nas conclusões.

7.4

Produtividade

Nesta seção, observam-se os dados obtidos em relação ao tempo que demandou o trabalho de tradução para cada uma das modalidades e tradutores.

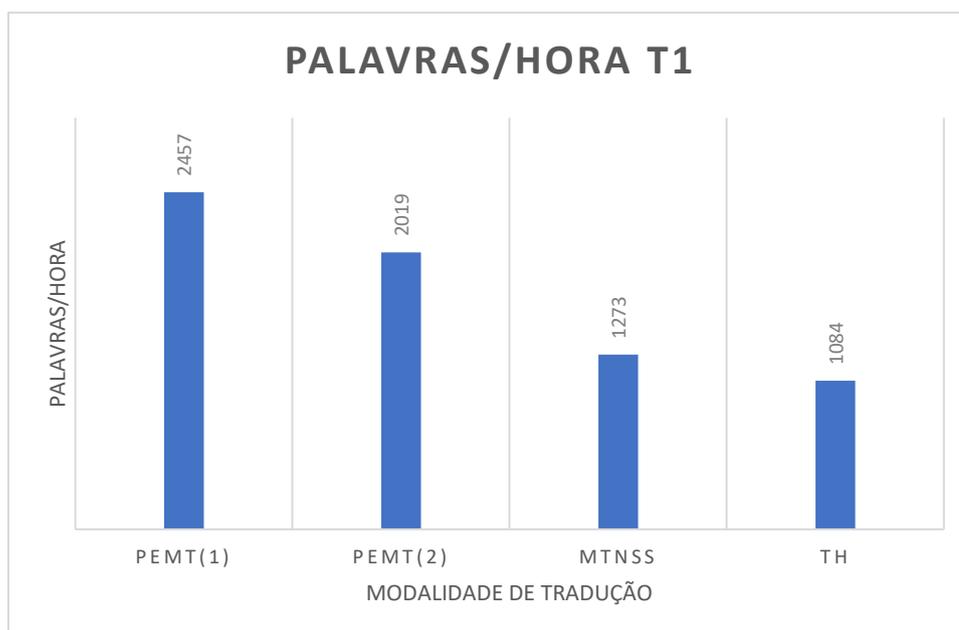


Figura 27 – Produtividade em palavras/hora (T1).

Fonte: Elaboração própria (2021).

Modalidade	Tempo	Palavras x hora
PEMT(1)	0:21	2457
PEMT(2)	0:29	2019
MTNSS	0:46	1273
TH	0:54	1084

Tabela 12 – Tempo consumido pela tradução + revisão da T1 em cada modalidade.

Fonte: Elaboração própria (2021).

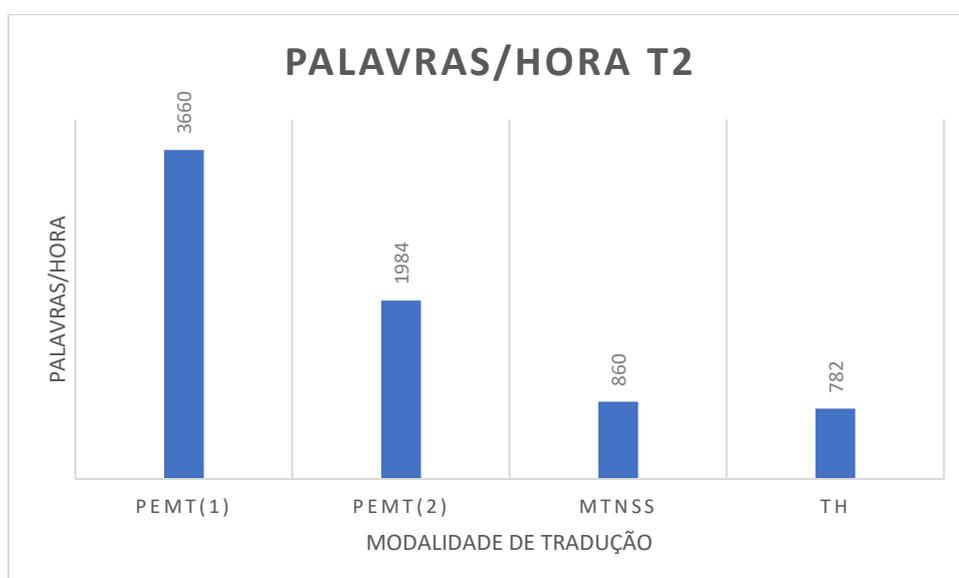


Figura 28 – Produtividade em palavras/hora (T1).

Fonte: Elaboração própria (2021).

Modalidade	Tempo	Palavras x hora
PEMT(1)	0:16	3660
PEMT(2)	0:26	1984
MTNSS	0:45	860
TH	1:06	782

Tabela 13 – Tempo consumido pela tradução + revisão da T2 em cada modalidade.
Fonte: Elaboração própria (2021).

A menor produtividade foi observada na TH, com 54 minutos para o T1 e 66 minutos para o T2 (média de 933 palavras/hora), seguida da MTNSS, com 46 minutos para T1 e 45 minutos para T2 (média de 1066,5 palavras/hora). As traduções feitas por PEMT tiveram tempos de 21 e 29 minutos para T1 e 16 e 26 minutos para T2 (média de 2530 palavras/hora). Aqui é necessário destacar que a PEMT(1) teve uma média significativamente mais baixa que a PEMT(2), o que será ampliado nas conclusões.

7.5

Preferências dos leitores

A seguir, são apresentadas as preferências dos leitores críticos em cada uma das comparações realizadas, seguidas do total das vezes em que cada modalidade de tradução foi preferida.

Modalidade	Avaliador 1	Avaliador 2
PEMT(1) X TH	PEMT(1)	TH
PEMT(2) X TH	PEMT(2)	PEMT(2)
PEMT(1) X MTNSS	MTNSS	MTNSS
PEMT(2) X MTNSS	MTNSS	PEMT(2)
MTNSS X TH	TH	TH

Tabela 14 – Resultado das avaliações.
Fonte: Elaboração própria (2021).

Preferência por PEMT(1) = 1

Preferência por PEMT(2) = 3

Preferência por MTNSS = 3

Preferência por TH = 3

Os resultados da avaliação dos leitores críticos foi um empate em três das quatro traduções obtidas (dois PEMT, uma MTNSS e uma TH). Uma das traduções por PEMT teve a pior avaliação, recebendo apenas um voto.

8

Conclusões

O estudo experimental realizado nesta dissertação teve o objetivo de comparar duas maneiras de utilizar a MT em um ambiente de TM (ferramenta CAT), a pós-edição tradicional, em que o tradutor edita um texto pré-traduzido por um sistema de MT, e a pós-edição no nível de subsegmento, em que o tradutor tem a opção de ir incorporando fragmentos de diferentes tamanhos oferecidos pelo sistema de MT na medida em que vai traduzindo. No primeiro caso, o tradutor edita, isto é, corrige e/ou altera o que achar preciso no texto pré-traduzido; no segundo, o tradutor inicia e organiza a frase e vai inserindo os fragmentos que achar conveniente para poupar digitação e aumentar a produtividade, podendo inclusive optar por introduzir a frase completa, se achar que a sugestão da MT é boa o suficiente. O sistema de MT utilizado foi o Google Tradutor, incorporado no ambiente do SDL Trados Studio. Tanto o Google Tradutor como o SDL Trados Studio são ferramentas disponíveis e amplamente utilizadas pelos tradutores *freelance* especializados. O estudo experimental incluiu a participação de quatro tradutores profissionais que traduziram do inglês para o espanhol, a sua língua nativa, dois textos redigidos originalmente em inglês com características diferentes, sendo um deles uma publicação de um blog de tecnologia, com um estilo de redação jornalístico, e o outro um texto altamente técnico.

