



Gisela de Almeida Rodrigues

A TRADUÇÃO AUTOMÁTICA E A PÓS-EDIÇÃO NA RWS

UM ESTUDO DE CASO NA INDÚSTRIA DO VESTUÁRIO

Relatório de Estágio do Mestrado em Tradução, orientado pela Doutora Phillippa May Bennett,
apresentado ao Departamento de Línguas, Literaturas e Culturas da Faculdade de Letras da
Universidade de Coimbra

Outubro de 2021

FACULDADE DE LETRAS

A TRADUÇÃO AUTOMÁTICA E A PÓS-EDIÇÃO NA RWS UM ESTUDO DE CASO NA INDÚSTRIA DO VESTUÁRIO

Ficha Técnica

Tipo de trabalho	Relatório de Estágio
Título	A tradução automática e a pós-edição na RWS
Subtítulo	Um estudo de caso na indústria do vestuário
Autor/a	Gisela de Almeida Rodrigues
Orientador/a(s)	Doutora Phillipa May Bennett
Júri	Presidente: Doutor Jorge Manuel Costa Almeida e Pinho Vogais: 1. Doutor Fernando Gonçalves Ferreira Alves 2. Doutora Phillipa May Bennett
Identificação do Curso	Mestrado em Tradução
Área científica	Tradução
Especialidade/Ramo	Português e uma língua estrangeira (Inglês)
Data da defesa	25-11-2021
Classificação do Relatório	16 valores
Classificação do Estágio e Relatório	17 valores

Agradecimentos

Quero começar por agradecer à minha irmã, à minha mãe e ao meu pai que me acompanharam durante todo o percurso pelo incentivo e apoio, e por me ajudarem a tornar na pessoa que sou hoje.

À Dra. Phillipa May Bennett por me orientar e ajudar durante todo este processo.

À Daniela por fazer-me fazer sentido e se disponibilizar a ajudar-me mesmo quando não podia.

You are awesome!

À Lobo por ter sido a minha companheira de relatório, por me ajudar quando as coisas ficaram apertadas e por ser um dos meus ombros amigos.

À Alice e à Leonor pelas nossas aventuras durante o curso e fora dele, pela vossa ajuda e paciência. Já está!

À minha madrinha de curso, Carolina, por ter sido uma madrinha fantástica.

À Daniela Silvestre por todos os momentos de amizade.

À RWS Portugal e a toda a equipa de *Electronics* por me acolherem de braços abertos e proporcionarem esta incrível oportunidade de estágio.

À Cristina pela disponibilidade, pela orientação durante e depois do estágio e pela confiança depositada em mim.

Ao Cláudio e à Ana por me ajudarem e orientarem a tornar numa melhor profissional.

RESUMO

A introdução de sistemas de tradução automática no mundo profissional veio modificar a tarefa de tradução e o papel do tradutor, exigindo novas formas de trabalhar. Juntamente com a utilização das memórias de tradução, surgiu uma nova prática denominada pós-edição e as empresas de tradução como a RWS começaram a adotá-la com o objetivo de aumentar a produtividade, sem comprometer a qualidade.

Nesse sentido, o presente relatório visa descrever o processo de pós-edição realizado durante o meu estágio curricular na RWS e analisar o desempenho do sistema de tradução automática utilizado nos projetos para um cliente da indústria do vestuário. Este trabalho começa por abordar o conceito de pós-edição e os processos envolvidos, como a utilização de um sistema de tradução automática e memórias de tradução. Posteriormente, apresenta os métodos de pós-edição recomendados pela empresa, bem como o perfil do cliente e os respetivos critérios de qualidade. Tudo isto serve de contexto para o estudo de caso realizado que envolve uma análise das sugestões de tradução automática dos textos trabalhados durante o estágio. Os erros identificados são categorizados segundo a métrica de qualidade SAE J2450, para comparar a qualidade inicial com a qualidade desejada. Os resultados revelam que, embora o desempenho do sistema esteja sempre dependente do tipo de texto, as sugestões de tradução automática não apresentam a qualidade exigida, obrigando o tradutor a intervir e a realizar a tarefa de pós-edição.

Palavras-chave: pós-edição, tradução automática, memórias de tradução, qualidade

ABSTRACT

The emergence of machine translation systems in the professional world has changed translation processes and the role translators have to play, requiring new ways of working. Alongside the use of translation memories, post-editing practices developed and translation companies like RWS began to implement this new practice into their environment, to increase productivity without compromising quality.

In this regard, this project aims to outline the post-editing process performed during my curricular internship at RWS and investigate the performance of the machine translation system applied in the projects for a client in the clothing industry. It begins by addressing the post-editing concept and the processes involved, such as the use of machine translation systems and translation memories. The next section of my project introduces the post-editing methods recommended by the company, as well as the client profile and their respective quality criteria. This will provide the context for the study case which involves an analysis of the machine translation suggestions for the projects translated during the internship. Errors will be categorized according to the quality metrics SAE J2450, to compare the initial quality to the required quality. The results show that, while the performance of the system is always dependent on the type of text, the suggestions made by the system used do not have the quality required, forcing the translator to intervene and perform post-editing.

Keywords: post-editing, machine translation, translation memories, quality

Índice

Introdução.....	1
1. O estágio curricular	3
1.1. Motivação para realizar o estágio.....	3
1.2. Caracterização da entidade de acolhimento	3
1.2.1. A antiga SDL, agora RWS	3
1.2.2. A RWS Portugal.....	4
1.3. O estágio.....	5
1.3.1. As formações iniciais	6
1.3.2. O trabalho realizado	7
1.3.3. O <i>feedback</i> e as avaliações	8
1.3.4. As expectativas.....	9
1.4. Os programas e ferramentas CAT utilizados.....	9
1.4.1. <i>SDL Trados Studio 2019</i>	10
1.4.2. <i>SDL Trados GroupShare 2019</i>	11
1.4.3. <i>SDL Machine Translation</i>	12
1.4.4. <i>Timesheet</i>	12
1.4.5. <i>Tasks</i>	13
1.4.6. Outros programas (<i>Xbench, Xliff Compare</i>).....	13
2. A pós-edição.....	15
2.1. Conceito	16
2.2. A tradução automática.....	16
2.2.1. Contexto da tradução automática	17
2.2.2. Tipos de tradução automática.....	18
2.2.3. O panorama da tradução automática.....	20
2.2.4. Obstáculos à tradução automática	20
2.3. As memórias de tradução	21
2.3.1. A utilização de memórias de tradução.....	23
3. Pós-edição na RWS.....	25
3.1. Perfil do Cliente A.....	26
3.1.1. O cliente e a pós-edição.....	27
3.1.2. Orientações do cliente	28
3.2. Pós-edição de tradução automática	32
3.2.1. Orientações para a pós-edição de tradução segundo a RWS.....	33
3.3. As métricas de qualidade.....	34
3.3.1. O sistema de qualidade da RWS (SAE J2450).....	34
4. Estudo de caso	41
4.1. Metodologia.....	41
4.2. Os erros da tradução automática neuronal.....	42

4.2.1.	Projetos de descrição de produtos	42
4.2.2.	Projetos de descrição de produtos e <i>marketing</i>	49
4.2.3.	Projetos de artigos e entrevistas	53
4.3.	Análise final	61
5.	Da tradução automática à qualidade desejada	63
5.1.	A edição das sugestões da tradução automática	63
	Conclusão	66
	Bibliografia.....	68

Índice de diagramas

Diagrama 1 - Cronologia da SDL no mercado de tradução	4
Diagrama 2 - Processo de tradução/pós-edição realizado durante o estágio.....	8
Diagrama 3 - Tipos de projetos do cliente	27
Diagrama 4 - Orientações do cliente	29
Diagrama 5 e 6 - Exemplos de termos homónimos	31

Índice de figuras

Figura 1 - Escala de qualidade das avaliações	9
Figura 2 - Registo da tarefa no Timesheet	12
Figura 3 - Informações da Tasks	13
Figura 4 - Exemplo de um relatório de comparação do Xliff Compare.....	15
Figura 5 - Pesquisa de concordance no SDL Trados Studio 2017.....	24
Figura 6 - Pesquisa concordance pouco específica	25
Figura 7 - Pesquisa concordance mais específica	25
Figura 8 - Memórias de tradução de um projeto	64
Figura 9 - Exemplo de tradução automática neuronal aplicada	64

Índice de gráficos

Gráfico 1 - Trabalho realizado durante o estágio	7
Gráfico 2 - Comparação dos projetos realizados	26
Gráfico 3 - Comparação de tipos de textos traduzidos/editados	27
Gráfico 6 - Erros identificados no projeto.....	43
Gráfico 7 - Erros identificados no projeto.....	47
Gráfico 8 - Erros identificados no projeto.....	50
Gráfico 9 - Erros identificados no projeto.....	52
Gráfico 10 - Erros identificados no projeto.....	54
Gráfico 11 - Erros identificados no projeto.....	57
Gráfico 12 - Total de erros identificados	62

Índice de tabelas

Tabela 1 - Sugestões sobre como evitar o uso de género.....	30
Tabela 2 - Alternativas ao uso do gerúndio.....	30
Tabela 3 - Exemplos do uso de conjunções no início da frase.....	32

Tabela 4 - Exemplos de como evitar a utilização de pronomes pessoais.....	32
Tabela 5 - Exemplo de alterações preferenciais.....	33
Tabela 6 - Categorias de erros de acordo com SAE J2450	35
Tabela 7 - Exemplos de erros de terminologia.....	37
Tabela 8 - Exemplo de significado errado	37
Tabela 9 - Exemplos de erros ortográficos.....	38
Tabela 10 - Exemplos de erros estilísticos	39
Tabela 11 - Gravidade dos erros consoante a categoria.....	40
Tabela 12 - Dados do projeto	42
Tabela 13 - Erros de tradução literal	44
Tabela 14 - Erros segundo as orientações do cliente	45
Tabela 15 - Dados do projeto	46
Tabela 16 - Exemplo de significado errado	47
Tabela 17 - Dados do projeto	48
Tabela 18 - Exemplo de erros identificados.....	48
Tabela 19 - Dados do projeto	49
Tabela 20 - Exemplo de erro estilístico e tradução incorreta.....	51
Tabela 21 - Dados do projeto	51
Tabela 22 - Exemplos de tradução literal.....	53
Tabela 23 - Dados do projeto	54
Tabela 24 - Exemplo de tradução literal	55
Tabela 25 - Outros exemplos	56
Tabela 26 - Dados do exemplo.....	57
Tabela 27 - Erros identificados	59
Tabela 28 - Exemplos de erros identificados	60

Termos e definições

Autopropagação	Funcionalidade do <i>SDL Trados Studio</i> que permite traduzir de forma automática todos os segmentos idênticos no mesmo projeto.
Context matches/ correspondência de contexto	Acontece quando o segmento, para além de ser idêntico, encontra-se no mesmo contexto que o segmento presente na memória de tradução.
Ferramentas CAT	Ferramentas de apoio ao processo de tradução.
Fragment matching	Funcionalidade do <i>SDL Trados Studio</i> que reconhece a existência de fragmentos do segmento semelhantes nas memórias de tradução e sugere a tradução armazenada.
Fuzzy matches	Segmentos que apresentam alguma semelhança com os segmentos guardados nas memórias de tradução e precisam de ser modificados para corresponder ao texto de partida.
MQM	<i>Mutidimensional Quality Metric</i>
NMT	Tradução automática neuronal
QA	<i>Quality Assurance</i>
SAE J2450	Métrica de qualidade de tradução
Texto de chegada	Texto traduzido na língua alvo
Texto de partida	Texto original que vai ser traduzido

Introdução

O presente relatório de estágio tem dois objetivos: (i) descrever a utilização da tradução automática e o processo de pós-edição na RWS¹, focando-se num cliente da indústria do vestuário, doravante designado como Cliente A, trabalhado durante o estágio curricular e; (ii) analisar o desempenho do sistema de tradução automática nos projetos desse mesmo cliente, categorizando os erros identificados e comparando a sua qualidade inicial com a qualidade desejada pela empresa e pelo cliente.

Dada a frequência da tarefa de pós-edição de sistemas de tradução automática durante o estágio com a empresa RWS, decidi abordar o tema no contexto deste relatório de estágio, uma vez que a tradução automática e o seu uso não foram explorados no primeiro ano de mestrado. Embora o curso de mestrado tenha oferecido as bases necessárias para traduzir, as tarefas realizadas durante o estágio exigiram diferentes capacidades para trabalhar com todas as ferramentas de apoio à tradução. Assim, este relatório pretende abordar o tema de pós-edição e analisar os diferentes passos necessários para chegar ao produto final. Além disso, incluirá um estudo de caso que se irá focar no trabalho realizado para o Cliente A² durante o estágio, mais especificamente nos resultados de tradução automática e o processo de pós-edição envolvido.

No primeiro capítulo será apresentada a instituição onde o estágio se realizou, bem como os detalhes do estágio, o trabalho e as formações realizados, e os programas e ferramentas de apoio à tradução utilizados. De seguida, será explicado o que é a pós-edição, introduzindo os sistemas de tradução automática e memórias de tradução, fundamentais ao processo. Além disso, o perfil do cliente também será abordado, uma vez que é importante para definir a qualidade final do texto. No quarto capítulo, far-se-á uma análise do desempenho do sistema de tradução automática *SDL Machine Translation* utilizado durante o estágio, identificando os erros presentes e comparando a qualidade inicial com a qualidade desejada do produto final. Com a realização deste estudo de caso espera-se descobrir quais são os erros mais comuns do sistema de tradução automática utilizado, se a qualidade das suas sugestões é suficiente para produzir um texto de qualidade desejada e se algum dos exemplos analisados exigiu mais

¹ O estágio foi realizado na SDL Portugal. Durante a realização do estágio, a SDL passou a fazer parte do grupo RWS. Assim, sempre que a empresa for mencionada, será utilizado o nome RWS ou RWS Portugal, de forma a manter o presente relatório atualizado.

² Por motivos de confidencialidade, nenhum nome será mencionado e qualquer termo que possa levar à identificação do cliente será alterado.

esforço de edição. Por último, será descrito o processo de pós-edição e os passos a seguir para editar as sugestões de tradução automática e concluir um projeto.

1. O estágio curricular

1.1. Motivação para realizar o estágio

O Mestrado em Tradução da Faculdade de Letras oferece o estágio curricular como uma das propostas para concluir o curso. O estágio é uma boa opção para alargar os conhecimentos e a perceção sobre a tradução profissional, bem como ganhar experiência na área. Também proporciona oportunidades para entrar em contacto com tradutores profissionais, entender como o mercado da tradução funciona no mundo real e explorar diferentes especialidades para além daquelas que se podem estudar durante o primeiro ano de mestrado. Até à realização do estágio não tinha tido qualquer experiência de tradução profissional. Uma vez que durante o curso de mestrado realizei apenas traduções de áreas como medicina, direito e literatura, parte da minha decisão para realizar um estágio curricular assentava na vontade de trabalhar numa empresa que lidasse com outros domínios de tradução. O estágio permitiu realizar outras tarefas, como a pós-edição, traduzir vários tipos de textos e reconhecer o funcionamento da tradução profissional atual.

A RWS Portugal, antiga SDL Portugal, foi a identidade escolhida para realizar o estágio, uma vez que é uma empresa multinacional que trabalha em variadas áreas de tradução e pares de línguas.

1.2. Caracterização da entidade de acolhimento

1.2.1. A antiga SDL, agora RWS

A antiga SDL (*Software and Documentation Localization*), agora parte do grupo RWS, era uma empresa global de tecnologia e serviços linguísticos com sede em Maidenhead, no Reino Unido. Foi fundada em 1992 por Mark e Cristina Lancaster e, a janeiro de 2021, contava com mais de 63 escritórios em 39 países diferentes e mais de 4300 profissionais (SDL plc, s.d.).

A empresa prestava os seus serviços a um grupo global de clientes, de vários setores, incluído o setor automóvel, médico e farmacêutico, bem como telecomunicações, entre outros (Davies, 2021) e pretendia ajudar a criar, traduzir e a distribuir conteúdos globalmente, com os seus serviços linguísticos, tecnologias linguísticas e conteúdo tecnológico (SDL plc, s.d.).

A SDL cresceu através de aquisições de *software*, como o *Trados*, e empresas, como o *Language Weaver*. A Figura 1 representa a cronologia da empresa até à atualidade e inclui alguns dos momentos mais importantes da entidade, como a aquisição do *software* Trados, que veio mais tarde a originar o *SDL Trados Studio*, ou o lançamento do próprio sistema de tradução automática.

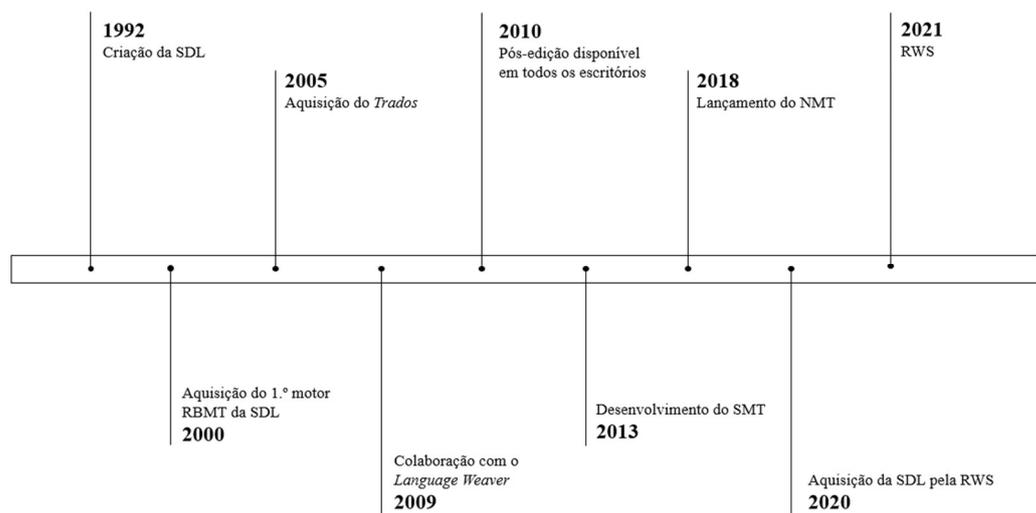


Diagrama 1 - Cronologia da SDL no mercado de tradução

A novembro de 2020, foi anunciada a aquisição da SDL por parte da RWS, outra grande empresa de serviços de tradução. Esta aquisição veio beneficiar ambas as empresas e os consumidores, combinando as capacidades de localização e tradução da RWS com o *software*, a tradução automática e a inteligência artificial da SDL (Davies, 2021; RWS, 2020).

Atualmente, as tecnologias linguísticas da RWS são usadas por grandes organizações e empresas. Algumas das ferramentas tecnológicas incluem o *SDL Trados Studio*, uma ferramenta de apoio à tradução, o *SDL Machine Translation*, um sistema de tradução automática neuronal, o *SDL MultiTerm* e o *SDL Trados GroupShare* (SDL plc, s.d.).

1.2.2. A RWS Portugal

A SDL Portugal foi criada a partir do departamento de português com escritório em Granada, Espanha, mas baseou-se em Portugal apenas em 2016, no Porto. Atualmente, enquanto RWS, conta com cerca de 40 trabalhadores, em 4 equipas diferentes: a equipa de *Life Sciences*, que se especializa na área da medicina e farmacêutica, a equipa de *Industrial Manufacturing & Automotive*, onde se trabalha com documentos pertencentes ao setor automóvel, a equipa de *Electronics* que trata de documentos de natureza técnica, de *software* e de *marketing*, e por

último a equipa de *High-Tech*, uma equipa um pouco mais recente dedicada a empresas específicas de tecnologia.

Para além das equipas, existem diferentes cargos dentro do escritório. Cada tradutor interno tem uma função diferente, com o respetivo nível, classificado de P2 a P7, sendo P2 o nível mais baixo. P2 ou *translator intern*, é o nível dado aos tradutores estagiários. O principal objetivo deste nível é desenvolver competências profissionais, adquirir experiência de trabalho no mundo real da tradução e ter a oportunidade de contribuir para os objetivos da equipa e dos projetos (Clarke, 2019). Para além dos níveis de tradução, existe também o nível de gestão, onde se inserem os *line managers* e os *language office directors* (Career Framework Library, 2019).

1.3. O estágio

Aquando do processo de escolha de uma entidade para realizar o estágio, a RWS não tinha uma parceria estabelecida com a Universidade de Coimbra. No entanto, foi possível recolher algum *feedback*, uma vez que antigos alunos do curso também tinham tido a oportunidade de estagiar na empresa. A candidatura foi relativamente simples: foi requisitada a elaboração de um pequeno teste de admissão, no qual era necessário traduzir um texto de *marketing* e um texto técnico, e responder a duas perguntas em português. Nos dias seguintes, foi realizada uma entrevista por telefone a explicar como tinha corrido o teste, como iria decorrer o estágio e quais as prioridades da empresa para com o estagiário. Além disso, também foram acordados os detalhes sobre o início do estágio, bem como a entrega do equipamento necessário.

O meu estágio teve início no dia 2 de setembro de 2020 e terminou a 17 de dezembro de 2020, completando um total de 496 horas. Foi combinado não trabalhar durante as sextas-feiras, uma vez que era um dia de aulas.

O horário de trabalho era bastante flexível e permitia aos tradutores e estagiários adaptarem-no consoante o seu trabalho. Pude escolher começar a trabalhar entre as 8h30 e as 9h30 e tinha de completar 8 horas de trabalho, o que significa que a partir das 17h poderia terminar o meu dia. Também o horário do almoço era flexível, podendo escolher almoçar entre as 12h30 e as 14h30. Além disso, a empresa dava a possibilidade de se completar 15 minutos adicionais todos os dias, de forma a sair uma hora mais cedo todas as sextas-feiras.

Como resposta à pandemia de COVID-19, a RWS deu prioridade à segurança de todos, colocando os seus colaboradores em regime de teletrabalho. Os materiais necessários para se poder trabalhar em casa, nomeadamente um portátil da empresa, foram disponibilizados e

remetidos por correio e todas as comunicações foram feitas remotamente, através da plataforma *Microsoft Teams*.

O estágio foi realizado com a equipa de *Electronics*, uma equipa de 9 pessoas, na altura. Embora fosse designada como a equipa de eletrónica, muitos dos projetos realizados tinham uma vertente de *marketing*. Além disso, a equipa recebia vários projetos por dia, normalmente, de pequena e média dimensão.

Antes de o estágio começar, foi entregue um plano de estágio formulado pela empresa, que incluía o programa e descrevia os objetivos a cumprir e as competências a desenvolver pelo estagiário durante o período de estágio. O principal objetivo era que o estagiário ficasse a conhecer a empresa e o seu funcionamento, se familiarizasse com as várias ferramentas de apoio à tradução e tivesse uma experiência de tradução variada e intensiva.

1.3.1. As formações iniciais

Durante o estágio, foram cumpridas cerca de 27 horas de formação iniciais. A empresa oferecia várias plataformas para a realização e acompanhamento das mesmas, nomeadamente o *MyLX* e o *Atlas*. O *Atlas* é uma plataforma que orienta os estagiários e novos membros da RWS, focando-se na aprendizagem e na introdução da empresa. Esta plataforma de formação encontra-se dividida em três módulos que abordam o funcionamento da empresa, desde os serviços à gestão da qualidade e segurança, para facilitar a adaptação inicial, e inclui formações com documentos importantes, vídeos educativos e pequenos questionários para testar aquilo que tinha sido aprendido. O *MyLX* é uma plataforma que contém vários tipos de formações focadas em diversos temas e áreas, desde segurança e ética de trabalho até o trabalho em equipa, pós-edição e ferramentas CAT. Por último, foi disponibilizado um documento chamado *Induction Checklist*, o qual apresentava o planeamento das formações do estágio que precisavam de ser feitas durante as semanas e meses iniciais. Este plano incluía também planos para formações sobre o *SDL Trados Studio* e outras ferramentas necessárias no dia a dia, como o *SDL Trados Groupshare*, o *Timesheet*, o *Xliff Compare* e o *Xbench*.

Outra formação realizada e relevante, tanto para o estágio como para este estudo de caso, estava relacionada com a pós-edição. Visto que a pós-edição era crucial no trabalho da RWS, era necessário ensinar os novos tradutores a pós-editarem e a trabalharem com as ferramentas de apoio à tradução de forma a aumentarem a sua produtividade.

1.3.2. O trabalho realizado

Durante o período de estágio, foram realizados projetos de pós-edição e tradução de diversos temas dentro da área de eletrónica e *marketing*. Contudo, foram realizados, em média, mais projetos de pós-edição e os projetos de tradução foram substancialmente menos comuns. Enquanto *translator intern*, o trabalho realizado abrangeu apenas produção, permitindo produzir e entregar, por norma, seis a oito projetos por dia³. Todos os projetos realizados eram de pequena e média dimensão, não ultrapassando as 2000 palavras. A maioria dos projetos incluía quase sempre tradução automática (NMT) integrada e era acompanhada por memórias de tradução da empresa ou fornecidas pelo cliente. No total, foram realizados cerca de 234 projetos com 119 513 palavras (novas e *fuzzy matches*). Destes, 214 foram de pós-edição e apenas 20 de tradução⁴ (Gráfico 1).

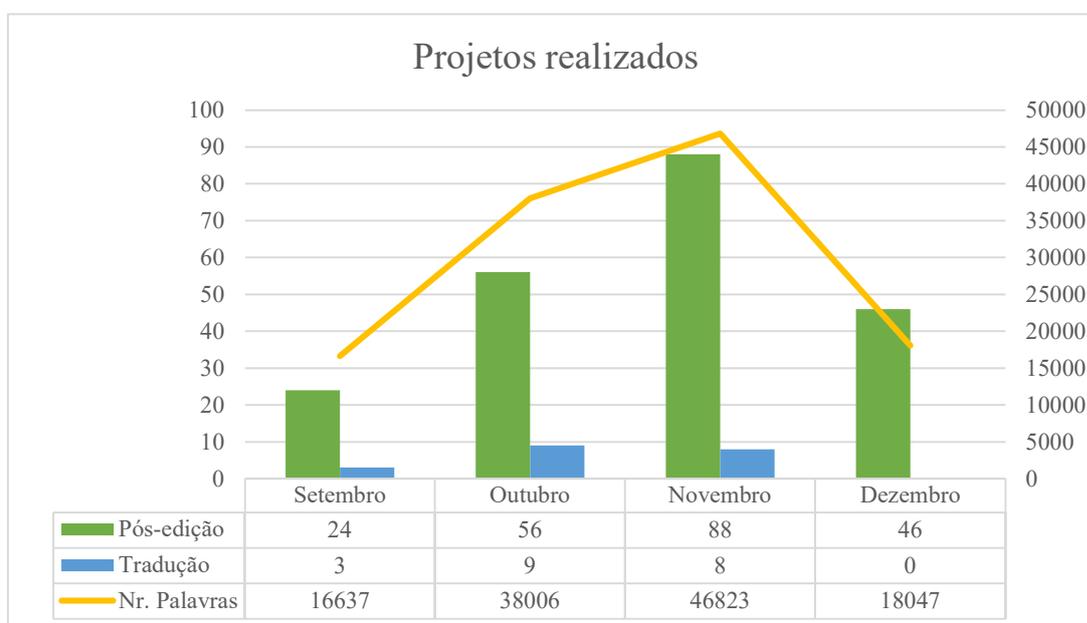


Gráfico 1 - Trabalho realizado durante o estágio

O Diagrama 2 demonstra o processo de tradução realizado. Dentro da equipa, todos os projetos eram atribuídos de acordo com a disponibilidade do tradutor ou do estagiário. Enquanto os tradutores recebiam, normalmente, os projetos de clientes com que estavam familiarizados, eu fui recebendo diversos projetos para poder conhecer um pouco de todas as contas.

³ Este número será considerado o ritmo padrão do trabalho realizado durante o estágio no presente relatório.

⁴ Durante o mês de setembro foram realizados menos projetos, pois foi o mês de adaptação. Os projetos realizados em dezembro correspondem apenas à primeira quinzena do mês, até à conclusão do estágio.

Normalmente, eram recebidos projetos para o dia e para o dia seguinte, visto que a maioria era criada nesse mesmo dia e tinha um prazo para o fim do dia ou para o dia seguinte.

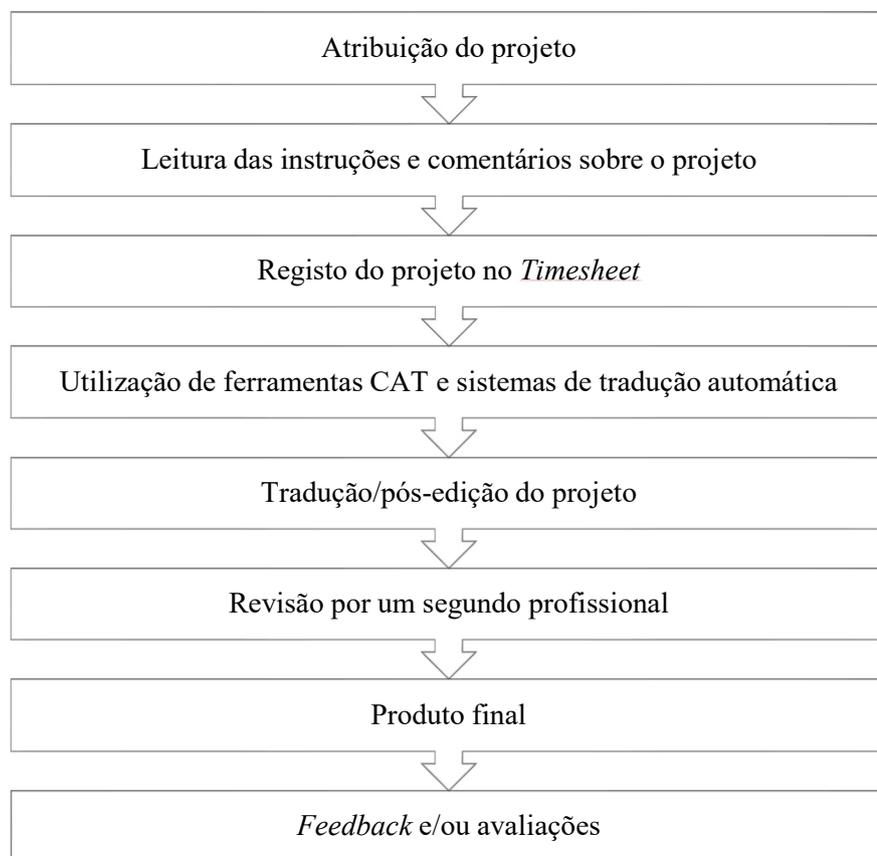


Diagrama 2 - Processo de tradução/pós-edição realizado durante o estágio

Através de todo o trabalho realizado, observou-se que, sendo a pós-edição uma tarefa comum, são necessárias ferramentas capazes de ajudar o tradutor a aumentar o seu nível de produtividade, ao mesmo tempo que mantém o nível de qualidade elevado. Essas ferramentas incluem um sistema de tradução automática, aplicado antes de o projeto ser enviado ao tradutor, e as memórias de tradução utilizadas durante o processo de tradução sobretudo para editar as sugestões da tradução automática e manter a consistência do vocabulário. Assim, de forma a manter a qualidade pretendida e a ajudar a melhorar o trabalho feito, foram realizadas avaliações mensais, assim como era recebido *feedback* frequente.

1.3.3. O *feedback* e as avaliações

O *feedback* foi uma parte integrante do estágio. Como tal, foram recebidos comentários e observações, assim como realizadas algumas avaliações internas, em concordância com os parâmetros da RWS, por membros da empresa. Os relatórios de comparação foram talvez o documento mais útil para me ajudar a verificar os erros e entender como poderia melhorar em

projetos futuros. Estes eram enviados com todas as alterações feitas pelo revisor e a respetiva explicação e serviam apenas como complemento para a minha melhoria enquanto estagiária e futura tradutora.

Já as avaliações internas eram realizadas mensalmente pelos revisores das traduções e avaliavam o trabalho e a qualidade dos trabalhos produzidos. Permitiam aos *line managers* monitorizar se o tradutor fazia o seu trabalho consoante os padrões de qualidade da RWS. As avaliações seguiam a métrica de qualidade SAE J2450. Todos os erros encontrados eram avaliados pela categoria de erro e a gravidade da mesma. Depois de identificados e somados os tipos de erros, o projeto era pontuado numa escala de 0 a 7 (ver Figura 1), sendo as traduções de nível 0 consideradas de qualidade abaixo do padrão e de nível 7 as traduções de qualidade acima do padrão. Todos os erros identificados eram acompanhados da respetiva correção e o motivo pelo qual foram considerados problemas.



Figura 1 - Escala de qualidade das avaliações

Todo o *feedback* e sugestões recebidos ajudaram-me a melhorar não só o trabalho realizado, mas também as estratégias de escrita da RWS Portugal e as minhas competências linguísticas.

1.3.4. As expectativas

Comecei o estágio com a expectativa de que iriam ser realizadas maioritariamente traduções de raiz, mas, contrariamente ao esperado, a pós-edição foi mais comum. Foi durante o mesmo que aprendi o conceito de pós-edição e a realidade na RWS, bem como o novo papel do tradutor que, devido aos avanços tecnológicos das memórias de tradução e às traduções automáticas, modifica para pós-editor.

Com o estágio, desenvolvi as minhas capacidades criativas, principalmente de *marketing* e publicidade, para as quais pensara não ter qualquer talento. Aprendi a gostar deste lado mais criativo da tradução e fui desafiada a sair da minha zona de conforto enquanto tradutora.

1.4. Os programas e ferramentas CAT utilizados

As ferramentas CAT são ferramentas de apoio à tradução que facilitam e automatizam o processo e foram concebidas para ajudar especialmente na tradução do conteúdo digital (Garcia, 2009; Kurniawati, *et al.*, 2016). Permitem reutilizar traduções já feitas de forma a manter a

consistência e têm como objetivo aumentar a produtividade, bem como a rapidez do tradutor, enquanto mantêm uma qualidade elevada dos serviços prestados (Alotaibi, 2020; Garcia, 2014; SDL Limited, 2018; Vela, *et al.*, 2019). As ferramentas incluem funcionalidades como memórias de tradução, glossários e gestão de terminologia, pesquisa de *concordance*, sugestões de *auto-completion* e muito mais (Vela, *et al.*, 2019, p.9).