Grande parte das pesquisas sobre pós-edição avaliam as traduções obtidas por meio da contagem e classificação de erros a partir de uma taxonomia determinada. Essa é a maneira em que as grandes agências de tradução costumam avaliar o trabalho dos tradutores (em geral, de maneira amostral). Porém, a opção no presente estudo foi analisar as características dos textos traduzidos a partir dos chamados “universais da tradução” (BAKER, 1993), mais especificamente, a simplificação, a normalização/homogeneização e a lei da interferência de Toury (2012). No caso, utilizamos essas características simplesmente como categorias

analíticas, não como indicadores de qualidade, para observar as diferenças entre a PEMT e a MTNSS, utilizando como referências a tradução 100% humana e a MT “crua”, isto é, sem nenhum tipo de edição. Portanto, a perspectiva adotada foi a de um tradutor especializado com o desejo de avaliar um método de trabalho alternativo à PEMT tradicional, para aumentar a produtividade mantendo o controle do processo.

A seguir resumiremos as respostas que achamos para as perguntas que orientaram a pesquisa:

1) Em que medida o trabalho em um ambiente de TM + MT oferece resultados diferentes usando a MT no nível de segmento (PEMT) e no nível de subsegmento (MTNSS)?

Foi possível observar diferenças entre a PEMT e MTNSS em algumas das características analisadas. Em certos casos, como a interferência, isso foi constatado nos dois tipos de textos traduzidos. Com relação a outras características, como a simplificação/homogeneização, foram observados impactos diferentes dependendo do tipo de texto. No caso do texto mais técnico (T2), a aproximação da MTNSS com a TH foi bastante significativa.

2) Em que consistem essas diferenças, caso existam?

Uma das diferenças observadas foi com relação à interferência do original, medida por meio da distância de edição considerando a sequência de PoS. Em tal sentido, a distância entre a MTNSS e a MT foi maior do que entre as PEMT e a MT em ambos os textos. Considerando que os sistemas de MT realizam um reordenamento menor das frases (Toral e Sánchez-Cartagena, 2017), isso permite inferir que a MTNSS seria a que apresenta menor interferência da MT e, portanto, do original. Também foram observadas diferenças com relação à simplificação e homogeneização, que serão descritas na resposta à pergunta que segue.

3) O uso de MT no nível de subsegmento é mais adequado para certos gêneros de texto do que para outros?

Esperava-se uma maior diversidade e densidade lexicais da MTNSS com relação a PEMT em ambos os tipos de texto; no entanto, no texto mais jornalístico (T1) foram observadas uma menor diversidade lexical e uma maior densidade lexical na MTNSS. No texto mais técnico (T2), contudo, tanto a densidade lexical como a diversidade lexical foram menores que as das PEMT, aproximando-se ainda mais da TH. Isso pode estar relacionado com o fato de que nas traduções

mais técnicas essas características podem não ter tanta variabilidade devido as limitações que impõe o estilo de comunicação, mais direto e objetivo, e inclusive a uniformidade lexical pode até ser considerada como uma característica positiva devido à necessidade de respeitar uma terminologia específica. Como afirmamos anteriormente, a principal diferença observada entre as modalidades de trabalho em ambos os tipos de textos foi a interferência. Aqui podemos considerar que a interferência de edição com relação às sequências de PoS é potencialmente mais notória em textos de tipo jornalístico que em textos técnicos, onde as possibilidades de utilizar estruturas frasais diferentes é mais limitada.

4) As duas maneiras de trabalhar com a MT, PEMT e MTNSS, permitem observar diferenças em termos de produtividade?

A produtividade da PEMT em palavras/hora foi muito superior à da MTNSS, com uma diferença média de 137,22%. Porém, é importante frisar que as duas traduções feitas com PEMT apresentaram diferenças muito significativas entre elas (no caso do T2, a diferença entre ambas PEMT atingiu um 45,8%). Também foi observada uma diferença de produtividade menor entre a PEMT e a MTNSS no T1 quando comparada ao T2. Isso indicaria que a MTNSS seria mais produtiva em textos de tipo jornalístico que em textos técnicos mais “duros”, onde a PEMT teria uma vantagem ainda mais significativa.

A MTNSS apresentou uma produtividade superior à da TH em ambos os tipos de texto, sendo mais significativa no T1 (17,44%) que no T2 (9,97%).

5) Os leitores avaliadores têm preferência pelos textos traduzidos com alguma das modalidades mencionadas (PEMT, MTNSS e TH)?

Os leitores avaliadores não manifestaram preferência por nenhuma modalidade de tradução específica. De fato, a TH, a MTNSS e uma das PEMT (PEMT(2)) ficaram empatadas com três votos cada uma. A outra PEMT, que foi também o caso em que se verificou a maior produtividade em termos de palavras/hora, obteve só um voto. Nesse caso uma explicação possível seria considerar que a diferença não esteve na modalidade de tradução, mas nessa tradução específica.

Para além das perguntas orientadoras da pesquisa, o trabalho permitiu fazer outras observações que vale a pena mencionar. A primeira é a constatação da influência da MT no trabalho de pós-edição, seja PEMT ou MTNSS, e da maneira como ela se manifesta. Em tal sentido, se considerarmos a maior diversidade

lexical e a maior densidade lexical como características desejáveis, especialmente para textos de tipo jornalístico e criativos em geral, podemos constatar que a MT provoca uma queda nesses indicadores, e que essa queda é compensada pelo trabalho de pós-edição. Um fenômeno similar pôde ser observado com relação à influência que a MT provoca em termos das sequências de PoS, onde uma proximidade inicial maior também é compensada no trabalho de pós-edição. No sentido contrário, no caso do texto mais técnico (T2), o que foi observado é que, em termos de diversidade lexical, todas as intervenções dos pós-editores seguiram a mesma tendência que a tradução humana, isto é, no sentido de reduzi-la. Isso pode ser atribuído à característica de uniformidade terminológica esperada nesse tipo de texto, além das especificidades da redação dos textos técnicos mais “duros”. Considerando isso, um aporte do presente trabalho é fornecer insumos para a conscientização do trabalho que faz o pós-editor em termos de intervenções capazes de definir certas características dos textos que produz. Essa conscientização tem potencialidade de ser utilizada no desenvolvimento das competências dos pós-editores. Além disso, o estudo também mostra a utilidade do uso de análises computacionais, simples e acessíveis, para observar e comparar características que seriam mais difíceis de constatar a olho nu.