Estes sistemas começaram a ser desenvolvidos no início dos anos 80 e vieram marcar a transição da tradução tradicional para o mundo digital (Alotaibi, 2020, p.2). Não só mudaram o nível de produção, mas também o ambiente de trabalho dos tradutores. Essencialmente, os sistemas passaram a dividir o texto em “segmentos” (frases definidas por pontuação) e procurar segmentos idênticos ou semelhantes nas memórias de tradução (Garcia, 2014, p.68). Esta nova funcionalidade permitiu simplificar o processo de pesquisa durante a tradução, uma vez que o texto é guardado em frases, conseguindo manter a consistência do estilo e do vocabulário do projeto e do cliente em questão (Han, 2020, p.5).

Na RWS, as ferramentas CAT são utilizadas em conjunto com os sistemas de tradução automática, alterando assim as prioridades do tradutor, cujas funções giram agora em torno da pós-edição e da consistência das suas traduções. Além disso, os serviços foram transmitidos para *software* com base na nuvem, sendo comum partilhar e publicar os projetos em sistemas baseados na nuvem que disponibilizam uma gestão de projetos de tradução mais fácil (Alotaibi, 2020, p.1). Isto significava que os documentos podiam ser acedidos em qualquer dispositivo, bem como guardados em tempo real para que todos tivessem acesso ao mesmo conteúdo (Han, 2020, pp.4-5).

A utilização da tecnologia na área da tradução tem-se tornado cada vez mais indispensável no processo da tradução. As tecnologias vieram afetar o local de trabalho dos tradutores (O’Hagan, 2013, p.503). Durante o estágio foi possível comprovar que as ferramentas CAT eram imprescindíveis para realizar os projetos, mesmo nos que não utilizavam sistemas de tradução automática. Nos capítulos seguintes (1.4.1. a 1.4.6) encontram-se os programas utilizados durante o estágio e uma breve descrição das suas funcionalidades.

1.4.1. *SDL Trados Studio 2019*

O *SDL Trados Studio 2019* foi uma das principais ferramentas utilizada exclusivamente no apoio à tradução durante o estágio. Esta versão foi criada para melhorar a experiência do utilizador com funcionalidades que ajudam a aumentar os níveis de produtividade de toda a equipa de tradução e permite realizar traduções de alta qualidade.

O *SDL Trados Studio* teve origem em 1984, quando Jochen Hummel e Iko Knyphausen tornaram o *Trados* num *Language Service Provider*. No entanto, foi só em 1992 que a primeira versão do *Translator's Workbench* foi lançada, uma aplicação de *software* simples que tinha como objetivo ajudar os tradutores. O *Trados Translator's Workbench* era utilizado principalmente para trabalhar com documentos técnicos ou componentes de *software* por empresas e freelancers (Arenas & Moorkens, 2019, p.218). Em 2009, a SDL lançou uma versão capaz de oferecer várias funcionalidades de tradução num só *software* (Polo, 2013, p.164). Desde então, o programa tem sido atualizado a cada 2 anos, garantindo uma melhoria no serviço em relação à primeira edição.

Aquando da redação deste relatório, o modelo mais atual é o *SDL Trados Studio 2021*, lançado em agosto de 2020. Durante o estágio, esta versão começou a ser testada por muitos tradutores, mas o *SDL Trados Studio 2019* foi a única versão utilizada.

O *SDL Trados Studio 2019* foi lançado em julho de 2018 (Ghislandi, 2018). Inclui funcionalidades como o *AutoSuggest*, o *AutoCorrect*, a pesquisa de *concordance* e muito mais. Conta também com uma gestão das memórias de tradução *Quality Assurance (QA)*. O *software* trabalha em conjunto com o *SDL GroupShare* para que se possam partilhar projetos e documentos *online* e em tempo real. Também oferece várias possibilidades para realizar projetos *offline*, disponibilizando a opção para partilhar apenas os *xliff* dos projetos traduzidos. Além disso, sempre que os projetos não se encontravam no *Groupshare*, eram criados pacotes que continham os projetos e todo o material necessário para o completar, como as memórias de tradução e os documentos de referência. Através destes pacotes, era capaz de partilhar com o revisor, não só o projeto concluído, mas também as memórias de tradução atualizadas, por exemplo.

1.4.2. *SDL Trados GroupShare 2019*

O *GroupShare* é uma plataforma completamente integrada no *SDL Trados Studio* que requer ligação à internet e permite a publicação dos projetos na nuvem. Isto significa que os projetos podem ser partilhados com vários utilizadores, assim como as suas memórias de tradução e textos de referência. Através desta ferramenta, os profissionais têm a possibilidade de trabalhar em simultâneo nos projetos, podendo atualizar as memórias de tradução em tempo real, garantindo assim a consistência do trabalho. Permite criar, gerir e partilhar projetos de forma segura, assim como personalizar as fases do projeto e seguir o progresso do projeto. O *GroupShare* também permite reduzir a quantidade de emails recebidos, uma vez que envia notificações automáticas dos projetos atribuídos aos membros da equipa (RWS Trados, 2020).

1.4.3. *SDL Machine Translation*

Durante o estágio, a tradução automática neuronal foi disponibilizada através do *SDL Machine Translation Cloud*⁵, um serviço de tradução automatizada baseado na nuvem da SDL (RWS, s.d.). Este serviço é um sistema de tradução automática de autoaprendizagem que pode ser utilizado em conjunto com o *SDL Trados Studio* para ajudar na tarefa de pós-edição, visto ser capaz de produzir traduções sem o auxílio de um tradutor humano (SDL, s.d.-b).

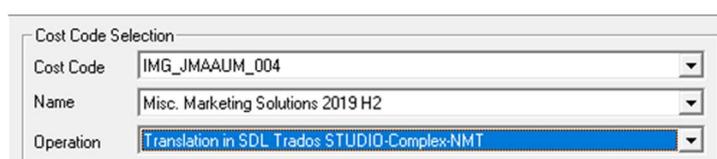
Atualmente, este servidor faz parte do *SDL Language Cloud*, uma plataforma com base na nuvem que oferece não só motores de tradução automática, mas também outros tipos de serviços, como editores *online* (SDL, s.d.-d). Esta plataforma inclui vários tipos de motores de tradução automática, cada um com uma funcionalidade diferente. Desde motores treinados para adaptarem a terminologia e o estilo do tradutor em tempo real com base na pós-edição de cada segmento (*AdaptiveMT*) e dicionários personalizados até motores de tradução automática industrial que oferecem uma tradução automática mais especializada, o *SDL Language Cloud* foi desenhado para gerar melhores resultados de tradução de forma a haver menos pós-edição envolvida (SDL, s.d.-e).

Nem todos os clientes permitiram a utilização da funcionalidade na tradução dos seus projetos. No entanto, a utilização da ferramenta foi aplicada em todos os projetos de pós-edição abordados mais à frente neste estudo de caso.

1.4.4. *Timesheet*

O *Timesheet* é um *software* de monitorização desenvolvido pela SDL. Este programa permite o registo das horas de trabalho pessoais, bem como monitorizar e documentar o tempo investido em cada tarefa (SDL, s.d.-c). O principal propósito deste *software* é calcular a produtividade do tradutor em cada tarefa.

Para registar cada tipo de tarefa é necessário introduzir o código do cliente, o tipo de tarefa que será feita e o número de palavras novas e *fuzzy matches*. É também necessário distinguir se o projeto é, por exemplo, uma pós-edição com NMT ou se é uma tradução simples sem tradução automática (Figura 2).



Cost Code Selection	
Cost Code	IMG_JMAAUM_004
Name	Misc. Marketing Solutions 2019 H2
Operation	Translation in SDL Trados STUDIO-Complex-NMT

Figura 2 - Registo da tarefa no Timesheet

⁵ *SDL Machine Translation* refere-se à antiga marca *BeGlobal* da empresa.

No final, quando um projeto é concluído, basta terminar a tarefa e o próprio programa irá avisar se a tarefa foi realizada dentro do tempo de produtividade pretendido, se o tradutor realizou a tarefa mais rápido do que o estimado ou se demorou mais tempo.

1.4.5. *Tasks*

A *Tasks* é uma folha *Excel online* onde se organizam as tarefas da equipa. Neste documento partilhado encontram-se registados os projetos atribuídos, os respetivos prazos, o número de palavras totais, novas e *fuzzy matches*, o progresso, o tipo de tradução, quem gere a conta do projeto em questão, ou seja, quem está encarregue de receber, atribuir e entregá-los ao cliente, o tradutor, o revisor e, se necessário, observações a ter em conta.

Código		ID Projeto													
Início	Prazo	Hora	Tipo tarefa	Total	Novas	75% - 84%	85% - 99%	100%	Progresso	Lead	Tradutor	PO T	Revisor	PO R	PMO PO
Comments															

Figura 3 - Informações da *Tasks*

Geralmente, depois de um projeto ser atribuído, é enviado um *handover* com todas as instruções dadas pelo cliente e, por vezes, questões resolvidas sobre o projeto. Por outro lado, o gestor da conta pode optar por apenas escrever o que for necessário na coluna dos comentários e disponibilizar o *link* para o projeto, sem precisar de enviar um email com instruções. Nestes casos, o tradutor apenas recebe uma notificação por email do *GroupShare* a informar que o projeto em questão foi atribuído. Quando uma tarefa é terminada, o tradutor deve indicar que a tarefa está feita e escondê-la da secção, de forma a manter o documento relevante e atualizado para toda a equipa.

1.4.6. Outros programas (*Xbench*, *Xliff Compare*)

Para além das ferramentas já mencionadas, a RWS utiliza ferramentas adicionais para ajudar o profissional a desenvolver as suas competências e a produzir a qualidade desejada nos projetos de tradução. Algumas dessas ferramentas utilizadas durante o estágio foram o *Xbench* e o *Xliff Compare*.

O *Xbench* é um *software* que serve de complemento ao *Studio Trados* utilizado para verificar e controlar a qualidade⁶ do projeto e fazer a gestão de terminologia (ApSIC, S.L., 2021). O tradutor tem à sua disponibilidade uma opção no *Studio* que permite verificar a

⁶ Quality Assurance (QA)

qualidade da tradução. Esta funcionalidade, apesar de ser bastante completa e mostrar os problemas encontrados no projeto, desde falta de terminologia a omissão de símbolos, pode demorar algum tempo a analisar o documento na íntegra, principalmente quando estes apresentam um tamanho consideravelmente grande. Assim, os tradutores acabam por optar por um *software* mais rápido como o *Xbench*. O *plug-in* da aplicação pode ser adicionado ao *software* e utilizado como alternativa.

Este programa é utilizado especialmente na última fase de tradução, quando é necessário verificar se existem algumas questões que precisam de ser modificadas. O seu papel é avisar sobre quaisquer erros de terminologia, espaços a mais, palavras repetidas ou números traduzidos de forma diferente do texto de partida, por exemplo.

Outro *software* utilizado é o *Xliff Compare*, uma ferramenta capaz de comparar ficheiros ou projetos e apresentar os resultados das comparações com todas as alterações feitas na revisão da tradução (SDL AppStore, 2021). Existem outros programas semelhantes, como o *Post-edit Compare*, o *ChangeTracker* ou o *Transistor*. Apesar de ser um *software* facultativo, uma vez que não está diretamente ligado ao processo de tradução, é uma ferramenta muito útil e importante para o desenvolvimento do profissional. Com esta ferramenta, é possível compreender o tipo e a quantidade de mudanças feitas na fase de tradução/revisão do texto (SDL AppStore, 2021). O *feedback* é uma mais-valia para o desenvolvimento do tradutor, que também é encorajado pela empresa. O que esta ferramenta faz é comparar o texto traduzido pelo primeiro tradutor e a revisão do mesmo, apresentando então as mudanças feitas no processo da revisão para melhorar o conteúdo.

Através desta ferramenta, o tradutor consegue perceber que tipo de mudanças foram realizadas e como poderá melhorar a sua próxima tradução. No documento são apresentados apenas os segmentos que sofreram alterações, assim como as colunas com o texto de partida, o texto traduzido, a tradução final e as alterações feitas. As palavras rasuradas a vermelho são os elementos que foram eliminados da tradução inicial. As palavras sublinhadas a amarelo representam os elementos acrescentados à tradução. Este relatório foi obtido durante o estágio curricular (ver Figura 4).

Source	Status	Match	Target (Original)	Target (Updated)	Target (Comparison)
Safe Mode on and	Rejected Translation Approved		Modo de segurança nã : a	Modo de segurança i e	Modo de segurança na a
Safe Mode allows you to start your basic functions active.	Rejected Translation Approved	92%	O Modo de segurança permite iniciar tua mais básicas ativas.	O Modo de segurança permite iniciar a mais básicas ativas.	O Modo de segurança permite iniciar tuas mais básicas ativas.
Be sure to back up data to a USB drive or Cloud Storage before selecting these options.	Rejected Translation Approved		Cria cópias de segurança dos dados numa unidade USB ou no armazenamento em nuvem antes de selecionar estas opções.	Cria cópias de segurança dos dados numa unidade USB ou no armazenamento em nuvem antes de selecionares estas opções.	Cria cópias de segurança dos dados numa unidade USB ou no armazenamento em nuvem antes de selecionar estas opções.
How do I start my in Safe Mode?	Rejected Translation Approved	84%	Como posso iniciar em Modo de segurança?	Como posso iniciar a minha em Modo de segurança?	Como posso iniciar a minha em Modo de segurança?
Release it after you hear the second beep — one beep will sound when you first press, and another seven seconds later.	Rejected Translation Approved	98%	Solta-o depois de ouvires o segundo sinal sonoro; irás ouvir um sinal sonoro quando primes o botão pela primeira vez e outro sete segundos depois.	Solta-o depois de ouvires o segundo sinal sonoro; irás ouvir um sinal sonoro quando premires o botão pela primeira vez e outro sete segundos depois.	Solta-o depois de ouvires o segundo sinal sonoro; irás ouvir um sinal sonoro quando primes o botão pela primeira vez e outro sete segundos depois.
Update System Software Allows you to update the system software via direct download or USB storage device.	Rejected Translation Approved		Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta ou dispositivo de armazenamento USB.	Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta ou dispositivo de armazenamento USB.	Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta ou dispositivo de armazenamento USB.
Update System Software Allows you to update the system software via direct download, USB storage device or disc.	Rejected Translation Approved	90%	Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta, dispositivo de armazenamento USB ou disco.	Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta, dispositivo de armazenamento USB ou disco.	Atualizar o software do sistema; Permite-te atualizar o software do sistema através de transferência direta, dispositivo de armazenamento USB ou disco.

Figura 4 - Exemplo de um relatório de comparação do Xliff Compare

2. A pós-edição

O fenómeno da globalização veio permitir às empresas disponibilizar os seus produtos em várias línguas para captar a atenção de mais clientes a nível mundial. A necessidade de lançar os produtos em simultâneo fez disparar a procura pela tradução profissional, exigindo às empresas de tradução e aos tradutores um aumento da produtividade, sem comprometer a qualidade⁷. Com a introdução da tradução automática no mercado da tradução, e juntamente com as ferramentas CAT, tanto os clientes como as empresas começaram a requisitar mais trabalho e a exigir o mesmo nível de produtividade.

Atualmente, é impossível falar de pós-edição sem mencionar o uso dos sistemas de tradução automática em conjunto com as memórias de tradução, principalmente em empresas como a RWS, que recorrem a sistemas de tradução automática para aumentar a produtividade e dependem das memórias de tradução para manter a consistência. O’Hagan (2013, p.505) denomina este tipo de tarefa “traduções de memória assistidas por tradução automática”. A pós-edição desta tradução semiautomática tornou-se uma prática bastante comum no mundo da tradução e, embora ambas as tecnologias comessem a ser utilizadas separadamente nos anos 80 e 90, a sua relação continuou a desenvolver-se, modificando o papel do tradutor ao

⁷ É importante referir que a definição de qualidade da tradução automática é subjetiva, pois está sempre dependente dos critérios impostos na análise do texto/tradução. No presente relatório “qualidade” refere-se aos textos produzidos que seguem os parâmetros linguísticos considerados “fluentes” de acordo com os critérios sugeridos não só pelo Cliente A, como também pela métrica SAE J2450 utilizada pela RWS.

incorporar sistemas de tradução automática nas ferramentas CAT (Aziz, Sousa, & Specia, 2012, p.3982; O'Hagan, 2013, p.505).

2.1. Conceito

A tarefa de pós-edição é uma área recente no mundo da tradução. Este processo tem um objetivo completamente diferente da tradução clássica. Na sua essência, a pós-edição é uma forma de revisão, onde o tradutor corrige texto pré-traduzido em vez de fazer uma tradução de raiz (Allen, 2003, p.297; Beßler, 2021).