O estudo também permitiu determinar que, à diferença do que muitos tradutores acreditam, nem sempre os leitores comuns são capazes de distinguir textos traduzidos de textos pós-editados ou, pelo menos, não manifestam preferências específicas pela tradução humana. Em tal sentido, os resultados não se alinham nem com trabalhos como os de Toral (2019), Fiederer e O'Brien (2009) e Farrel (2018), que observaram que os avaliadores preferiram traduções humanas a traduções obtidas por PEMT, nem com trabalhos como os de Green *et al.* (2013) e Bowker e Buitrago-Ciro (2015), onde os avaliadores expressaram preferências pela PEMT. No caso do presente estudo, o empate entre PEMT, MTNSS e TH permite pensar que as preferências dos leitores não estão determinadas pelo método de trabalho, mas por outros fatores.

Também cabe mencionar algumas limitações do estudo experimental realizado. Em primeiro lugar, as decorrentes do pequeno número de textos incluídos e tradutores participantes, o que limita as possibilidades de generalização dos resultados. Em segundo lugar, especialmente no que se refere à interferência nas sequências de PoS, teria sido útil incluir um modelo de

sequências de PoS típicas a partir de um corpus de textos similares para compará-las às sequências dos textos analisados. Isso teria possibilitado uma avaliação mais detalhada da influência da MT nas duas modalidades de trabalho estudadas. Com relação à avaliação por leitores críticos, também teria sido interessante criar um questionário para tentar determinar os motivos das preferências de forma mais detalhada, tentando isolar as características que provocam mais impacto para determinar a preferência de um texto sobre outro. Do mesmo modo, teria sido interessante avaliar a aceitação dos tradutores de um e outro método de trabalho (PEMT e MTNSS), uma vez que a produtividade é um fator importante, mas não é o único que pesa na hora de escolher uma modalidade de trabalho, especialmente quando não existem exigências específicas do cliente e/ou prazos muito ajustados.

Finalmente, gostaria de destacar que a presente dissertação pretende ser mais uma contribuição à construção de pontes entre os tradutores profissionais e a Universidade. É muito comum ouvir reclamações de tradutores profissionais sobre a distância que existe entre a teoria e os estudos acadêmicos e o seu trabalho cotidiano, mas se nós mesmos, tradutores, não levarmos para a Universidade as nossas preocupações e perspectivas, ninguém fará isso por nós.

Referências bibliográficas

ALCINA, A. Translation technologies: scope, tools and resources. **Target: International Journal on Translation Studies**, Netherlands, v. 20, n. 1, p. 79-102, 2008.

ALMEIDA, G. **Translating the post-editor**: an investigation of post-editing changes and correlations with professional experience across two Romance languages. 2013. 350 f. Tese (Doctor of Philosophy) – Dublin City University, Dublin, 2013.

ALONSO, E; CALVO, E. Developing a Blueprint for a technology-mediated approach to translation studies. **Meta**, Montreal, v. 60, n. 1, p. 135-157, abr. 2015. Disponível em: <https://doi.org/10.7202/1032403ar>. Acesso em: 12 mar. 2021.

ANGELELLI, C. (ed.). **The sociological turn in translation and interpreting studies**. Amsterdam/Philadelphia: John Benjamins, 2014.

ARENAS, A. G. Productivity and quality in MT post-editing. *In*: MACHINE TRANSLATION SUMMIT – WORKSHOP: BEYOND TRANSLATION MEMORIES: NEW TOOLS FOR TRANSLATORS MT, 12., 2009, [s.l.]. **Anais [...]** [S.l.]: MTSummit, 2009.

ARENAS, A. G. The role of professional experience in post-editing from a quality and productivity perspective. *In*: O'BRIEN, S.; CARL, M.; WINTHER BALLING, L.; SPECIA, L.; SIMARD, M. **Post-editing of machine translation: processes and applications**. Newcastle Upon Tyne: Cambridge Scholars Publishing, 2014.

CHESTERMAN, A; WAGNER, E. **Can theory help translators?:** a dialogue between the ivory tower and the wordface. Nova York: St. Jerome Publishing, 2010.

KOPONEN, M. Is machine translation post-editing worth the effort?: a survey of research into post-editing and effort. **The Journal of Sepcialised Translation**, Online, n. 25, jan. 2016.

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. Amsterdam/Filadelfia: Benjamins, 1993.

BAKER, M. Towards a methodology for investigating the style of a literary translator. **Target: International Journal of Translation Studies**, Netherlands, v. 12, n. 2, p. 241-266, 2000.

BAR-HILLEL, Y. The present state of research on mechanical translation. **Research Laboratory of Eletronics**, Massachusetts, n. 2, p. 229-237, 1951. Disponível em: <https://doi.org/10.1002/asi.5090020408>. Acesso em: 12 mar. 2021.

BASSNETT, S. **Translation studies**. Londres: Methuen & Co. Ltd., 1980.

BENTIVOGLI, L.; BISAZZA, A.; CETTOLO, M.; FEDERICO, M. Neural versus phrasebased machine translation quality: a case study. *In*: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, EMNLP 2016, 2016, Austin, Texas. **Anais [...]** Austin, Texas: Association for Computational Linguistics, 2016. p. 257–267.

BIAU GIL, J.; PYM, A. Technology and translation: a pedagogical review. *In*: PYM, A.; PERESTRENKO, A.; STARINK, B. (Ed.). **Translation technology and its teaching**. Tarragona: Intercultural Studies Group; Universitat Rovirai Virgili, 2006. Disponível em: https://usuaris.tinet.cat/apym/online/translation/BiauPym_TechnologyAndTranslation.pdf. Acesso em: 12 mar. 2021.

BOWKER, L.; BUITRAGO CIRO, J. Investigating the usefulness of machine translation for newcomers at the public library. **Translation and Interpreting Studies**, v. 10, n. 2, p. 165-186, 2015.

BUNDGAARD, K. Translator attitudes towards translator-computer interaction: findings from a workplace study. **HERMES: Journal of Language and Communication in Business**, n. 56, p. 125-144, 2017.

CARMO, F.; SHTERIONOV, D.; MOORKENS, J.; WAGNER, J.; HOSSARI, M.; PAQUIN, E.; SCHMIDTKE, D.; GROVES, D.; WAY, A. A review of the state-of-the-art in automatic post-editing. **Machine Translation**, n. 35, p. 101-143, dec. 2020. Disponível em: <https://doi.org/10.1007/s10590-020-09252-y>. Acesso em: 17 fev. 2021.

CHAUME, F. The turn of audiovisual translation: new audiences and new technologies. **Translation Spaces**, v. 2, p. 105-123, 2013.

CHESTERMAN, A. The name and nature of translator studies. **HERMES: Journal of Language and Communication in Business**, v. 22, n. 42, p. 13-22, 2009.

CHRISTENSEN, T.; FLANAGAN, M.; SCHJOLDAGER, A. Mapping translation technology research in translation studies: an introduction to the thematic section. **HERMES: Journal of Language and Communication in Business**, n. 56, p. 7-20, 2017. Disponível em: https://www.researchgate.net/publication/320326624_Mapping_Translation_Technology_Research_in_Translation_Studies_An_Introduction_to_the_Thematic_Section. Acesso em: 17 jul. 2019.

CRONIN, M. **Translation in the digital era**. New York: Routledge, 2013.

DAEMS, J.; DE CLERCQ, O.; MACKEN, L. Translationese and post-edited: how comparable is comparable quality? **Linguistica Antverpiensia New Series: Themes In Translation Studies**, v. 16, p. 89-103, 2017.

DENERO, J.; GREEN, S. Lilt: an interactive & adaptive mt based translator assistant or cat tool. **Blog eMpTy Pages**, 10 nov. 2016. Disponível em: <http://kv-emptypages.blogspot.com/2016/11/lilt-interactive-adaptive-mt-based.html>. Acesso em: 20 jan. 2020.

DEPALMA, D. Augmenting human translator performance. **CSA Research, Blogs e Events**, 10 feb. 2021. Disponível em: <https://csa-research.com/Blogs-Events/Blog/Augmenting-Human-Translator-Performance>. Acesso em: 10 mar. 2021.

DIÑO, G. Reader Polls: post-editing, interactive prediction, translation devices, and M&A. **Blog Slator, News, Features**, 30 may 2019. Disponível em: <https://slator.com/features/reader-polls-post-editing-interactive-prediction-translation-devices-and-ma/>. Acesso em: 17 jul. 2019.

ESTILL, D. A. **De artesanato a indústria – a tradução globalizada**: autoria, texto de partida, tradutor e texto de chegada na era da informação. 2017. 130 f. Tese (Doutorado em Estudos da Linguagem) – Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, 2017.

FARRELL, M. Machine translation markers in post-edited machine translation output. *In*: CONFERENCE TRANSLATING AND THE COMPUTER, 40., 2018, London. **Anais [...]** London: AsLing, 2018. p. 50-59.

FEDERICO, M.; CATTELAN, A.; TROMBETTI, M. Measuring user productivity in machine translation enhanced computer assisted translation. *In*: CONFERENCE OF THE ASSOCIATION FOR MACHINE TRANSLATION IN THE AMERICAS: RESEARCH PAPERS, 10., 2012, California. **Anais [...]** California: AMTA, 2012.

GREEN, S.; HEER, J.; MANNING, C. The efficacy of human post-editing for language translation. *In*: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, CHI 13, 2013, Paris. **Anais [...]** Paris: SIGCHI, 2013. p. 439-448.

GREEN, S.; WANG, S. I.; CHUANG, J.; HEER, J.; SCHUSTER, S.; MANNING, C. D. Human effort and machine learnability in computer aided translation. *In: CONFERENCE ON EMPIRICAL METHODS IN NATURAL LANGUAGE PROCESSING, EMNLP 2014, 2014, Doha. Anais [...]* Doha: EMNLP, 2014. p. 1225-1236.

GREEN, S.; HEER, J.; MANNING, C. Natural language translation at the intersection of AI and HCI: old questions being answered with both AI and HCI. *ACM Queue*, v. 13, n. 6, p. 1-13, 2015. Disponível em: <https://queue.acm.org/detail.cfm?id=2798086>. Acesso em: 17 jul. 2019.

GRÖHN, A. **Suitability of neural machine translation for different types of texts: a study on potential predictors**. 2019. 57 f. Thesis (Pro Gradu English Translation) – Department of Modern Languages, Universidad de Helsinki, Helsinki, 2019.

GRUDIN, J. Moving Target: the evolution of human-computer interaction. *In: JACKO, J. A. (Ed.). Human-computer interaction handbook: fundamentals, evolving technologies, and emerging applications*. 3. ed. Boca Raton: Taylor & Francis Group, 2012.

HARDT, M.; NEGRI, A. **Império**. 2. ed. Rio de Janeiro: Record. 2002.

HICKEY, S.; AGULLÓ GARCÍA, B. The 2021 Nimdzi 100: the ranking of Top 100 Largest Language Service Providers. **Nimdzi**, Nimdzi Research, Market Research, Featured market research, 1 mar. 2021. Disponível em: <https://www.nimdzi.com/nimdzi-100-top-lsp/#growth-projection>. Acesso em: 17 may 2021.

HU, K.; CADWELL, P. A Comparative study of post-editing guidelines. **Baltic Journal of Modern Computing**, v. 4, n. 2, p. 346-353, 2016.

HUANG, C.; YANG, P.; CHEN, K.; CHANG, J. S. TransAhead: a computer-assisted translation and writing tool. *In: CONFERENCE OF THE NORTH AMERICAN CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS: HUMAN LANGUAGE TECHNOLOGIES, 2012, Montreal. Anais [...]* Montreal: ACL, 2012. p. 352–356. Disponível em: https://www.researchgate.net/publication/236965527_TransAhead_A_Computer-Assisted_Translation_and_Writing_Tool. Acesso em: 22 jul. 2019.

HUTCHINS, W. J. **The first public demonstration of machine translation: the Georgetown-IBM system, 7th January 1954**. 2005. Disponível em: <http://www.hutchinsweb.me.uk/GU-IBM-2005.pdf>. Acesso em: 10 jul. 2019.

HUTCHINS, W. J. **Machine translation: past, present, future**. West Sussex, England: Ellis Horwood Limited, 1986.

HUTCHINS, W. J. Machine translation: a brief history. *In*: KOERNER, F.; ASHER, R. (Ed.). **Concise history of the language sciences: from the Sumerians to the cognitivists**. Oxford: Pergamon Press, 1995.

KINGSCOTT, G. Technical translation and related disciplines. **Perspectives: Studies in translatology**, [S.l.], v. 10, n. 4, p. 247-255, 2010.

KOBY, G. S.; FIELDS, P.; HAGUE, D. R.; LOMMEL, A.; MELBY, A. Defining translation quality. **Tradumàtica: tecnologies de la traducció**, n. 12, p. 413-420, dec. 2014. Disponível em: https://ddd.uab.cat/pub/tradumatica/tradumatica_a2014n12/tradumatica_a2014n12p413.pdf. Acesso em: 10 jul. 2019.

KOEHN, P. **Statistical machine translation**. Nova York: Cambridge University Press, 2010.

KOEHN, P.; KNOWLES, R. Six challenges for neural machine translation. *In*: WORKSHOP ON NEURAL MACHINE TRANSLATION, 1., 2017, Vancouver, Canadá. **Anais [...]** Vancouver, Canadá: Neural Machine Translation, 2017. p. 28-39.

LICKLIDER, J. C. R. Man-computer symbiosis. **IRE Transactions on Human Factors in Eletronics**, v. HFE-1, p. 4-11, 1960.

MALVERN, D.; RICHARDS, B. J.; CHIPERE, N. **Lexical diversity and language development: quantification and assessment**. Nova York: Palgrave Macmillan, 2004.

MARTINS, M. A. P. Descriptive translation studies: uma revisão crítica. **Gragoatá**, Niterói, v. 7, n. 13, p. 33-52, 2002. Disponível em: <https://periodicos.uff.br/gragoata/article/view/33512>. Acesso em: 10 jul. 2019.

MATEO, R. A deeper look into metrics for translation quality assessment (TQA): a case study. **Miscelânea: A Journal of English and American Studies**, n. 49, p. 73-94, 2014.

MUNDAY, J. **Introducing translation studies: theories and applications**. Routledge: Londres; Nova York, 2016.

O'BRIEN, S. Translation as human-computer interaction. **Translation Spaces**, Netherlands, v. 1, n. 1, p. 101-122, 2012.

OLOHAN, M. Sociological approaches to translation technology. *In*: O'HAGAN, M. (Ed.). **The routledge handbook of translation and technology**. London: Routledge, 2019.