A RWS considera que existem dois tipos de pós-edição: a pós-edição simples ou compreensível e a pós-edição completa ou publicável. O nível publicável é a qualidade padrão exigida pela maioria dos clientes e foi o único tipo de pós-edição utilizado durante o estágio. É usualmente requisitado para textos de grande visibilidade onde é esperado um texto com qualidade elevada. Este nível de pós-edição pode ser comparado com o nível de qualidade de uma tradução. Por outro lado, o nível compreensível tem apenas como objetivo transmitir o mesmo significado que o texto de partida. Contrariamente à pós-edição publicável, o nível de qualidade exigido não é elevado, uma vez que é utilizado principalmente em textos internos, de pouca visibilidade.

2.2. A tradução automática

A tradução automática é uma tecnologia de tradução que tem vindo a ser desenvolvida nos últimos 20 anos e é atualmente uma ferramenta bastante comum na área de tradução, seja profissionalmente, com o uso de *software* como o *SDL Machine Translation* ou para utilização diária, como o *Google Translate* e o *DeepL*⁸.

Contrariamente a outras ferramentas utilizadas para auxiliar o processo de tradução, a tradução automática é uma ferramenta capaz de traduzir frases e textos sem intervenção humana (Zhou & Ma, 2021, p.213). O sistema sugere uma tradução “bruta” do texto de partida, que acaba por ser um ponto de partida para a intervenção do tradutor/pós-editor (Koglin & Cunha, 2019, p.40), evitando fazer traduções de raiz. Profissionalmente, recorre-se a esta ferramenta para aumentar o nível de produtividade, principalmente em contexto de traduções técnicas, e

⁸ Em 2017, mais de 500 milhões de utilizadores utilizou o *DeepL Translator* (DeepL, 2020) e em 2018, a ferramenta *Google Translate* contava com cerca de 200 milhões de utilizadores por dia (Prates, Avelar, & Lamb, 2019, p.6365).

reduzir os custos da tarefa, uma vez que o mesmo trabalho é realizado em menos tempo e de forma mais rápida (Plitt e Masselot, 2010 *apud* Gaspari, *et al.* 2014, p.60).

Tal como Heyn (1996, p.117) afirma, a tradução automática é descrita como um sistema de sugestões que pode ser ou não utilizado pelo tradutor, consoante as condições de tradução para o projeto em específico. A tradução automática não é um sistema essencial de utilização obrigatória, mas sim uma tecnologia opcional que assiste e ajuda na tarefa de tradução.

Vários estudos revelam que o uso de tradução automática em segmentos mais pequenos parece produzir uma tradução mais eficiente e mais controlada (Kartsaklis & Kanavos, 2010, p.17), bem como ajudar, de certa forma, na produtividade (O'Brien & Moorkens, 2014, p.3; Gaspari, *et al.* 2014, p.61). Zaretskaya (2015, p.1) afirma que o *software* é barato e fácil de utilizar, resultando em traduções de rascunho mais rápidas que apenas precisam de alguma edição para chegar à qualidade pretendida. No entanto, pode, por vezes, apresentar alguns problemas, pelo que são precisos profissionais para editar e aperfeiçoar o resultado gerado, de forma a atingir uma qualidade compreensível ou excelente (CERTT, 2012, p.4).

2.2.1. Contexto da tradução automática

A tradução automática começou a ser estudada nos anos 50, juntamente com outras disciplinas de computação. Entretanto, as investigações foram interrompidas, uma vez que as abordagens estudadas não eram capazes de desempenhar o papel pretendido (Freitas & Liu, 2017, p.1). Nos anos 80, o interesse pela tradução automática ressurgiu. Desde então, esta ferramenta tem sido incorporada nos trabalhos de localização como um sistema auxiliar da tarefa de tradução e é cada vez mais comum as grandes empresas utilizarem este *software* para traduzir os projetos mesmo antes de os enviarem para os tradutores/pós-editores (Arenas, 2008, p.11).

O conceito de tradução automática surgiu há algumas dezenas de anos, mas a Segunda Guerra Mundial foi um ponto crucial para a criação de um sistema de tradução automática. Através da construção dos primeiros computadores, bem como dos conhecimentos de Alan Turing e da sua equipa, as bases para a tradução automática foram fundamentadas (Stein, 2013, pp.5-6). Esta tecnologia foi inicialmente criada para automatizar completamente o processo de tradução, considerado repetitivo, intuitivo e mecânico, e era capaz de produzir textos legíveis e lógicos. Contudo, os sistemas de tradução automática continuaram a progredir. Em:

- **1949**, Warren Weaver publicou uma das primeiras teorias sobre a tradução automática no seu memorando sobre tradução.

- **1951**, Yehosha Bar-Hillel tornou-se no primeiro investigador da área com a sua pesquisa realizada no Instituto de Tecnologia de Massachusetts (MIT).
- **1954**, o primeiro sistema de tradução automática foi apresentado ao público.
- **1996**, a tradução automática começou a ficar disponível *online*, com o SYSTRAN, um programa gratuito destinado a traduzir textos simples.
- **2007**, surgiu o motor de tradução automática estatística MOSES.
- **2008/2009**, foi lançado um serviço de tradução de texto para dispositivos móveis no Japão e um telemóvel capaz de traduzir diálogo entre as línguas inglesa, japonesa e chinesa (Okpor, 2014, p.159).
- **2013**, a empresa Google liderou investigações sobre a tradução automática neuronal (NMT) e lançou o *Google Translate*.

A indústria da tradução continuou a ver grandes progressos nesta área com o aumento da qualidade das traduções automáticas, resultando no cenário perfeito capaz de equilibrar o custo das traduções com níveis de produtividade e qualidade ideais.

2.2.2. Tipos de tradução automática

Desde o final dos anos 90 que a área de tradução automática tem desenvolvido vários tipos de sistemas de tradução automática (O'Hagan, 2013, p.510). De sistemas que tentaram imitar a forma como o ser humano processa a linguagem através de regras gramaticais e sistemas baseados em dados, a sistemas que utilizam probabilidades estatísticas ou utilizam exemplos previamente traduzidos como modelos (CERTT, 2012, p.4), todos apresentaram soluções e fórmulas matemáticas diferentes para poderem traduzir os textos na língua de partida de forma automática e autónoma. O capítulo seguinte apresenta uma lista dos vários sistemas de tradução automática utilizados antes do surgimento da tradução automática NMT.

2.2.2.1. Os vários tipos de tradução automática

1. **EBMT**, ou tradução automática com base em exemplos, procura exemplos semelhantes para traduzir o texto de partida (Tripathi & Sarkhel, 2010, p.391). Este sistema processa o corpus dos exemplos antes de iniciar a tradução e utiliza esses mesmos exemplos para traduzir tipos de frases semelhantes da língua de partida para a língua de chegada (Forcada, 2010, p.220; Okpor, 2014, p.163).
2. **RBMT**, ou tradução automática baseada em regras, é um sistema de tradução automática que se baseia em regras gramaticais e palavras comuns que constam no dicionário. Este sistema é capaz de lidar com vários fenómenos linguísticos e é

extensível e sustentável. O seu principal objetivo é converter as estruturas da língua de partida para a estrutura da língua de chegada (Tripathi & Sarkhel, 2010, p.389).

3. **SMT**, ou tradução automática estatística, é um sistema que recorre ao corpus alinhado para treinar o sistema e utiliza probabilidades estatísticas, ao contrário da tradução RBMT (O'Hagan, 2013, p.511). São aplicados métodos estatísticos através de corpora bilingue para traduzir o conteúdo (Tripathi & Sarkhel, 2010, p.390). Até ao surgimento da tradução automática neuronal, a RWS recorria a sistemas de tradução automática estatística para traduzir os seus projetos.

2.2.2.2. *O avanço da tradução automática neuronal*

A tradução automática neuronal, ou NMT, é uma nova abordagem à tradução automática que simula o funcionamento do cérebro e aprende a traduzir através de uma grande rede neuronal. De acordo com Koglin e Cunha (2019, p.38), o sistema NMT surgiu como uma revolução ao paradigma criado pelas conquistas dos sistemas SMT. O sistema aprende através da observação das correlações entre o texto de partida e o texto de chegada e modifica-se de forma a aumentar a probabilidade de produzir traduções mais corretas. Este sistema utiliza uma abordagem de codificação e descodificação, convertendo as palavras em sequências de representação contextual e produzindo uma sequência de palavras (Koehn, 2017 *apud* Koglin & Cunha, 2019, p.38). Consegue realizar tal tarefa porque capta não só informação sintática, mas também o sentido e a semântica das frases, e analisa as frases completas, em vez de palavras ou expressões. Por conseguinte, pode determinar o conteúdo e o significado do texto de partida. Isto permite ao sistema produzir traduções com mais qualidade, bem como resultados mais fluentes em relação aos sistemas anteriormente utilizados por profissionais. Contrariamente ao sistema de tradução automática estatística, quando o sistema NMT tem acesso a uma grande base de dados de vocabulário e é treinado numa determinada língua, é capaz de apresentar melhores resultados (Moorkens & Way, 2019, p.2; Liu, *et al.* 2020, p.1).

Esta tecnologia tem vindo a mostrar cada vez mais resultados satisfatórios em termos de desempenho em comparação com os sistemas anteriores (RWS Trados, s.d.). Como resultado, é atualmente o tipo de tradução automática preferido na tradução de projetos. Alguns dos sistemas populares que contam com tradução automática neuronal incluem o *Google Translate*, o *DeepL Translate* e o *SDL Machine Translation*.

2.2.3. O panorama da tradução automática atual

Atualmente, em contexto profissional, a tradução automática é utilizada em conjunto com sistemas de memória de tradução. Com a tradução automática neuronal, a integração dos *fuzzy matches* tornou-se menos óbvia, uma vez que a ferramenta de tradução automática não precisa de criar uma base de dados de sequências alinhadas ou não utiliza regras para descodificar o texto (Xu, Crego, & Senellart, 2020, p.1580). O trabalho do tradutor/pós-editor passa essencialmente por corrigir os erros gramaticais existentes, tornar o estilo do texto apropriado e, no geral, tornar o texto o mais compreensível e natural possível (Loureiro, 2019, p.21).

Embora a tradução automática ainda não apresente resultados perfeitos, estes são muitas vezes considerados de boa qualidade e capazes de transmitir a mensagem geral do texto original (Loureiro, 2019, p. 19). Contudo, por mais avançada que seja, ainda é necessária intervenção humana para produzir textos de qualidade desejada (Han, 2020, p.3), uma vez que a tradução automática é um sistema desenhado para servir como uma extensão das capacidades do utilizador e não como uma substituição de profissionais capaz de traduzir textos autonomamente (Lipmann, 1971 *apud* Hutchins, 1998, p.293).

2.2.4. Obstáculos à tradução automática

A tradução automática neuronal, apesar de ser uma tecnologia bastante avançada na área da tradução automática, ainda apresenta alguns problemas de desempenho. Um deles está relacionado com certos pares de línguas menos utilizados. Se não são utilizados de forma contínua, o sistema não é treinado o suficiente para produzir traduções mais fluentes (Kartsaklis & Kanavos, 2010, p.18). Independentemente de o sistema ter capacidade para produzir traduções com a qualidade desejada, os tradutores continuam a depender das bases terminológicas e memórias de tradução disponíveis para poderem cumprir todos os requisitos impostos pelo cliente, manter a consistência com as traduções já feitas e, em muitos casos, manter uma linguagem natural (Kartsaklis & Kanavos, 2010, p.19).

Alguns tradutores afirmam ainda que a tradução automática pode abrandar o seu ritmo de trabalho (O'Brien & Moorkens, 2014, p.3). Segundo Jia, *et al.* (2019, p.79), observou-se que a pós-edição com NMT gerou traduções significativamente mais rápidas em relação às traduções feitas de raiz. Todavia, esta diferença só se observou de forma mais significativa na tradução/pós-edição de textos técnicos. É comum os sistemas de tradução automática sugerirem traduções de qualidade muito baixa. Dependendo do tipo de qualidade, poderá requerer mais esforço por parte do tradutor (Aziz, *et al.*, 2012, p.3986), fazendo com que seja necessário descartar as sugestões de tradução. O estilo da tradução também acarreta um papel muito

importante nesta tecnologia de tradução, o que significa que a tradução automática não é adequada para todo o tipo de texto. Liu, *et al.* (2020, p.8) acrescenta ainda que, quando apresentada a opção de escolher entre duas palavras com significados semelhantes, o sistema de tradução automática neuronal sugere aquela que é utilizada com mais frequência, ou seja, é possível evidenciar uma correlação entre a inserção das palavras e a frequência das palavras. Logo, pode-se afirmar que o *software* é ideal para traduções de textos repetitivos com vocabulário bastante recorrente. Contudo, o mesmo não acontece no que toca às traduções mais “livres”⁹, como por exemplo, a tradução literária, que inclui expressões e vocabulário com mais de uma interpretação e menos repetitivos.

Outro obstáculo que impede o sistema de tradução automática de atingir um bom resultado é a sua capacidade para lidar com o discurso e as características textuais mais complexos¹⁰ e dependentes do contexto. Também a capacidade de encontrar soluções de tradução adequadas está dependente do tipo de texto e da forma como o sistema é construído. Como resultado, é possível averiguar quais os segmentos que irão apresentar problemas de tradução e que precisarão de intervenção humana através da pós-edição e quais irão sugerir uma tradução considerada correta e adequada (Koglin & Cunha, 2019, p.38).

Ainda assim, e embora a utilização da tradução automática continue a ser um tema bastante debatido na comunidade de tradução, não podemos negar o avanço rápido desta tecnologia (O’Brien & Moorkens, 2014, p.1). Mais à frente, iremos explorar um pouco mais estes erros e outros erros comuns identificados em exemplos de projetos realizados durante o estágio que afetam a qualidade geral da tradução.

2.3. As memórias de tradução

As memórias de tradução foram criadas com o objetivo de armazenar o conteúdo digital em bancos de memória para poder reciclar os segmentos traduzidos anteriormente e permitir ao tradutor corresponder as traduções antigas com o novo texto (Reinke, 2018, p.55; Xu, *et al.*, 2020, p.1581).

Esta ferramenta é um dos componentes mais desenvolvidos e populares das ferramentas de apoio ao processo de tradução, juntamente com os sistemas de tradução automática (Vela, *et*

⁹A linguagem livre remete para o tipo de escrita mais descontraído e informal, semelhante ao discurso diário. Normalmente, inclui mais idiomas, e expressões de duplo significado e menos diretas.

¹⁰ São considerados segmentos complexos aqueles que apresentam frases de tamanho maior e que contém mais informação.

al., 2019, p.8; Zaretskaya, *et al.*, 2015, p.85). Tornou-se na ferramenta CAT mais utilizada em contexto profissional na área de tradução, especialmente em textos especializados e na localização de informação digital (Reinke, 2018, p.55; Schjoldager & Christensen, 2017, p.89). Alguns autores sugerem que esta ferramenta é comum porque garante diversas vantagens aos tradutores, desde o aumento da produtividade e qualidade à redução dos custos (Reinke, 2018, p.56; Schjoldager & Christensen, 2017, p.1), bem como ajuda a poupar tempo e esforço no processo de tradução de documentos mais repetitivos (CERTT, 2012, p. 9). Igualmente, a RWS afirma que a utilização destes sistemas aumenta drasticamente o nível de qualidade, produtividade e eficiência da tarefa de tradução (RWS, s.d.), bem como ajuda a manter a consistência não só do vocabulário do cliente, mas também do tipo de linguagem requerida pelo cliente e pela empresa. No entanto, Flanagan, *et al.* (2010, p.190) e Reinke (2018, p.56) demonstram que os tradutores recorrem às memórias de tradução porque estas são úteis na verificação de terminologia e contexto, bem como, permitem manter uma maior consistência de vocabulário e expressões.

As memórias de tradução assistidas por tradução automática vieram mudar o processo de tradução para sempre, principalmente porque o tradutor não precisa de traduzir os projetos de raiz (Bundgaard & Christensen, 2019, p.15). As memórias de tradução também têm um grande impacto nos sistemas de tradução automática. Segundo Kartsaklis e Kanavos (2010, p.18), quanto mais dados as memórias de tradução contiverem, mais eficiente será o processo de tradução automática, melhorando a escolha de vocabulário e a qualidade geral do sistema no projeto. Por exemplo, será muito mais produtivo traduzir um texto técnico e repetitivo, cuja linguagem é considerada menos complexa sintaticamente (Jia, *et al.*, 2019, p.68) e cujo sistema aprende com as correções, do que com um mais expressivo.

Além disso, a divisão dos segmentos presente em alguns sistemas de tradução atuais dá-se devido às memórias de tradução, pois permite facilitar o uso de memórias de tradução. Normalmente, cada segmento corresponde a uma frase ou unidades de texto semelhantes a uma frase. Através do alinhamento¹¹, estes segmentos são então associados aos segmentos correspondentes no texto de chegada. Em seguida, a memória de tradução compara cada segmento e o conteúdo guardado na sua base de dados. Se encontrar segmentos semelhantes de traduções antigas, esse mesmo segmento é recuperado da base de dados e sugerido ao tradutor

¹¹ A ferramenta de alinhamento compara os textos de partida e chegada e determina a correspondência do segmento original com a sua respetiva tradução (Trados, s.d.).

(Bowker, 2005, p.14). No caso das tarefas realizadas durante o estágio, essas sugestões eram apresentadas em forma de *fuzzy matches* ou *fragment matches*.