PAIE, K. E2f and Lilt case study: first large-scale application of auto-adaptive machine translation. **E2f**, News, Technology, 13 jun. 2016. Disponível em: <https://e2f.com/7539/>. Acesso em: 17 jul. 2019.

PLITT, M.; MASSELOT, F. A productivity test of statistical machine translation post-editing in a typical localisation context. **The Prague Bulletin of Mathematical Linguistics**, Praga, n. 93, p. 7-16, jan. 2010. Disponível em: <<http://ufal.mff.cuni.cz/pbml/93/art-plitt-masselot.pdf>>. Acesso em: 17 jul. 2019.

POPIOLEK, M. Industry Standards Demystified: Part 5. **American Translators Associations**, ATA-TCD News Articles, 22 jan. 2021. Disponível em: <http://ata-tcd.com/iso-185872017-translation-services-post-editing-of-machine-translation-output-requirements/>. Acesso em: 3 maio 2021.

PORSIEL, J. Machine Translation at Volkswagen: a case study. **MultiLingual Magazine**, dez. 2008. Disponível em: <https://ufal.mff.cuni.cz/project/tle2009/presentations/porsiel-multilingual-mt-vw-case-study-2008.pdf>. Acesso em: 17 jul. 2019.

QUAH, C. K. **Translation and Technology**. Hampshire, Nova York: Palgrave Macmillan, 2006.

RISKU, H. A cognitive scientific view on technical communication and translation: do embodiment and situatedness really make a difference? **Target: International Journal on Translation Studies**, v. 22, n. 1, p. 94-111, 2010.

ROWDA, J. Inside eBay's translation machine. **Slator**, News, Features, 1 nov. 2016. Disponível em: <https://slator.com/features/inside-ebay-translation-machine/>. Acesso em: 17 jul. 2019.

SALDANHA, G.; O'BRIEN, S. **Research methodologies in translation studies**. Londres, Nova York: Routledge, 2014.

SIN-WAI, C. (Ed.). **The routledge encyclopedia of translation technology**. Londres, Nova York: Routledge, 2016.

SCARPA, F. Corpus-based quality assessment of specialist translation: a study using parallel and comparable corpora in english and italian. *In*: SARVECIC, S.; GOTTI, M. **Insights into specialized Translation**. Berna: Peter Lang, 2006. p. 154-172.

SIN-WAI, C. **The future of translation technology: towards a world without Babel**. Londres, Nova York: Routledge, 2017.

SPARCK JONES, K. Natural language processing: a historical review. *In*: **Computational Linguistics**, 2001.

SPINZI, C.; RIZZO, A.; ZUMMO, M. (Ed.). **Translation or transcreation?: discourses, texts and visuals**. Newcastle Upon Tyne: Cambridge Scholars Publishing, 2018.

STATISTA RESEARCH DEPARTAMENT. Market size of the global language services industry 2009-2021. **Statista**, Services, Business Services, 16 sep. 2021. Disponível em: <https://www.statista.com/statistics/257656/size-of-the-global-language-services-market/>. Acesso em: 4 fev. 2021.

STRAKA, M.; STRAKOVÁ, J. Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. *In*: CONFERENCE ON COMPUTATIONAL NATURAL LANGUAGE LEARNING, CoNLL 2017, 2017, Vancouver, Canada. **Anais** [...] Vancouver, Canada: ACL, 2017. Disponível em: <https://www.semanticscholar.org/paper/Tokenizing%2C-POS-Tagging%2C-Lemmatizing-and-Parsing-UD-Straka-Straková/ea40fba2dc52bde32dedd45b05f83d27dc2ea6c7>. Acesso em: 3 ago. 2021.

TEIXEIRA, M. S. **O jogo da avaliação**: um estudo prático sobre tradução automática. 2018. 134 f. Dissertação (Mestrado em Estudos da Linguagem) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, 2018.

TORAL, A. Post-editeese: an exacerbated translationese. *In*: MACHINE TRANSLATION SUMMIT, 17., 2019, Dublin. **Anais** [...] Dublin: MTSummit, 2019. p. 273-281.

TSC 3016 Miguel A Jiménez Crespo. Salamanca: [s.n.], 2018. 1 vídeo (45 min). Publicado pelo canal Ediciones Universidad de Salamanca. Disponível em: <https://bit.ly/3Bb1nTd>. Acesso em: 22 jul. 2019.

VAN DER MEER, J. Translation Technology: past, present and future. *In*: ANGELONE, E.; EHRENSBERGER-DOW, M.; MASSEY, G. **The Bloomsbury Companion to Language Industry Studies**. Londres: Bloomsbury, 2020.

VAN DER MEER, J. The future does not need translators. **TAUS: the language data network**, Insights, Blog, Food for Thought, 24 feb. 2016. Disponível em: <https://blog.taus.net/the-future-does-not-need-translators>. Acesso em: 23 jul. 2019.

VANDEGHINSTE, V.; VANALLEMEERSCH, T.; AUGUSTINUS, L.; BULTÉ, B.; VAN EYNDE, F.; PELEMANS, J.; VERWIMP, L.; WAMBACQ, P.; HEYMAN, G.; MOENS, M.-F.; VAN DER LEK-CIUDIN, I.; STEURS, F.; RIGOUTS TERRY, A.; LEFEVER, E.; TEZCAN, A.; MACKEN, L.; HOSTE, V.; DAEMS, J.; BUYSSCHAERT, J.; COPPERS, S.; VAN DEN BERGH, J.; LUYTEN, K. Improving the translation environment for professional translators advances. **Informatics**, v. 6, n. 2, p. 24, 2019. Disponível em: <https://doi.org/10.3390/informatics6020024>. Acesso em: 20 jul. 2019.

VASHEE, K. The quest for human parity machine translation. **eMpTy Pages**, 29 mar. 2021. Disponível em: <https://kv-emptypages.blogspot.com>. Acesso em: 20 jul. 2021.

VERMEER, H. Skopos and commission in translational action. *In*: VENUTI, L. (Ed.). **The Translation Studies Reader**. Routledge: Nova York, 2012.

WAY, A. Machine translation: where are we at today? *In*: ANGELONE, E.; EHRENSBERGER-DOW, M.; MASSEY, G. **The Bloomsbury Companion to Language Industry Studies**. Londres: Bloomsbury, 2020.

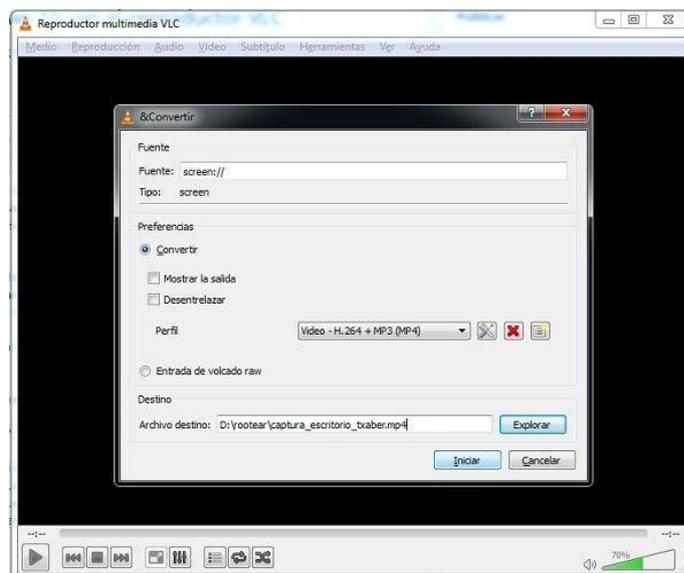
Apêndice A

Instruções para a gravação da tela do PC

Para realizar la grabación del escritorio con el reproductor se deben seguir estos pasos:

- a) abrir el reproductor VLC previamente instalado;
- b) pulsar en Medio;
- c) pulsar en Abrir dispositivo de captura (o sustituir los pasos 2 y 3 por la combinación de teclas Control+C);
- d) en Modo de captura seleccionar Escritorio. En la parte inferior donde dice Reproducir, cambiar esa opción por Convertir;
- e) seleccionar el tipo de archivo en el que se guarda la grabación;
- f) seleccionar la ruta de destino del vídeo;
- g) pulsar en Iniciar.