Apesar das vantagens que os sistemas de memória de tradução possam oferecer, é necessário ter em conta o lado menos produtivo desta ferramenta, principalmente se as memórias de tradução estiverem poluídas. Quando bem gerida, esta ferramenta é essencial para manter a consistência da linguagem e permitir ao tradutor ser mais produtivo. Se tal não acontecer e existirem várias traduções para um só segmento, isto pode causar problemas de inconsistência e diminuir a qualidade do texto a ser traduzido (Bowker, 2005, p.15; LeBlanc, 2013, p.8). Uma vez que o processo de utilização de uma memória de tradução é interativo, cada nova tradução aumenta o tamanho da respetiva memória de tradução, e se esta for gerida e atualizada de forma apropriada, irá tornar-se cada vez mais útil para futuras traduções (LeBlanc, 2013, p.8; Vela, *et al.*, 2019, p.8).

2.3.1. A utilização de memórias de tradução

Um sistema de memória de tradução inclui uma variedade de funcionalidades e ferramentas que ajudam o tradutor no processo de tradução, desde gestão de terminologia, a sistemas de reconhecimento automático de termos e ferramentas de alinhamento (Reinke, 2018, pp.58-59).

De entre as funcionalidades que o sistema de memória de tradução oferece, as mais utilizadas para realizar a tarefa de pós-edição são: a pesquisa de *concordance*, *fuzzy matching* e *fragment matching*. Devido a questões de relevância, abaixo será apenas abordada a funcionalidade mais utilizada e importante para o trabalho realizado durante o período de estágio: a *pesquisa de concordance*.

2.3.1.1. Pesquisa de concordance

Segundo O'Hagan (2013, p.504), a maioria das ferramentas de memória de tradução comercial contém a pesquisa de *concordance*. Esta funcionalidade permite aos tradutores pesquisar e utilizar fragmentos de texto previamente traduzidos e armazenados nas memórias de tradução (Bundgaard & Christensen, 2019, p.17). Dado que os meses de estágio serviram como um período de adaptação aos clientes da empresa, a utilização da pesquisa de *concordance* foi crucial na tradução dos projetos realizados durante o estágio, devido à necessidade de manter a consistência com a linguagem da empresa e de certos clientes, principalmente nos textos mais repetitivos.

Esta funcionalidade foi utilizada especialmente para verificar termos, procurar fragmentos dos segmentos ou expressões, assim como termos ou frases semelhantes ao texto de partida que o sistema não conseguiu sugerir devido à segmentação ou à diferença de escrita. A imagem abaixo retrata a pesquisa de *concordance* no *SDL Trados Studio*¹².

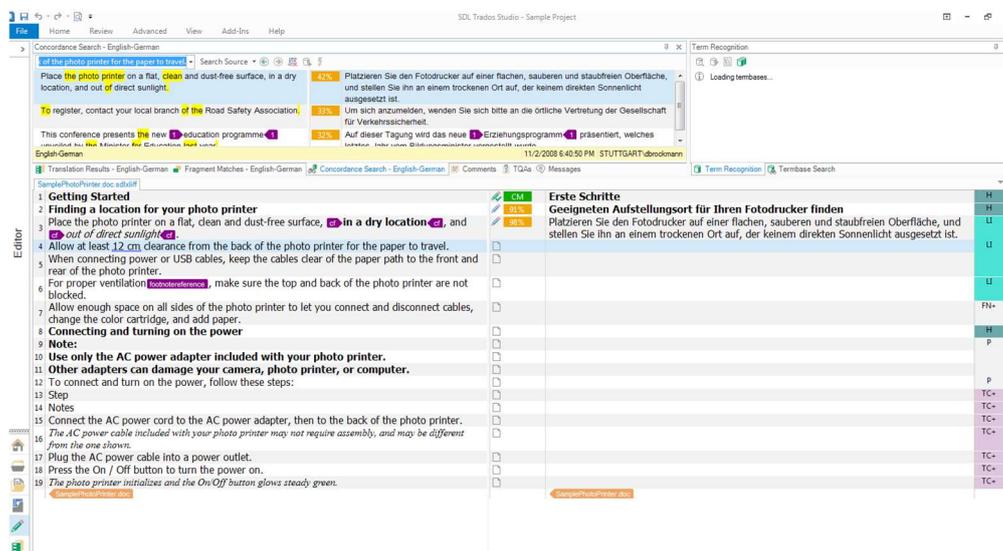


Figura 5 - Pesquisa de concordance no *SDL Trados Studio* 2017

A *concordance* pode ser utilizada através da pesquisa de uma palavra ou frase. Quanto maior for a frase, menos específica será a pesquisa e mais abrangentes serão os resultados (ver Figura 6). No entanto, quanto menor for a frase, ou se for apenas pesquisada uma palavra, a probabilidade de obter resultados semelhantes à pesquisa é maior (ver Figura 7).

¹² A imagem apresentada é uma imagem do *SDL Trados Studio* 2017. Contudo foi utilizada porque não existem diferenças significativas entre as versões de 2017 e 2019.

<https://community.sdl.com/product-groups/translationproductivity/f/studio/9451/sdl-trados-studio-2017-automated-concordance-search-if-no-tm-results-not-working>

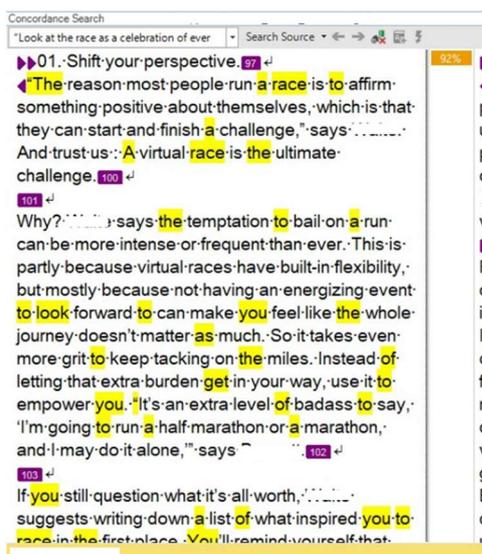


Figura 6 - Pesquisa concordance pouco específica

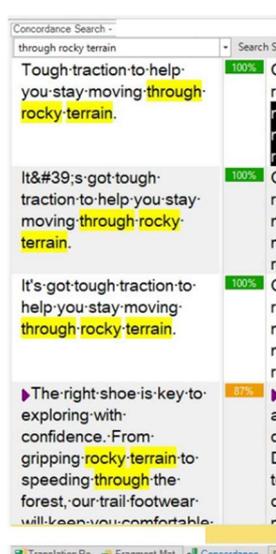


Figura 7 - Pesquisa concordance mais específica

No *SDL Trados Studio*, é possível usar um atalho: ao selecionar a palavra ou frase que o tradutor pretende pesquisar, basta carregar na tecla F3. Também é possível pesquisar na própria janela. A partir daí, o sistema realiza uma pesquisa na sua base de dados e dispõe dos resultados, apresentando os segmentos da língua de partida e a respetiva tradução na língua de chegada. Caso a sugestão seja útil para o tradutor, basta copiar e colar a parte do texto que pretende utilizar (Bundgaard & Christensen, 2019, p.17).

3. Pós-edição na RWS

Enquanto empresa de tradução, a RWS recorre frequentemente ao uso de sistemas de tradução automática e memórias de tradução para traduzir os vários projetos de forma eficiente, consistente e rápida. A pós-edição é um processo complexo que requer bastantes orientações e estratégias para produzir um texto com a qualidade desejada. No entanto, as orientações poderão sempre variar de cliente para cliente e exigir diferentes tipos de esforço e trabalho por parte do tradutor. Dado que este relatório inclui um estudo de caso do Cliente A, seria imprescindível mencionar as várias orientações do cliente e da RWS necessárias para produzir textos com a qualidade desejada segundo a métrica de qualidade SAE J2450.

3.1. Perfil do Cliente A

O cliente e os projetos mencionados estão relacionados com a indústria do vestuário e produtos desportivos. Uma das principais razões por escolher realizar um estudo de caso sobre este cliente em particular é por ter sido a conta mais trabalhada durante os meses de estágio (ver Gráfico 2). Além disso, o sistema de tradução automática, bem como memórias de tradução e bases terminológicas foram utilizados com frequência na realização dos projetos.

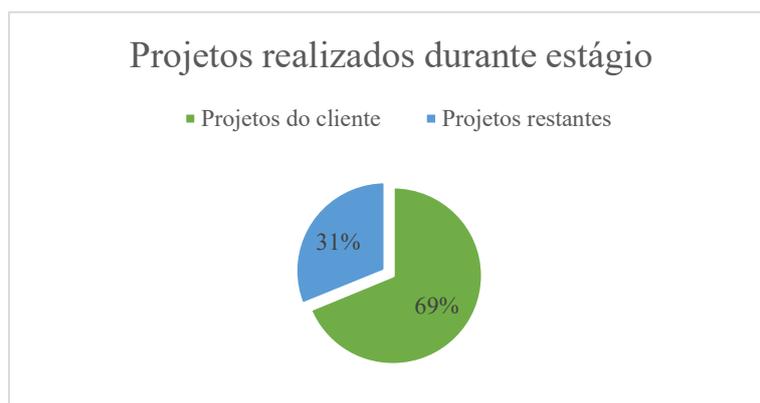


Gráfico 2 - Comparação dos projetos realizados

O cliente forneceu orientações específicas (que serão apresentadas no capítulo 3.1.2) e exigia um nível de qualidade um pouco diferente dos restantes clientes. Por esse motivo, os seus projetos eram normalmente enviados aos membros da equipa que já se encontravam familiarizados com o mesmo. Desta forma, conseguiam ser mais produtivos, uma vez que não precisavam de investir parte do seu tempo a ler todas as orientações disponibilizadas. Assim, por mostrar aptidão na realização dos projetos, comecei a traduzi-los em maior quantidade, tornando-me num dos membros familiarizados com o cliente.

O principal objetivo do cliente era criar projetos cujo texto de chegada fosse amigável, apelativo, o mais inclusivo possível e que tivesse uma linguagem natural. De forma a simplificar os vários tipos de projetos do cliente, estes serão divididos em duas categorias: os projetos que continham linguagem informal e os projetos de linguagem formal (Diagrama 3).



Diagrama 3 - Tipos de projetos do cliente

A maioria dos projetos realizados eram de linguagem informal (ver Gráfico 3) e exigiram uma combinação de linguagem apelativa e linguagem específica da área de vestuário desportivo. Dentro da linguagem informal, os principais textos trabalhados abrangiam (i) projetos de descrição de produtos, onde o tradutor precisava de traduzir de forma objetiva, mas apelativa e utilizar terminologia específica da área de vestuário e calçado; (ii) projetos que utilizavam uma linguagem mais livre, onde se incluíam textos de artigos e entrevistas e; (iii) uma combinação dos dois primeiros tipos de textos, onde era utilizada uma linguagem mais informal, mas ao mesmo tempo mais apelativa e descritiva. Dentro do espectro da linguagem mais formal, foram realizados projetos de natureza descritiva, expositiva ou instrucional, abrangendo principalmente formações para os funcionários, termos e condições e artigos de assistência.

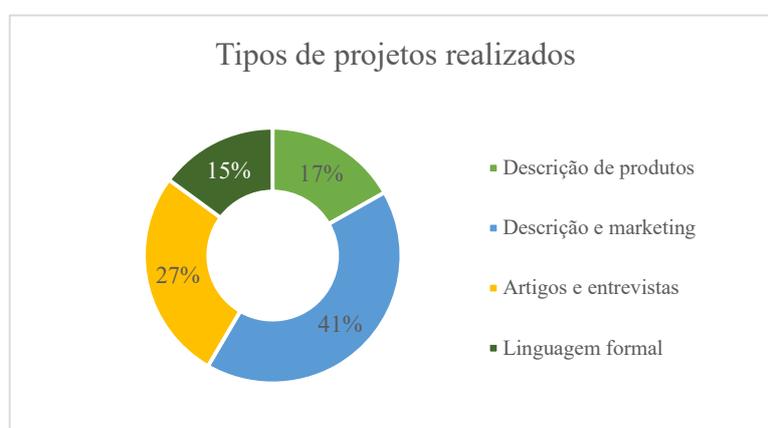


Gráfico 3 - Comparação de tipos de textos traduzidos/editados

3.1.1. O cliente e a pós-edição

Todos os projetos¹³ realizados relativos ao Cliente A foram pós-edições e foram acompanhados de tradução automática neuronal, fornecida pelo serviço *SDL Machine Translation Cloud*, memórias de tradução e bases terminológicas. As memórias de tradução

¹³ No caso da legendagem, esta era sempre realizada como tradução, mas com o auxílio das memórias de tradução e suas funcionalidades.

foram uma funcionalidade crucial para o cliente, uma vez que este enviava textos repetitivos que precisavam de manter a consistência. No total, foram editadas cerca de 101 922 palavras, tanto palavras novas como *fuzzy matches* (Gráfico 4). O Gráfico 5 representa o número total de palavras editadas durante o estágio. Como é possível ver, o número de *fuzzy matches* foi consideravelmente maior do que o número de palavras novas, comprovando que a conta apresentava vários segmentos semelhantes ou repetidos.

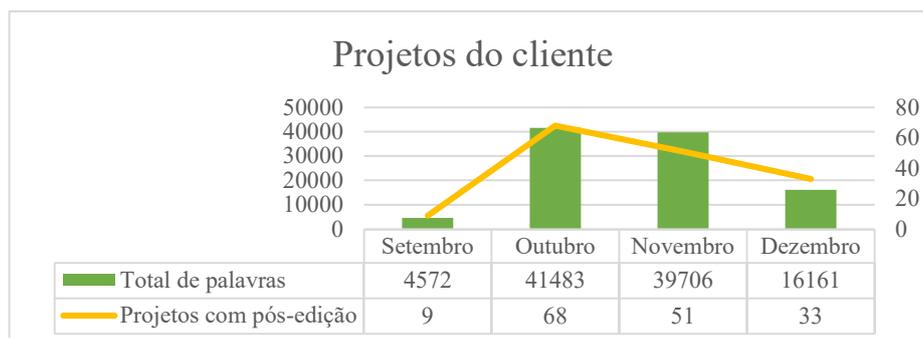


Gráfico 4 - Dados dos projetos realizados durante o estágio

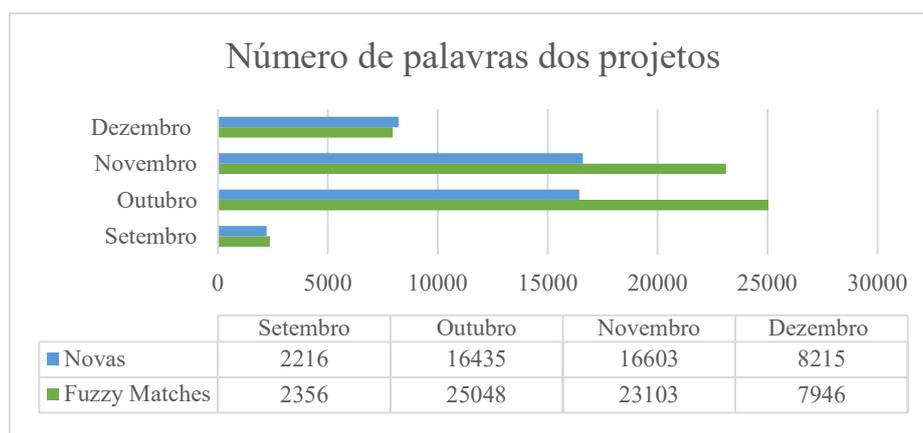


Gráfico 5 - Número de palavras novas e fuzzy matches editados durante o estágio

3.1.2. Orientações do cliente

Este capítulo irá abordar as orientações do Cliente A destinadas a atingir o nível de qualidade desejado, incidindo mais detalhadamente nos erros gramaticais, tempos verbais, números e unidades de medida, terminologia e formas de tratamento (ver Diagrama 4).

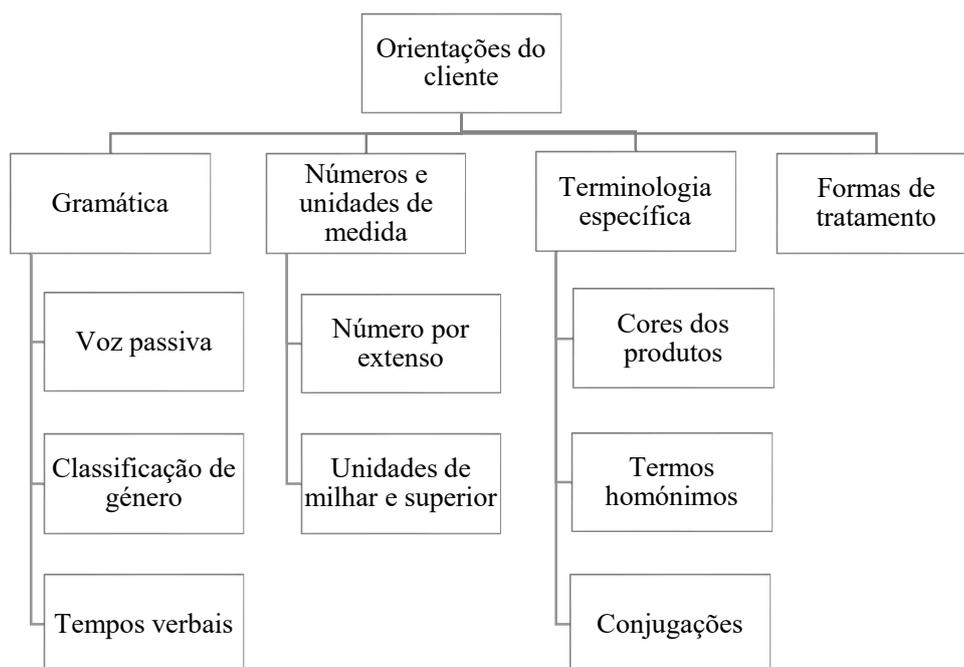


Diagrama 4 - Orientações do cliente

Este destaque está relacionado com os possíveis erros dos sistemas de tradução automática, visto que muitos dos erros identificados durante o estágio não cumpriam algumas das orientações do cliente. Consequentemente, sendo este estudo de caso assente nos erros da tradução automática e o seu efeito na tarefa de pós-edição, foi considerado que estas orientações específicas seriam as mais pertinentes para o estudo em questão.