Esta es una captura de la última pantalla a la que llegarán antes de comenzar con la grabación:



Apêndice B

Ficha do projeto T1 para pós-edición de Machine Translation (PEMT)

Ficha del proyecto

Documento: Tottenham Hotspur (Word)

Número de palabras: 976 (Ver log file)

Valor: USD 25,00

Fecha de entrega: Diciembre de 2019

Descripción del servicio: Posedición hard inglés > español latinoamericano
(Posedición hard = texto con calidad de traducción humana)

Material incluido en el paquete:

- TM
- Glosario

Deliverables: Pacote de retorno + Target

Observaciones:

- 1) Por favor, empezar a grabar la pantalla antes de abrir el paquete y terminar de grabar cuando el proyecto esté listo;
- 2) Por favor, no usar ningún recurso (TM adicionales, MT, otros glosarios, etc.) fuera del paquete, excepto el diccionario de AutoSuggest (uso opcional).

Totals											
Total	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags	
Files:1	PerfectMatch	0	0	0	0.00%	0	0	0		0	
Chars/Word:5.13	Context Match	0	0	0	0.00%	0	0	0		0	
	Repetitions	0	0	0	0.00%	0	0	0		0	
	Cross-file Repetitions	0	0	0	0.00%	0	0	0		0	
	100%	0	0	0	0.00%	0	0	0		0	
	95% - 99%	0	0	0	0.00%	0	0	0		0	
	85% - 94%	0	0	0	0.00%	0	0	0		0	
	75% - 84%	0	0	0	0.00%	0	0	0		0	
	50% - 74%	0	0	0	0.00%	0	0	0		0	
	New/AT	54	976	5011	100.00%	16	34	0		0	
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0		0	
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0	
	Total		54	976	5011	100%	16	34	0	0.00%	0

¡Muchas gracias!

Apêndice C

Ficha do projeto T1 para tradução humana (TH)

Ficha del proyecto

Documento: Tottenham Hotspur (Word)

Número de palabras: 976 (Ver log file)

Valor: USD 25,00

Fecha de entrega: 21/12

Descripción del servicio: Traducción inglés > español latinoamericano

Material incluido:

- TM
- Glosario

Deliverables: SDLXLIFF + Target

Observaciones:

- 1) Por favor, empezar a grabar la pantalla antes de abrir el paquete y terminar de grabar cuando el proyecto esté listo;
- 2) Por favor, no usar ningún recurso (TM adicionales, MT, otros glosarios, etc.) fuera de lo que se envía en el zip.

Totals											
Total	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags	
Files:1	PerfectMatch	0	0	0	0.00%	0	0	0		0	
Chars/Word:5.13	Context Match	0	0	0	0.00%	0	0	0		0	
	Repetitions	0	0	0	0.00%	0	0	0		0	
	Cross-file Repetitions	0	0	0	0.00%	0	0	0		0	
	100%	0	0	0	0.00%	0	0	0		0	
	95% - 99%	0	0	0	0.00%	0	0	0		0	
	85% - 94%	0	0	0	0.00%	0	0	0		0	
	75% - 84%	0	0	0	0.00%	0	0	0		0	
	50% - 74%	0	0	0	0.00%	0	0	0		0	
	New/AT	54	976	5011	100.00%	16	34	0		0	
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0		0	
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0	
	Total		54	976	5011	100%	16	34	0	0.00%	0

¡Muchas gracias!

Apêndice D

Ficha do projeto T2 para pós-edição de Machine Translation (PEMT)

Ficha del proyecto

Documento: FAQ - Insight Online (Word)

Número de palabras: 897 (Ver log file)

Valor: USD 25,00

Fecha de entrega: Diciembre de 2019

Descripción del servicio: Posedición hard inglés > español latinoamericano
(Posedición hard = texto con calidad de traducción humana)

Material incluido en el paquete:

- TM
- Glosario

Observaciones:

- 1) Por favor, empezar a grabar la pantalla antes de abrir el paquete y terminar de grabar cuando el proyecto esté listo;
- 2) Por favor, no usar ningún recurso (TM adicionales, MT, otros glosarios, etc.) fuera del paquete, excepto el diccionario de AutoSuggest (uso opcional).

File Details											
File	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags	
Insight_Online_FAQ.docx.sdxliff	PerfectMatch	0	0	0	0.00%	0	0	0		0	
Chars/Word:5.69	Context Match	0	0	0	0.00%	0	0	0		0	
	Repetitions	0	0	0	0.00%	0	0	0		0	
	Cross-file Repetitions	0	0	0	0.00%	0	0	0		0	
	100%	15	37	217	4.12%	0	0	0		0	
	95% - 99%	8	8	128	0.89%	11	0	0		8	
	85% - 94%	1	1	3	0.11%	1	0	0		0	
	75% - 84%	2	5	32	0.56%	0	2	0		0	
	50% - 74%	3	10	58	1.11%	6	2	0		5	
	New/AT	69	836	4666	93.20%	98	142	0		43	
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0		0	
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0	
	Total		98	897	5104	100%	116	146	0	0.00%	56

¡Muchas gracias!

Apêndice E

Ficha do projeto T2 para tradução humana (TH)

Ficha del proyecto

Documento: FAQ - Insight Online (Word)

Número de palabras: 897 (Ver log file)

Valor: USD 26,00

Fecha de entrega: Diciembre de 2019

Descripción del servicio: Traducción inglés > español latinoamericano

Material incluido en el paquete:

- TM
- Glosario

Observaciones:

- 1) Por favor, calcular el tiempo exigido desde que se empieza a traducir hasta el final de la revisión;
- 2) Por favor, no usar ningún recurso (TM adicionales, MT, otros glosarios, etc.) fuera del paquete.

File Details										
File	Type	Segments	Words	Characters	Percent	Recognized Tokens	Fragment words (whole TU)	Fragment words (TU fragment)	AdaptiveMT Impact	Tags
Insight Online FAQ.docx.sdxliff	PerfectMatch	0	0	0	0.00%	0	0	0	0	0
Chars/Word:5.69	Context Match	0	0	0	0.00%	0	0	0	0	0
	Repetitions	0	0	0	0.00%	0	0	0	0	0
	Cross-file Repetitions	0	0	0	0.00%	0	0	0	0	0
	100%	15	37	217	4.12%	0	0	0	0	0
	95% - 99%	8	8	128	0.89%	11	0	0	0	8
	85% - 94%	1	1	3	0.11%	1	0	0	0	0
	75% - 84%	2	5	32	0.56%	0	2	0	0	0
	50% - 74%	3	10	58	1.11%	6	2	0	0	5
	New/AT	69	836	4666	93.20%	98	142	0	0	43
	AdaptiveMT Baseline	0	0	0	0.00%	0	0	0	0	0
	AdaptiveMT with Learnings	0	0	0	0.00%	0	0	0	0.00%	0
	Total	98	897	5104	100%	116	146	0	0.00%	56

¡Muchas gracias!