▪ Gramática

Dentro das orientações relacionadas com a gramática, existem vários aspetos a ter em conta aquando da tradução de cada projeto.

- *Voz ativa e passiva*: o uso da voz passiva é algo pouco comum na língua portuguesa. O cliente sugere tentar limitar o uso da mesma.
- *Géneros (classificação do género de indivíduos)*: o tradutor deve sempre optar por uma tradução neutra e evitar o uso de género. Se for necessário, deve-se rescrever a frase de forma a evitar dirigir o texto apenas a um grupo. No entanto, nos casos em que o tradutor vê que não há qualquer alternativa, é sugerido referir-se ao público através do masculino. Nos casos em que os textos se dirigem a audiências específicas, opta-se pelo género em questão. O objetivo é incluir todas as identidades e comunidades. Eis alguns exemplos de como as frases podem ser escritas de forma inclusiva (Tabela 1).

Texto de partida	Sugestão de tradução automática	Proposta de tradução final
<i>To keep you fresh</i>	Para te manter fresco	Para manter a frescura
<i>Are you ready?</i>	Estás preparado?	Tudo a postos?

Tabela 1 - Sugestões sobre como evitar o uso de género

Este é um assunto complexo, principalmente para o sistema de tradução automática, uma vez que são escolhas muito específicas de cada língua, principalmente entre os pares de línguas inglês, que não precisa de determinantes, e português que tem dois géneros.

- *Tempos verbais*: o gerúndio deve ser evitado, pois esta conjugação não é muito recorrente no português europeu. Por vezes, a tradução automática apresenta frases com esta forma verbal, produzindo um texto mais próximo da estrutura do português do Brasil, cuja utilização é comum. É aconselhado reformular a frase para não utilizar esta estrutura. A tabela seguinte (Tabela 2) apresenta alguns exemplos de como as frases podem ser reescritas para evitar o uso do gerúndio.

Com gerúndio	Sem gerúndio
Aceda ao artigo clicando na ligação.	Clique na ligação abaixo para aceder ao artigo.
Define um novo objetivo acedendo à aplicação.	Define um novo objetivo através da aplicação.

Tabela 2 - Alternativas ao uso do gerúndio

▪ **Números, unidades de medida, datas e moeda**

É comum o sistema de tradução automática traduzir os números exatamente como se apresentam no texto de partida. Assim, deve-se ter em conta as seguintes orientações:

- *Números por extenso*: por regra, os números entre 1-9 são sempre escritos por extenso, exceto quando se trata de datas ou medidas.
- *Unidades de milhar e superior*: para os números com mais de quatro dígitos, basta separar com um espaço, ao contrário da língua inglesa que utiliza uma vírgula (por exemplo, 100 000 vs. 100,000).

▪ **Terminologia específica**

Tal como acontece com os números, o sistema tem tendência para traduzir literalmente o texto de partida. Existem muitos outros temas que não se enquadram num grupo específico,

mas que são igualmente importantes. Desde como escrever as cores dos produtos e termos homónimos até à utilização de conjunções, eis outros temas que o tradutor precisa de ter em atenção:

- *Cores dos produtos*: as cores genéricas como vermelho, branco, verde, devem ser escritas com letra minúscula. Se a cor for acompanhada por um substantivo que não esteja relacionado com nomes de cores, como por exemplo “Plant Green”, só se deve traduzir a cor e o substantivo fica em inglês: “verde Plant”. No entanto, se o substantivo estiver relacionado com o tópico de cores, por exemplo, “Metalic Purple”, traduz-se para a língua de chegada: “roxo metalizado”. Quando a cor faz parte do nome do produto, esta deve ser mantida na língua de partida.
- *Termos homónimos*: palavras com o mesmo significado em inglês, mas que contêm várias traduções na língua de chegada e, conseqüentemente, devem ser traduzidas consoante o contexto.

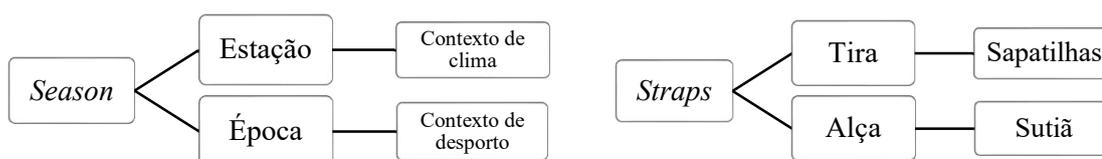


Diagrama 5 e 6 - Exemplos de termos homónimos

- *Conjunções “e”, “ou” e “mas” no início da frase*: o ideal é substituir estas conjunções por “além disso”, “em alternativa” e “no entanto” respetivamente ou, quando possível, não traduzir. É comum a língua inglesa começar as frases com estas conjunções, e conseqüentemente, o sistema de tradução automática segue o texto original. Na tabela seguinte são apresentados alguns exemplos da sugestão de tradução automática e a respetiva proposta de tradução para evitar a utilização destas conjunções no início das frases (Tabela 3).

Texto de partida	Tradução automática	Proposta de tradução final
<i>And you have to run lots of errands during the day.</i>	E você tem que executar lotes dos errands durante o dia	Além disso, precisas de realizar várias tarefas ao longo do dia.
<i>Or you could self-diagnose by looking at your most loved shoes.</i>	Ou podes fazer um autodiagnóstico olhando para as tuas sapatilhas mais adoradas.	Em alternativa, podes fazer um autodiagnóstico ao olhar para as tuas sapatilhas favoritas.
<i>But this month it's all about you.</i>	Mas este mês é tudo sobre ti.	No entanto, este é o teu mês.

Tabela 3 - Exemplos do uso de conjunções no início da frase

▪ Formas de tratamento

Por último, a utilização do tratamento informal e formal está dependente do tipo de texto com que o tradutor está a trabalhar. O tradutor deve ter em consideração que o sistema de tradução automática não consegue adivinhar quando deve utilizar cada forma de tratamento.

Também a utilização dos pronomes pessoais, principalmente das primeira e segunda pessoas deve ser evitada. A tabela seguinte demonstra alguns exemplos do uso de pronomes pessoais e como os evitar (Tabela 4).

Texto de partida	Tradução automática	Proposta de tradução final
<i>I am very excited.</i>	Eu estou muito entusiasmada.	Estou muito entusiasmada
<i>If you're feeling aches or pains, see a physical therapist.</i>	Se você estiver sentindo dores ou dores, veja um fisioterapeuta.	Se sentires dores, visita um fisioterapeuta.

Tabela 4 - Exemplos de como evitar a utilização de pronomes pessoais

3.2. Pós-edição de tradução automática

A utilização de um *software* capaz de produzir traduções automáticas de qualidade aceitável veio exigir um tipo de trabalho diferente. Por essa razão, surgiram práticas de pós-edição capazes de aproveitar as sugestões de tradução automática o máximo possível e traduzir mais projetos em menos tempo, sem comprometer o custo ou a qualidade do produto final.

Segundo Vardaro, *et al* (2019), os tradutores/pós-editores têm três possibilidades no que toca ao processo de pós-edição de tradução automática. Podem (Vardaro, *et al*, 2019, p.2):

- (i) Deixar a sugestão inicial da tradução automática sem alterar nada;
- (ii) Aplicar algumas correções para melhorar a compreensão do texto sugerido ou;
- (iii) Alterar o texto sugerido por completo.

Contudo, as escolhas do tradutor vão ser influenciadas não só pelo tipo de texto e pela qualidade da tradução automática do par de línguas, mas também pelas orientações do cliente e as sugestões da entidade, neste caso da RWS.

3.2.1. Orientações para a pós-edição de tradução segundo a RWS

Dada a frequência com que a pós-edição era praticada para realizar os projetos, a RWS focava-se em estratégias capazes de permitir o aumento da produtividade, a criação de projetos com qualidade elevada e a redução dos custos.

A RWS trabalha maioritariamente com textos técnicos. Isto significa que a qualidade do sistema de tradução automática pode ter tendência para ser mais consistente e apresentar melhores resultados, uma vez que os textos são mais repetitivos e a terminologia está definida. Tendo isto em conta, a empresa sugeria certos procedimentos para realizar a pós-edição de tradução automática de forma mais eficiente, como:

- “Tentar não efetuar alterações desnecessárias”, como por exemplo, substituir palavras pelo seu sinónimo. As alterações preferenciais acabam por invalidar o nível de produtividade pretendido. O ideal é não alterar a frase caso esta apresente vocabulário aceitável para o contexto, salvo por requisito do cliente. A Tabela 5 demonstra um exemplo de uma alteração preferencial segundo o Cliente A.

Texto original	Sugestão de tradução automática	Proposta de pós-edição
<i>One-piece hugs the foot, providing added comfort and lockdown.</i>	A peça única envolve o pé, proporcionando maior conforto e bloqueio .	A peça única envolve o pé, proporcionando maior conforto e firmeza .

Tabela 5 - Exemplo de alterações preferenciais

- “Utilizar a função de copiar e colar do *software*” o máximo possível, principalmente quando se recorre às memórias de tradução.
- “Não apagar ou reescrever palavras”.

Todas estas indicações têm como principal objetivo aproveitar as sugestões o máximo possível e tirar o maior partido das funcionalidades que o *SDL Trados Studio* oferece, para que o tradutor consiga poupar tempo durante a tarefa. No entanto, mesmo que a tradução automática

seja útil e utilizável, caso esta não cumpra os requisitos do cliente e não atinja a qualidade desejada e considerada publicável, necessitará de intervenção por parte do tradutor.

3.3. As métricas de qualidade

As métricas de qualidade permitem aos tradutores avaliar a qualidade de textos de forma consistente e objetiva (SDL, s.d.-a). São maioritariamente usadas para assegurar a qualidade dos projetos e o cumprimento dos padrões definidos, não só pela empresa, mas também pelo cliente. Podem ser utilizadas por tradutores, revisores e até pelos clientes para assegurar que todas as orientações foram cumpridas.

A RWS recorre principalmente ao modelo SAE J2450 para avaliar as traduções realizadas e assegurar a qualidade dos seus projetos. Por isso, a categorização dos erros neste estudo de caso irá seguir a mesma.

3.3.1. O sistema de qualidade da RWS (SAE J2450)

A métrica de qualidade de tradução SAE J2450 foi concebida para definir uma métrica de qualidade padrão para as traduções de indústria automóvel. Pode ser aplicada independentemente do par de línguas utilizado e do método de tradução (SAE International, 2016). No entanto, também pode ser utilizada para avaliar traduções de outras áreas. O sistema foi desenvolvido pela *Society of Automotive Engineers* (SAE) e caracteriza-se por ser um sistema simples, mas focado somente na qualidade linguística das traduções. Foi concebido para ser utilizado manualmente, geralmente em textos de grande dimensão e avalia sete categorias de erros e dois níveis de gravidade (Doherty & Gaspari, 2013, p.15), incluindo os erros de terminologia e significado, omissões, erros estruturais, ortográficos, de pontuação e estilo¹⁴ (Tabela 6).

¹⁴ Consultar anexo A: padrão de avaliação de qualidade J2450

Categorias dos erros
Erro de terminologia
Significado errado
Omissão
Erro estrutural
Erro ortográfico
Pontuação
Estilo
Erros diversos

Tabela 6 - Categorias de erros de acordo com SAE J2450

Sendo um sistema destinado a avaliar a qualidade de textos especializados, onde o estilo, o tom e o registo da língua de partida não são os aspetos mais importantes para o produto final, a métrica evita assim categorizar os “erros” estilísticos (SAE International, 2016), excluindo categorias de análise do estilo e do tom (thebigword, s.d.). Durante a análise dos exemplos, observou-se que, embora as sete categorias da métrica SAE J2450 tenham sido suficientes para identificar a maioria dos erros, concluiu-se que a categoria “erros diversos” era demasiado vasta e certos erros precisaram de ser definidos de forma mais precisa. Por esse motivo, para o presente estudo de caso, foram utilizadas algumas categorias da métrica de qualidade MQM¹⁵, uma vez que oferece uma categorização mais personalizável dos erros e pode ser utilizada em simultâneo nas traduções automática e realizada por tradutores (Doherty & Gaspari, 2013). São elas: tradução incorreta, tradução literal e por traduzir¹⁶.

É também importante referir que a análise foi realizada manualmente, podendo resultar em algumas dificuldades na categorização dos mesmos. No entanto, poderá ser vantajoso para categorizar certas questões que ainda constituem um problema para os programas de *software*, como a identificação de falta de contexto ou assumir uma expressão idiomática. Além disso, sendo a métrica criada por humanos para avaliar a qualidade, principalmente de traduções feitas por humanos, é possível que certos erros cometidos pelo *software* não sejam identificáveis consoante as categorias da métrica.

¹⁵ Sistema métrico de qualidade aprovado pela União Europeia.

¹⁶ Por razões de espaço, serão apenas abordadas de forma geral para poder identificar os erros de forma mais detalhada.

3.3.1.1. *Categorias dos erros*

Antes de qualificar todos os problemas e as categorias em que se enquadram, é importante definir o conceito de erro. Segundo Doherty & Gaspari (2013, p.4), um erro representa algum aspecto da tradução que pode ser corrigido. Quando a palavra não corresponde ao texto de partida ou é considerada incorreta na língua de chegada, esta é então designada como um erro (QTLaunchPad, 2015).

Eis as categorias de erros segundo a métrica SAE J2450, acompanhadas da sua definição retirada do formulário de avaliação da RWS e alguns exemplos retirados de projetos realizados durante o estágio.

a. Erros de terminologia

Para que se possa definir esta categoria, é necessário definir o conceito de “termo”. Segundo o formulário de avaliação da RWS, a palavra “termo” pode ser definida como:

- Uma palavra única, como “automóvel”.
- Uma expressão que corresponde a um significado, como “módulo de controlo de potência”.
- Abreviações, como “Hz”.
- Acrónimos, como “ABS”.
- Números e algarismos, como 3,15 e “três”.
- Nomes próprios, incluindo marcas, marcas registadas e locais.
- Nomes pessoais.
- Palavras funcionais, como as preposições, conjunções, determinantes e artigos.

Desta forma, são considerados erros de terminologia todos os termos que:

- violam o glossário do cliente (exemplo 1).
- estão em conflito com a tradução definida para o termo na língua de chegada na área em questão (exemplo 2).
- não mantêm a consistência com as traduções feitas no mesmo documento ou tipo de documento (exemplo 3) a não ser que o contexto justifique o uso de terminologia diferente.

	Texto de partida	Tradução automática sugerida	Pós-edição proposta
1	<i>The heel cushion allows for breathability.</i>	A almofada do calcanhar proporciona respirabilidade.	O amortecimento no calcanhar garante respirabilidade.
2	<i>This high-rise fit hits at your belly button.</i>	Este corte de cintura subida assenta no botão da barriga .	Este corte de cintura subida assenta ao nível do umbigo .
3	<i>Fresh Fits.</i>	Novos ajustes .	Novos estilos .

Tabela 7 - Exemplos de erros de terminologia

b. Significado errado

Acontece quando um termo ou expressão representa um conceito na língua de chegada diferente do significado transmitido pelo termo do texto de partida (exemplo 1) ou contradiz o significado do texto original.

	Texto de partida	Tradução automática sugerida	Pós-edição proposta
1	<i>What are your kids looking forward to the most?</i>	Quais são os seus filhos que mais procuram?	O que é que as crianças mais desejam?

Tabela 8 - Exemplo de significado errado

c. Omissão

Um erro de omissão pode ocorrer quando um texto ou unidades de texto não apresentam o equivalente do texto de partida no texto de chegada. Isto significa que o conteúdo presente no texto original não é traduzido no texto de chegada. No entanto, é importante referir que esta categoria não implica que a linguagem de chegada deva ser uma correspondência exata da linguagem de partida.

d. Erro estrutural

Os erros estruturais abrangem erros sintáticos, erros de concordância, entre outros. Podem ocorrer quando:

- Um termo do texto de partida surge na posição incorreta do texto de chegada.
- A estrutura frásica utilizada no texto de chegada é incorreta.
- Uma palavra ou termo corretos são traduzidos juntamente com uma forma morfológica incorreta, como por exemplo, número, género ou forma verbal.
- Duas ou mais palavras não cumprem com as regras gramaticais.

e. Erro ortográfico

Os erros ortográficos ocorrem quando o termo na língua de chegada não é escrito corretamente¹⁷ e viola a ortografia definida pelo glossário do cliente ou pelas normas de ortografia da língua de chegada, bem como se for escrito de forma incorreta ou inapropriada. Esta categoria inclui problemas de capitalização e relativos ao novo acordo ortográfico.