Apêndice F

Texto do e-mail para os avaliadores

¡Hola, XXXX!

Muchas gracias por colaborar con mi investigación de maestría.

La tarea que te pido es simple, que hagas una lectura de dos versiones de un mismo artículo para después responder la pregunta:

¿Cuál de los dos textos preferís?

No hace falta ninguna fundamentación ni usar ningún criterio específico para tu respuesta.

Los enlaces para bajar los textos o leerlos online son:

TX Link

TY Link

En lo posible, sería bueno recibir la respuesta hasta el XX de XXXXX de 2021.

Una vez más, ¡muchas gracias por la ayuda!

Apêndice G

Exemplos de simplificação e interferência T1

Exemplos de simplificação

Exemplos de explicitação

Exemplos de interferência

Exemplos de alteração da ordem da frase

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
Managing Director for UK&I Marc Waters explains how HPE technology built into the fabric of the stadium will transform fan experience.	El Director General de UK&I , Marc Waters, explica cómo la tecnología HPE incorporada en la estructura del estadio transformará la experiencia de los fanáticos .	El Director General de Reino Unido e Irlanda , Marc Waters, explica cómo la tecnología HPE incorporada en la estructura del estadio transformará la experiencia de los aficionados.	El Director General de UK&I , Marc Waters, explica cómo la tecnología HPE incorporada en la estructura del estadio transformará la experiencia de los aficionados.	Marc Waters, director ejecutivo para el Reino Unido e Irlanda , explica cómo la tecnología HPE incorporada en la estructura del estadio transformará la experiencia de los aficionados.	El Director Ejecutivo de GB e Irlanda , Marc Waters, explica cómo la tecnología de HPE, incorporada al tejido del estadio, va a transformar la experiencia de los fanáticos .
Aside from the lack of queues and incredible bottom filling pints of beer (oh yes), one of the first advances the fans will notice is the quality and availability of the free stadium Wi-Fi.	Además de la falta de colas y las increíbles pintas de cerveza que llenan el fondo (oh sí), uno de los primeros avances que notarán los fanáticos es la calidad y disponibilidad del Wi-Fi gratuito del estadio.	Además de la falta de colas y las increíbles pintas de cerveza (¡así es!), uno de los primeros avances que notarán los aficionados es la calidad y disponibilidad de Wi-Fi gratuita en el estadio.	Además de que no se verán filas y de las increíbles pintas de cerveza (oh, sí), uno de los primeros avances que notarán los aficionados es la calidad y la disponibilidad del wifi gratuito del estadio.	Además de la eliminación de las colas y los increíbles vasos de cerveza que se llenan por abajo (¿original, no?), una de las principales mejoras que los aficionados notarán es la calidad y la disponibilidad del Wi-Fi gratuito del estadio.	Además de la ausencia de filas, y las increíbles pintas de cerveza que se llenan por abajo (así es), uno de los primeros avances que los fanáticos van a notar es la calidad y la disponibilidad de la red Wi-Fi gratuita del estadio.

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
This includes the Core Wired Network infrastructure that enables the stadium's critical services, such as CCTV , Building Management Systems, audio visual, ticketing and much more... to function.	Esto incluye la infraestructura de la red alámbrica central que permite que los servicios críticos del estadio, como CCTV , sistemas de gestión de edificios, audiovisuales, venta de entradas y mucho más... funcionen.	Esto incluye la infraestructura de la red cableada central que permite el funcionamiento de los servicios fundamentales del estadio, como la CCTV , los sistemas de administración de edificios, los sistemas audiovisuales y la venta de entradas, entre otros.	Esto incluye la infraestructura de red central cableada que permite que los servicios críticos del estadio, como CCTV , los sistemas de gestión de edificios, los sistemas audiovisuales, de venta de entradas y muchos otros funcionen de manera adecuada .	Esto incluye la infraestructura de red cableada central que permite que funcionen los servicios críticos del estadio, como el circuito cerrado de televisión (CCTV) , los sistemas de administración del edificio, los recursos audiovisuales, la venta de entradas y mucho más.	Esto incluye la infraestructura Core Wired Network , que permite el funcionamiento de los servicios críticos del estadio, como el Circuito Cerrado de Televisión , los Sistemas de Gestión del Edificio (BMS), el equipamiento audiovisual, la venta de entradas y mucho más.
We all know refreshments are an important element of sport and the stadium has 65 food and drink outlets , including its own microbrewery and Michelin-starred cuisine .	Todos sabemos que los refrescos son un elemento importante del deporte y el estadio cuenta con 65 locales de comidas y bebidas , incluida su propia microcervecería y cocina con estrellas Michelin .	Sabemos que la comida es un elemento importante del deporte y el estadio cuenta con 65 puntos de venta de comidas y bebidas, incluida su propia microcervecería y cocina con estrellas Michelin .	Todos sabemos que las bebidas son un elemento importante del deporte y el estadio cuenta con 65 locales de comida y bebida, que incluye su propia microcervecería y una cocina con estrellas Michelin	Todos sabemos que los refrigerios son un elemento importante de los deportes y el estadio cuenta con 65 locales de comidas y bebidas, que incluyen su propia microcervecería y una cocina con estrellas Michelin .	Ya sabemos que la comida y la bebida son elementos importantísimos para la experiencia deportiva , y el estadio cuenta con 65 quioscos , su propia microcervecería y una gastronomía premiada con estrellas Michelin .
The effectiveness of retail outlets versus footfall can be better understood and revenue opportunities optimized through planning and proactive notifications based on known interests.	La efectividad de los puntos de venta minoristas frente a las pisadas se puede entender mejor y las oportunidades de ingresos se pueden optimizar mediante la planificación y notificaciones proactivas basadas en intereses conocidos.	La efectividad de los puntos de venta se puede entender mejor y las oportunidades de ingresos se pueden optimizar mediante la planificación y notificaciones proactivas basadas en intereses conocidos.	La eficacia de los puntos de venta frente a la afluencia de público puede entenderse mejor, y de esta forma, las oportunidades de ingresos se pueden optimizar mediante la planificación y las notificaciones proactivas basadas en intereses conocidos.	La efectividad de los puntos de venta en comparación con el tráfico de clientes permite comprender mejor los procesos y aumentar las oportunidades de ingresos mediante planeamiento y avisos proactivos basados en los intereses conocidos.	La efectividad de los stands de ventas en contraste con el juego fútbol puede ser mejor comprendida para optimizar oportunidades de lucro por medio de planeamiento y notificaciones proactivas basadas en intereses conocidos.

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
Also, if staff notice the queue for one toilet is far larger than others, they can use the intelligent signage displays at the stadium and notifications in the app to direct fans to alternatives.	Además, si el personal nota que la cola para un inodoro es mucho más grande que otras, puede usar las pantallas de señalización inteligente en el estadio y las notificaciones en la aplicación para dirigir a los fanáticos a las alternativas.	Además, si el personal nota que la cola para un inodoro es mucho más grande que otras, puede usar las pantallas de señalización inteligente en el estadio y las notificaciones en la aplicación para dirigir a los aficionados a las alternativas.	Además, si el personal nota que la fila para un baño es mucho más larga que otras, puede usar las pantallas de señalización inteligente del estadio y las notificaciones de la aplicación para dirigir a los aficionados a otras alternativas.	Además, si el personal observa que la cola para un baño es mucho más larga que las de los demás, puede usar las pantallas de señalización inteligente en el estadio y avisos en la aplicación para dirigir a los aficionados a baños alternativos.	Además, si el personal nota que la fila de uno de los sanitarios está mucho más larga que las otras, puede usar los carteles digitales inteligentes del estadio, y las notificaciones de la app, para enviar a los fanáticos hacia otras opciones.

Apêndice H

Exemplos de simplificação e interferência T2

Exemplos de simplificação

Exemplos de explicitação

Exemplo de interferência

Exemplo de alteração de ordem da frase

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
Insight Online is a cloud-based infrastructure management and support feature powered by HPE remote support technology.	Insight Online es una característica de soporte y administración de infraestructura basada en la nube con tecnología de soporte remoto HPE.	Insight Online es un recurso de administración y soporte de infraestructura basado en la nube, con la tecnología de soporte remoto de HPE.	Insight Online es una función de gestión y soporte para infraestructura basada en la nube que funciona con la tecnología de soporte remoto HPE.	Insight Online es un recurso de administración de infraestructura y soporte basado en la nube que utiliza tecnología de soporte remoto de HPE.	Insight Online es un recurso de soporte y administración de infraestructura basado en la nube con la tecnología de soporte remoto de HPE.
Track: service events and support cases	Seguimiento: eventos de servicio y casos de soporte	Seguimiento: eventos de servicio y casos de soporte	Realizar un seguimiento de eventos de servicio y casos de soporte	Seguimiento: eventos de servicio y casos de soporte	Rastrear: eventos de servicio y casos de soporte
You will need an HPE Passport ID to sign in.	Necesitará un ID de HP Passport para iniciar sesión.	Necesitará un HPE Passport ID para iniciar sesión.	Necesitará un ID de HP Passport para iniciar sesión.	Necesitará un ID de HPE Passport para iniciar sesión.	Para iniciar sesión necesitará la ID de HPE Passport.

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
HPE Insight RS 7.x central connect software provides 24x7 remote monitoring, diagnosis, and notifications for servers, storage, and network devices.	El software de conexión central HPE Insight RS 7.x proporciona monitoreo, diagnóstico y notificaciones remotos las 24 horas del día, los 7 días de la semana , para servidores, almacenamiento y dispositivos de red.	El software de conexión central HPE Insight RS 7.x proporciona supervisión, diagnóstico y notificaciones remotos, 24x7 , para servidores, almacenamiento y dispositivos de red.	El software de conexión central HPE Insight RS 7.x proporciona supervisión, diagnóstico y notificaciones remotas sin interrupciones , para los servidores, el almacenamiento y los dispositivos de red.	El software HPE Insight RS 7.x Central Connect proporciona monitoreo, notificaciones y diagnóstico remotos 24x7 para servidores, almacenamiento y dispositivos de red.	El software de conexión directa HPE Insight RS 7.x proporciona 24x7 , monitoreo remoto, diagnóstico y notificaciones para servidores, almacenamiento y dispositivos de red.
Reduced risk thanks to continuous remote to identify issues before they become problems.	Riesgo reducido gracias al monitoreo remoto continuo para identificar problemas antes de que se conviertan en problemas.	Riesgo reducido gracias a la supervisión remota continua para identificar inconvenientes antes de que se conviertan en problemas.	Reducción del riesgo gracias a la supervisión remota continua que permite identificar los problemas antes de que aparezcan.	Reducción del riesgo gracias a monitoreo remoto continuo para identificar eventos antes de que se conviertan en problemas.	Se reduce el riesgo gracias al monitoreo remoto permanente para identificar los inconvenientes antes de que se conviertan en problemas.
You can use Insight Online to track service events and support cases, view device configurations, and proactively monitor your HPE contracts and warranties for all your devices monitored by HPE OneView 3.0.	Puede usar Insight Online para rastrear eventos de servicio y casos de soporte, ver configuraciones de dispositivos y monitorear proactivamente sus contratos y garantías de HPE para todos sus dispositivos monitoreados por HPE OneView 3.0.	Puede usar Insight Online para rastrear eventos de servicio y casos de soporte, ver configuraciones de dispositivos y supervisar de forma proactiva sus contratos y garantías de HPE referentes a todos sus dispositivos supervisados por HPE OneView 3.0.	Puede usar Insight Online para realizar un seguimiento de eventos de servicio y casos de soporte, ver las configuraciones de los dispositivos y supervisar de forma proactiva sus contratos y garantías de HPE para todos sus dispositivos controlados por HPE OneView 3.0.	Puede usar Insight Online para rastrear eventos de servicio y casos de soporte, ver configuraciones de dispositivos y monitorear sus contratos y garantías de HPE para todos sus dispositivos supervisados por HPE OneView 3.0.	Puede usar Insight Online para rastrear eventos de servicio y casos de soporte, ver configuraciones de dispositivos y supervisar proactivamente los contratos y garantías de HPE para todos sus dispositivos monitoreados por HPE OneView 3.0.

ORI	MT	PEMT(1)	PEMT(2)	MTNSS	TH
You can view Insight Online from any PC, tablet, or mobile phone browser.	Puede ver Insight Online desde cualquier PC, tableta o navegador de teléfono móvil.	Puede ver Insight Online desde cualquier PC, tablet o navegador de teléfono móvil.	Puede ver Insight Online desde cualquier navegador en su PC, tablet o teléfono móvil.	Puede ver Insight Online desde cualquier navegador de PC, tablet o teléfono móvil.	Puede ver Insight Online desde cualquier navegador en una PC, tablet o teléfono inteligente.
HPE SIM will continue to provide powerful health monitoring , simple and agentless firmware update distribution, and easy integration with plug-ins like HPE Insight Control and HPE Matrix Operating Environment.	HPE SIM continuará brindando un monitoreo de salud potente , distribución de actualizaciones de firmware simple y sin agentes, y una fácil integración con complementos como HPE Insight Control y HPE Matrix Operating Environment.	HPE SIM continuará brindando supervisión de estado robusta, distribución de actualizaciones de firmware simple y sin agentes, y fácil integración con complementos como HPE Insight Control y HPE Matrix Operating Environment.	HPE SIM continuará proporcionando un control poderoso del estado , una distribución de actualización de firmware simple y sin agentes, y una integración fácil con complementos como HPE Insight Control y HPE Matrix Operating Environment.	HPE SIM continuará brindando un monitoreo sólido del estado , una distribución de actualizaciones de firmware simple y sin agentes y una integración fácil con complementos como HPE Insight Control y HPE Matrix.	HPE SIM continuará proporcionando monitoreo de estado de salud , distribución de actualización de firmware sencilla y sin agentes e integración fácil con complementos como HPE Insight Control y HPE Matrix Operating Environment.