Texto de partida	Tradução automática sugerida
<i>So, here we go. Let's see. Oh my god! Wow.</i>	Então, aqui vamos nós. Vamos ver. OH meu deus! Uau.
<i>Soft, Stylish, Exactly what you need</i>	Suave, elegante, exactamente aquilo de que necessita
<i>Experience With Us</i>	Experiência conosco
<i>Be the first to hear about product launches, collaborations and more.</i>	Seja o primeiro a dar conta de lançamentos de produtos, collaboriações e muito mais.

Tabela 9 - Exemplos de erros ortográficos

f. Pontuação

Os erros de pontuação acontecem quando o texto de chegada não cumpre com as regras de pontuação da própria língua. A pontuação é, por regra, diferente entre línguas e é necessário adaptar a posição da mesma para produzir um texto que soe natural ao ler.

g. Estilo

Um problema pode ser categorizado como erro de estilo quando o tom ou o registo não são adaptados ao texto de chegada ou quando as orientações dadas sobre o estilo não são seguidas.

¹⁷ (American Translators Association, 2021)

Os exemplos apresentados na Tabela 10 não respeitam algumas das orientações do cliente mencionadas acima¹⁸.

Texto de partida	Tradução automática sugerida
<i>“For me, these are amazing, because after a long hard session at the track I just slide into them and I’m good to go.”</i>	“Para mim, elas são fantásticas, porque depois de uma longa sessão dura na pista, eu simplesmente faço deslizar para eles e estou bem para ir.”
<i>And if you have to run lots of errands during the day but still want to look cute, these are definitely a hit.</i>	E se você tiver que executar lotes dos errands durante o dia mas ainda quiser olhar bonito, estes são definitivamente um sucesso.

Tabela 10 - Exemplos de erros estilísticos

h. Erros diversos

Esta categoria abrange todo o tipo de erros que o avaliador pode encontrar no texto de chegada e que não encaixam nas categorias anteriormente referidas. Reconhece-se que nem todos os erros podem ser classificados apenas com as sete categorias mencionadas, bem como o facto de o avaliador perceber que há algo de errado na frase, mas não sabe como categorizar o que está incorreto. Os erros desta categoria podem incluir:

- Traduções literais de idiomas;
- Adição de palavras que não têm relevância para o texto;
- Falhas na adaptação de formatos de horas e data ou unidades de medida, etc.

De forma a fazer uma análise mais pormenorizada, os erros de tradução literal, tradução incorreta, escrita confusa e termos por traduzir da métrica MQM serão incluídos nesta última categoria.

- *Tradução literal* – dá-se quando a tradução é feita, tal como o nome indica, literalmente, descartando a expressão ou o contexto em que a palavra está incluída.
- *Tradução incorreta* – dá-se quando o conteúdo de chegada não representa o conteúdo do texto de partida (QT21, pg. 6). Dentro desta categoria inserem-se principalmente erros que, embora a tradução esteja correta de acordo com os dicionários, a mesma não se adequa ao contexto.

¹⁸ Ver capítulo 3.1.2.

- *Por traduzir* – dá-se quando o sistema de tradução automática não traduz o termo ou o conteúdo do texto de partida.

3.3.1.2. O nível de gravidade dos erros

A gravidade dos erros, o segundo passo na classificação dos erros, pode ser classificada entre os níveis crítico, sério, menor ou repetido e cada um terá um peso diferente consoante a categorização do erro e, conseqüentemente, uma pontuação diferente¹⁹. Para o efeito de manter o foco na métrica SAE J2450, serão apenas abordados os níveis sério e menor. A tabela abaixo demonstra o peso de cada categoria consoante a sua gravidade, numa escala de 1 a 5, sendo 5 a pontuação máxima (ver Tabela 11). No momento de avaliação, o projeto só pode somar, no máximo, 13 pontos²⁰ para que o texto seja considerado de qualidade desejável. Embora esta pontuação seja utilizada apenas durante as avaliações, pode servir de base para compreender o peso que cada erro tem quando cometido.

Categoria dos erros	Gravidade	
	Sério	Menor
Erro de terminologia	5	2
Significado errado	5	2
Omissão	4	2
Erro estrutural	4	2
Erro ortográfico	3	1
Pontuação	2	1
Estilo	2	1
Erros diversos	3	1

Tabela 11 - Gravidade dos erros consoante a categoria

De forma a distinguir a gravidade do erro, o avaliador deve ter em conta o seguinte: se o erro acarreta conseqüências para o utilizador da tradução (como um técnico), principalmente

¹⁹ A tabela apresentada contém apenas a pontuação relativa ao nível de qualidade *premium*, a qualidade utilizada nas avaliações dos trabalhos do cliente. A métrica de qualidade J2450 inclui ainda o nível de qualidade de valor e económico. As diferenças entre o nível de qualidade *premium* e por valor diferem apenas nas categorias de erro estrutural e ortográfico. Quanto ao nível de qualidade económica, só os erros sérios é que são pontuados e apenas nas primeiras cinco categorias da tabela.

²⁰ No nível de qualidade *premium*

no contexto de segurança, ou afeta o significado da tradução, este deve ser considerado um erro de gravidade séria. Caso contrário, o avaliador deve classificar o erro como menor.

- *Erros de pouca gravidade*: são os erros que, ainda que detetáveis, não provocam diferenças de significado consideráveis nem confundem ou iludem o leitor.
- *Erros sérios*: este nível de gravidade é selecionado quando o erro em questão modifica claramente o significado do texto de partida e pode induzir o utilizador em erro ou ter um impacto negativo grave.

Os erros mais graves terão sempre um peso maior na pontuação do que os erros de menor gravidade. Sempre que o avaliador tiver dificuldade em escolher o nível de gravidade do erro, deve considerá-lo como um erro sério (SAE International, 2001).

4. Estudo de caso

4.1. Metodologia

O objetivo deste capítulo é analisar a qualidade das sugestões da tradução automática neuronal fornecidas pelo *SDL Machine Cloud* em três tipos de projetos e quantificar os erros identificados.

Durante o estágio foram traduzidos vários tipos de textos e observou-se que as sugestões de tradução automática tinham desempenhos diferentes, dependendo do tipo de texto. Serão apenas analisados exemplos retirados de três tipos de textos: “descrição de produtos”, “descrição de produtos e *marketing*” e “artigos e entrevistas”. Estes textos inserem-se na categoria da linguagem informal²¹.

Iniciei o estudo de caso através da análise de exemplos dos erros identificados em diversos tipos de textos realizados durante o estágio. Todos os documentos utilizados para este estudo foram selecionados de forma a obter um leque variado de erros e problemas que o sistema de tradução automática cometeu, especialmente tendo em conta que a amostra em estudo é pequena. Foram escolhidos sete exemplos: três documentos de descrição de produtos, dois de descrição de produtos e *marketing*, um artigo e uma entrevista. Para facilitar a leitura e a análise do estudo de caso, este capítulo encontra-se dividido em quatro secções: três secções dedicadas a cada tipo de texto, cada um com a respetiva análise, e uma quarta secção que tem como objetivo analisar todos os dados adquiridos na secção anterior e retirar uma conclusão.

²¹ Consultar capítulo 3.1

Após a escolha dos textos, foi realizada uma análise dos erros que foram categorizados de acordo com a métrica SAE J2450 e as orientações específicas do Cliente A. Para além disso, decidi contabilizar quantos segmentos receberam algum tipo de edição, com intenção de comparar o número de erros presentes em cada texto com o número de segmentos inicial. Também será abordada a gravidade dos erros identificados de forma a avaliar o desempenho do sistema de tradução automática de cada tipo de texto, segundo o formulário de avaliação da RWS. Serão adicionados alguns exemplos dos erros identificados e a respetiva proposta de edição.

4.2. Os erros da tradução automática neuronal

4.2.1. Projetos de descrição de produtos

Os projetos de descrição de produtos exigiram a utilização de linguagem específica de vestuário. O seu objetivo era descrever detalhadamente a forma como os produtos eram fabricados, quais os materiais utilizados e o seu propósito. Os documentos tinham tendência a serem repetitivos e, usualmente, continham bastantes segmentos com *fuzzy matches* ou segmentos bloqueados²². Por esse motivo, era recorrente utilizar as funcionalidades de pesquisa de *concordance* e de autopropagação. Isto permitia diminuir o tempo de trabalho investido no projeto e manter a consistência.

Exemplo 1

O primeiro documento analisado contava com 36 ficheiros e cada ficheiro correspondia a um produto diferente. Alguns ficheiros continham conteúdo repetido.

	Nr. segmentos	<i>Context match</i> ²³	100%	<i>Fuzzy alta</i> ²⁴	<i>Fuzzy baixa</i> ²⁵	NMT
Total	751	4	586	47	19	67

Tabela 12 - Dados do projeto

²² Os segmentos bloqueados são normalmente correspondências de 100% ou de contexto que já foram aprovadas pelo cliente. Uma vez que os segmentos são idênticos, não é necessário trabalhar neles e, consequentemente, são bloqueados.

²³ Correspondência de contexto. Acontece quando o segmento, para além de ser idêntico, encontra-se no mesmo contexto que o segmento da memória de tradução.

²⁴ Segundo a RWS, *fuzzy* altas são segmentos que contêm correspondências entre 85% e 99%.

²⁵ Segundo a RWS, *fuzzy* baixas são segmentos que contêm correspondências entre 75% e 84%.

Como a Tabela 12 demonstra, o número de segmentos que continham tradução automática neuronal é muito pequeno comparado com o número de segmentos existente ou o número total de correspondências (*fuzzy matches*, 100% e *context matches*), comprovando a natureza repetitiva do projeto. Dos 67 segmentos que continham tradução automática neuronal, 39 eram segmentos repetidos. Além disso, pelo menos, 57 segmentos²⁶ sofreram algum tipo de alterações, preferenciais ou de correção:

- dois segmentos não sofreram qualquer alteração;
- 16 segmentos sofreram apenas alterações preferenciais;
- 27 segmentos sofreram correções de algum tipo de erro;
- 22 segmentos sofreram tanto alterações preferenciais como correções.

Ao analisar os erros, foi possível reconhecer que o desempenho do sistema de tradução automática era relativamente bom e que, em muitos casos, poderia passar sem qualquer pós-edição. Ainda assim, identificaram-se vários tipos de erros, incluindo 14 erros de tradução literal, 13 erros de terminologia, oito erros de tradução, seis erros estruturais, cinco erros estilísticos, cinco erros de significado, um erro de omissão e um erro de pontuação.

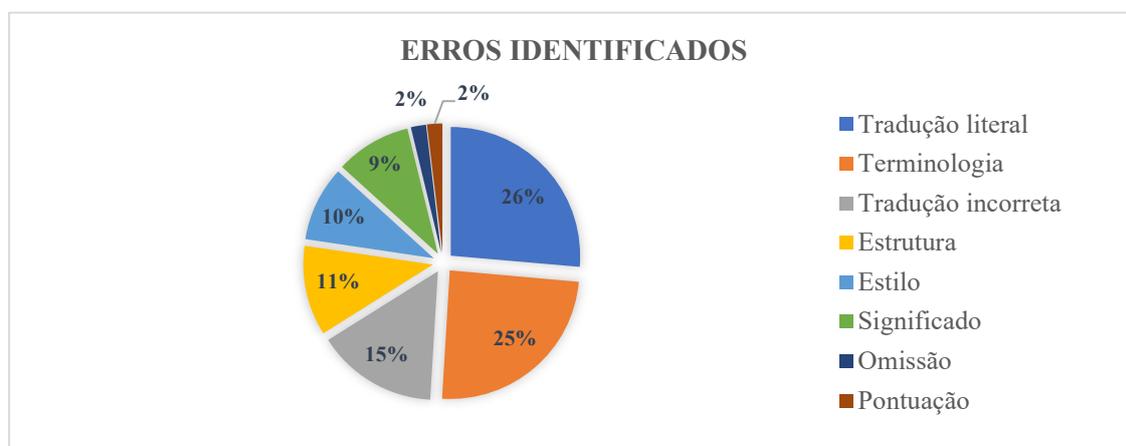


Gráfico 4 - Erros identificados no projeto

Nenhum dos erros identificados inclui segmentos repetidos, pois durante a tarefa de pós-edição, devido à funcionalidade de autopropagação, não foi necessário editar as frases repetidas.

Além disso, de todos os erros identificados, 15 foram considerados erros de gravidade maior. A maioria eram erros de tradução literal. Também foram identificados alguns erros de terminologia graves. Os restantes foram considerados como erros menores. No total, foram somados 84 pontos. Sendo o máximo permitido apenas 13 pontos, isto significa que, caso a

²⁶ Incluindo os segmentos repetidos

tradução automática sugerida no documento analisado fosse alvo de avaliação, esta não seria considerada aceitável.

A Tabela 13 inclui alguns dos erros identificados neste projeto e demonstra como o desempenho do sistema não conseguiu manter uma linguagem natural.

Alteração preferencial	Erro estilístico	Erro estrutural	Terminologia	Tradução incorreta	Tradução literal
Texto de partida		Sugestão de tradução automática		Proposta de tradução final	
<i>Here's to new beginnings between you and the pavement.</i>		Aqui 's aos começos novos entre você e o pavimento.		Aos novos começos entre ti e o asfalto.	
<i>Back with extra bounce that's perfect for hitting the pavement.</i>		Costas com ressaltado adicional que 's perfeito para bater no pavimento.		Contam com ressaltado adicional perfeito para enfrentar a estrada.	
<i>Whether you're racking up everyday miles or on your long run, feel the spring in your step with the same cushioned support as its predecessor.</i>		Quer estejas 's a acumular quilómetros todos os dias ou a correr muito tempo, sente a primavera na tua passada com o mesmo suporte amortecido que o seu antecessor.		Quer estejas a acumular quilómetros diários ou numa corrida longa, dá impulso às tuas passadas com o mesmo suporte amortecido do modelo anterior.	

Tabela 13 - Erros de tradução literal

A Tabela 14 demonstra segmentos que incluem erros que não foram de encontro às orientações específicas do cliente, desde erros estilísticos a erros específicos à terminologia.

Alteração preferencial	Erro estilístico	Erro estrutural	Terminologia	Tradução incorreta	Tradução literal
Texto de partida		Sugestão de tradução automática		Proposta de tradução final	
<i>The road is your runway.</i>		A estrada é a sua pista.		A estrada é a tua passarela.	
<i>Made with at least 20% recycled content by weight, it combines the flexibility you love with a contained design that will help keep you close to the ground for that barefoot feeling.</i>		Fabricado com, pelo menos, 20% de conteúdo reciclado em termos de peso, combina a flexibilidade que adora com um design compacto que o ajuda a manter-se próximo do solo para uma sensação de pé descalço.		Fabricadas com, pelo menos, 20% de materiais reciclados, combinam a flexibilidade que adoras com um design com contenção que mantém a proximidade ao solo para proporcionar uma sensação de pé descalço.	

Tabela 14 - Erros segundo as orientações do cliente

Dos erros identificados como tradução literal, seis continham a transcrição dos chamados *tags*, por exemplo “It's”. Este conjunto de palavras, símbolos e números leva o sistema a assumi-lo como parte do texto e, por conseguinte, tenta traduzi-lo. Muitas vezes, isto leva à adição de palavras não existentes, uma vez que acaba por traduzir tudo separadamente.

As alterações preferenciais presentes na Tabela 13 foram realizadas consoante as memórias de tradução pré-existentes.

O principal erro estilístico cometido foi a utilização de linguagem formal em textos que se enquadram na categoria de linguagem informal. No segundo exemplo da Tabela 14, analisando o contexto em que o segmento se encontra, a palavra “made” deveria ter sido traduzida no género feminino no plural.

Exemplo 2

O segundo documento era maior que o primeiro exemplo, mas continha um número de segmentos com implementação de tradução automática consideravelmente menor (ver Tabela 15).

	Nr. segmentos	<i>Context match</i>	100%	<i>Fuzzy alta</i>	<i>Fuzzy baixa</i>	NMT
Total	891	x	809	70	x	12

Tabela 15 - Dados do projeto

À semelhança do primeiro projeto analisado, o número de segmentos que continham tradução automática neuronal é mais pequeno do que o número total de correspondências (*fuzzy matches*, 100% e *context matches*). Dos 12 segmentos que continham tradução automática neuronal, apenas nove sofreram algum tipo de alterações, preferenciais ou de correção:

- três segmentos não sofreram qualquer alteração;
- dois segmentos sofreram apenas alterações preferenciais;
- dois segmentos sofreram correções de algum tipo de erro;
- cinco segmentos sofreram tanto alterações preferenciais como correções.

Comparando com o primeiro documento analisado, observou-se que o desempenho do sistema de tradução automática neste projeto foi consideravelmente melhor. No entanto, cometeu alguns erros de gravidade séria, sendo identificados, pelo menos, cinco erros de nível grave. Todos estes erros foram principalmente traduções literais. Somando todos os erros encontrados, obteve-se uma pontuação de 34, o que significa que, mais uma vez, se a tradução do sistema de tradução automática fosse submetida a avaliação, esta não seria considerada de qualidade publicável.

Foram identificados sete erros de tradução literal, três erros de significado, três erros de terminologia, um erro de pontuação e um termo por traduzir.

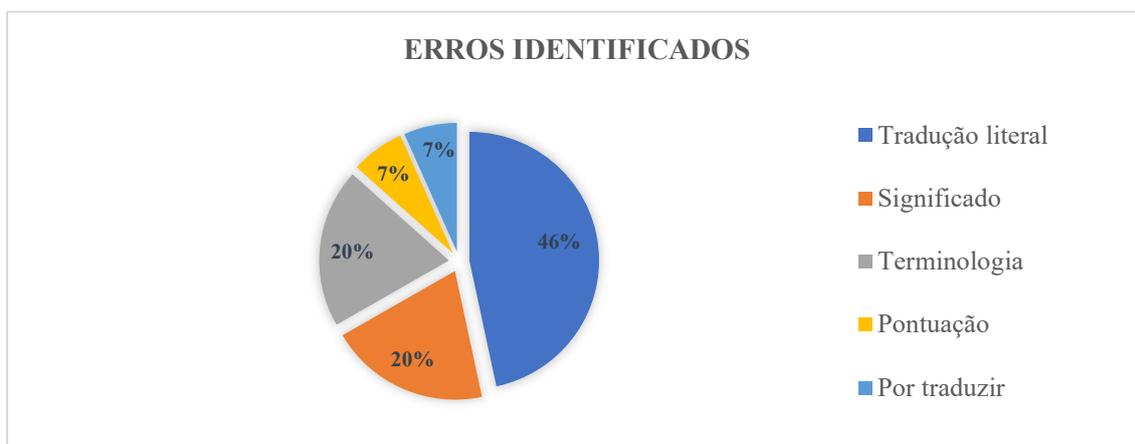


Gráfico 5 - Erros identificados no projeto

A Tabela 16 demonstra alguns dos exemplos de tradução literal sugerida pelo sistema de tradução automática.

Texto de partida	Sugestão de tradução automática	Proposta de tradução final
<i>Set fire to the field.</i>	Defina o fogo para o campo.	Entra em campo com um estilo arrojado.
<i>The socks deliver, keeping your feet comfortable through less than ideal conditions.</i>	As meias proporcionam , mantendo os pés confortáveis em menos do que as condições ideais.	As meias garantem aquilo de que precisas , mantendo os pés confortáveis quando as condições estão longe de ser as ideais.

Tabela 16 - Exemplo de significado errado

O primeiro exemplo demonstra como o sistema de tradução não consegue traduzir frases que contenham vários significados. O mesmo acontece no segundo exemplo, onde a palavra “deliver” tem uma função diferente daquela que foi traduzida.

Exemplo 3

Este terceiro exemplo é considerado um projeto pequeno, uma vez que só tinha cinco segmentos com tradução automática (ver Tabela 17). Mais uma vez, a maioria do projeto consistiu em segmentos bloqueados ou *fuzzy matches*.

	Nr. segmentos	Context match	100%	Fuzzy alta	Fuzzy baixa	NMT
Total	559	2	515	32	5	5

Tabela 17 - Dados do projeto

Dos cinco segmentos que continham tradução automática neuronal, todos sofreram algum tipo de alteração, preferenciais ou de correção:

- dois segmentos sofreram apenas alterações preferenciais;
- três segmentos sofreram tanto alterações preferenciais como correções.

No total foram identificados cinco erros. Mantendo a tendência dos projetos anteriores, três dos erros encontrados foram de tradução literal, um de terminologia e um de tradução incorreta.

A Tabela 18 demonstra um exemplo de tradução literal e tradução incorreta presente neste projeto.

Alteração preferencial	Tradução incorreta	Tradução literal
Texto de partida	Sugestão de tradução automática	Proposta de tradução final
<i>The tough rubber outsole was built for the rigors of running on pavement with an added bit of electric sparkle.</i>	A sola exterior em borracha resistente foi concebida para os rigores do running em pavimento com um toque adicional de brilho elétrico.	A sola exterior em borracha resistente foi criada para suportar as exigências de correr no pavimento com um toque adicional de brilho elétrico.

Tabela 18 - Exemplo de erros identificados

Como podemos observar, foi possível aproveitar grande parte da sugestão de tradução automática, sendo apenas preciso fazer algumas alterações preferenciais e correção de erros.

4.2.2. Projetos de descrição de produtos e *marketing*

Os projetos que combinavam descrição de produtos e *marketing* eram textos que, para além de exigirem uma linguagem específica da área de vestuário, utilizavam também a linguagem apelativa e criativa de *marketing*. O principal objetivo deste tipo de texto era descrever brevemente como o produto era feito e quem era o público-alvo. Geralmente, este tipo de textos continha menos segmentos bloqueados. Os projetos podiam apresentar conteúdo repetido ou não, o que significa que o número de segmentos com *fuzzies* e com tradução automática neuronal alterava consoante cada projeto. Tal como nos documentos de descrição de produtos, era recorrente utilizar a pesquisa de *concordance* para pesquisar principalmente as expressões utilizadas.

Exemplo 1

A Tabela 19 representa os dados do primeiro projeto a ser analisado desta categoria.

	Nr. segmentos	<i>Context match</i>	100%	<i>Fuzzy alta</i>	<i>Fuzzy baixa</i>	NMT
Total	121	6	80	12	6	17

Tabela 19 - Dados do projeto

Comparando este tipo de documento com os projetos de “descrição de produtos”, observou-se que este contém segmentos mais pequenos e com menos informação. Além disso, conta também com vários segmentos que não devem ser traduzidos, segundo o Cliente A. Ainda assim, o número de segmentos com palavras novas continua a ser mais baixo que o número de segmentos já com correspondência ou bloqueados. Isto significa que este tipo de projetos se inclui na categoria de projetos de conteúdo repetido.

Dos 17 segmentos que continham tradução automática neuronal, oito eram segmentos repetidos. Além disso, a maioria²⁷ dos segmentos sofreu alterações preferenciais e de correção:

- dois segmentos não sofreram qualquer alteração;
- 10 segmentos sofreram correções de algum tipo de erro;
- cinco segmentos sofreram tanto alterações preferenciais como correções.

²⁷ Incluindo os segmentos repetidos

De todos os erros identificados, a categoria de terminologia foi a mais comum, com seis erros, seguida de quatro erros estilísticos, dois erros de omissão, um significado errado e um erro ortográfico e uma tradução incorreta. Curiosamente, e ao contrário de muitos projetos deste tipo, este projeto não apresentou nenhum erro de tradução literal.

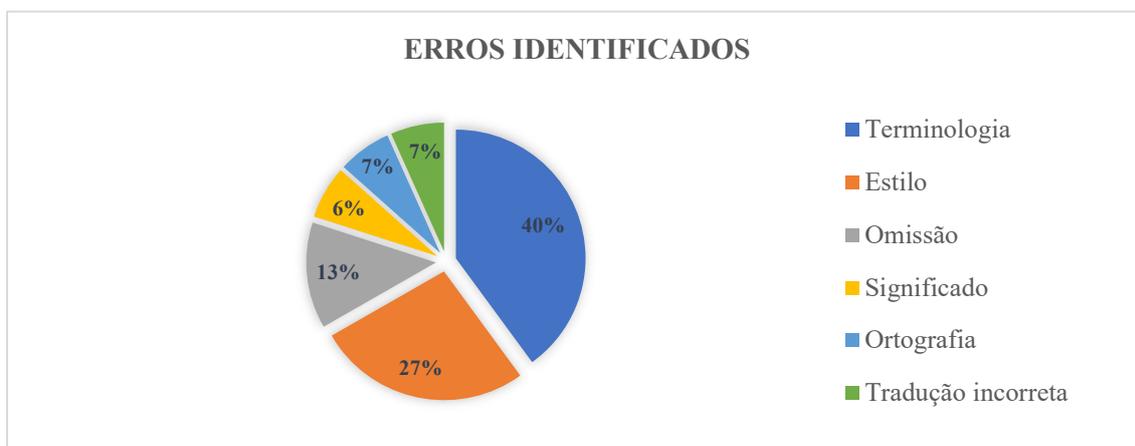


Gráfico 6 - Erros identificados no projeto

Nenhum dos erros identificados inclui os segmentos repetidos, pois durante a tarefa de pós-edição, devido à funcionalidade de autopropagação, não foi necessário editar as frases repetidas. Além disso, o sistema de tradução automática traduziu todos os segmentos que não deviam ser traduzidos, sendo que foram considerados erros estilísticos. Também a tradução dos nomes dos produtos foi considerada um erro estilístico porque não respeita as indicações do cliente.

Outro fator interessante deste projeto é que nenhum dos erros foi considerado de nível grave, uma vez que todos os segmentos conseguiram transmitir a mensagem desejada. Isto significa que, caso a tradução automática sugerida fosse alvo de avaliação, embora ainda não conseguisse atingir a qualidade desejada, a pontuação dos erros seria consideravelmente mais pequena, com 24 pontos.

A Tabela 20 apresenta alguns exemplos de como o sistema não segue as orientações do cliente e é necessária a intervenção do tradutor para as corrigir. A orientação em causa é o uso de linguagem neutra e inclusiva.

Erro estilístico		Significado
Texto de partida	Sugestão de tradução automática	Proposta de tradução final
<i>Designed to keep you warm and dry.</i>	Concebido para te manter quente e seco .	Concebido para manter o aquecimento e a secura .
<i>Why Do I Keep Getting Hurt?</i>	Por que motive me sinto magoado ?	Por que razão me estou sempre a lesionar ?

Tabela 20 - Exemplo de erro estilístico e tradução incorreta

Embora a tradução do primeiro exemplo esteja gramaticalmente correta, o sistema de tradução automática optou por traduzir a frase de forma a utilizar a combinação de pronome pessoal com um verbo conjugado, em vez de utilizar nominativos para evitar a utilização de géneros. No segundo exemplo, para além de ter traduzido a frase com o significado errado, o sistema voltou a sugerir a mesma opção.

Exemplo 2

O segundo exemplo é consideravelmente mais pequeno que o exemplo analisado anteriormente, com menos correspondências no geral. No entanto, incluiu mais segmentos com integração de tradução automática neuronal.

	Nr. segmentos	Context match	100%	Fuzzy alta	Fuzzy baixa	NMT
Total	67	x	26	2	3	34

Tabela 21 - Dados do projeto

Cerca de 50% do projeto continha palavras novas. Isto significa que o documento abrangia mais conteúdo que ainda não tinha sido traduzido. Devido a alguma falha por parte do sistema, dois segmentos não incluíram a tradução automática implementada. Dos 34 segmentos que continham tradução automática neuronal, oito eram segmentos repetidos. Desses segmentos, 33²⁸ sofreram algum tipo de alterações, preferenciais ou de correção:

- um segmento não sofreu qualquer alteração;
- sete segmentos sofreram apenas alterações preferenciais;

²⁸ Incluindo os segmentos repetidos

- 15 segmentos sofreram correções de algum tipo de erro;
- 11 segmentos sofreram tanto alterações preferenciais como correções.

De todas as categorias presentes, os erros de terminologia foram os mais predominantes, com 15 erros. Foram também identificados 11 erros de tradução literal, nove de tradução incorreta, nove erros estilísticos, sete significados errados, cinco erros de estrutura, dois erros de palavras por traduzir, dois erros de pontuação e um erro ortográfico.

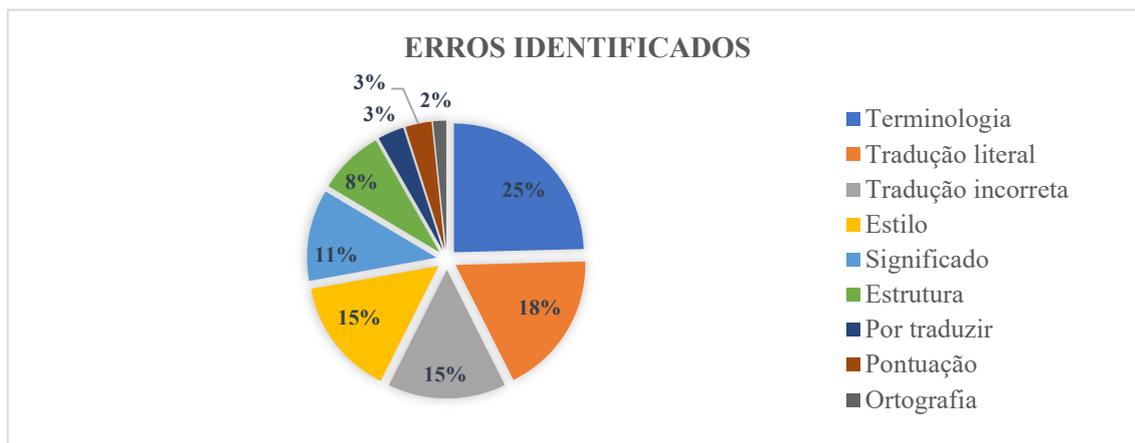


Gráfico 7 - Erros identificados no projeto

Este documento foi um dos que apresentou mais categorias de erros durante a análise. É interessante salientar que, ao contrário do primeiro exemplo, este segundo projeto pareceu apresentar erros mais diversos e menos previsíveis. Observou-se que este documento continha segmentos maiores e com mais informação. Além disso, apresentou mais erros relacionados com tradução literal. É possível relacionar estes dois fatores para justificar o desempenho mais pobre do sistema de tradução automática.

Além disso, de todos os erros analisados, 11 foram identificados como erros de gravidade séria. No total foram somados, pelo menos, 30 pontos. Mais uma vez, o sistema de tradução automática, se fosse avaliado, não conseguiria garantir a qualidade pretendida pelo formulário da empresa.

A Tabela 22 inclui alguns exemplos da tradução literal incluída no documento.

Por traduzir	Erro estilístico	Erro estrutural	Tradução incorreta	Tradução literal
Texto de partida	Sugestão de tradução automática		Proposta de tradução final	
<i>If you thought this staple couldn't get any better, we decked out the upper with white synthetic leather, made with at least 25% recycled polyester. Looks dope fresh out the box and breaks in beautifully, so you can keep these classics in rotation for a while.</i>	Se achaste que esta peça essencial não podia ser melhor, enganámos a parte superior com pele sintética branca, fabricada com, pelo menos, 25% de poliéster reciclado. Parece uma nova forma de descobrir a caixa e parte-se de forma maravilhosa , para que possa manter estes clássicos em rotação durante algum tempo.	Se achavas que este essencial não poderia ficar melhor, adornámos a parte superior com pele sintética branca, fabricada com, pelo menos, 25% de poliéster reciclado. Tem um look fantástico e adapta-se elegantemente , para que possas manter estes clássicos em rotação durante algum tempo.		
<i>Dance is a primal pastime dating back to the Stone Age. Jk, we just made that up. But what we know for certain is that it feels good to move.</i>	A dança é um passatempo primordial que remonta à idade da pedra. JK, acabamos de fazer isso. Mas o que sabemos para certo é que se sente bem a mexer.	A dança é um passatempo primordial que remonta à idade da pedra. Estamos a brincar, acabamos de inventar isso. No entanto , sabemos que sabe bem mexer.		

Tabela 22 - Exemplos de tradução literal

4.2.3. Projetos de artigos e entrevistas

Esta categoria, tal como o nome indica, abrangia artigos, textos de blogues e entrevistas e tinha como objetivo utilizar uma linguagem mais informal, amigável e livre. Geralmente, estes projetos continham mais palavras por traduzir e segmentos com integração de tradução neuronal, em relação aos textos mencionados anteriormente. Isto significava que a tradução não podia

dependem tanto das memórias de tradução, uma vez que os textos eram diferentes e era raro existirem frases completas ou parágrafos idênticos guardados nas bases de dados.

Exemplo 1

O documento continha menos segmentos com correspondência de 100%, menos *fuzzy matches* e mais segmentos com tradução automática NMT integrada. Além disso, o número de segmentos com palavras novas foi consideravelmente maior, em comparação com todos os exemplos já analisados.

	Nr. segmentos	Context match	100%	Fuzzy alta	Fuzzy baixa	NMT
Total	23	x	7	4	x	12

Tabela 23 - Dados do projeto

Dos 12 segmentos que continham tradução automática neuronal, todos sofreram algum tipo de alterações, preferenciais ou de correção:

- seis segmentos sofreram correções de algum tipo de erro;
- seis segmentos sofreram tanto alterações preferenciais como correções.

Entre os nove segmentos que apresentaram algum tipo de erro, foram identificados 17 erros de tradução literal, 15 erros estilísticos, nove erros de significado, seis erros de ortografia, cinco erros de terminologia, cinco erros estruturais, quatro erros de pontuação e três erros de texto deixado por traduzir. Além disso, dos erros identificados, pelo menos 10 foram considerados de gravidade séria.

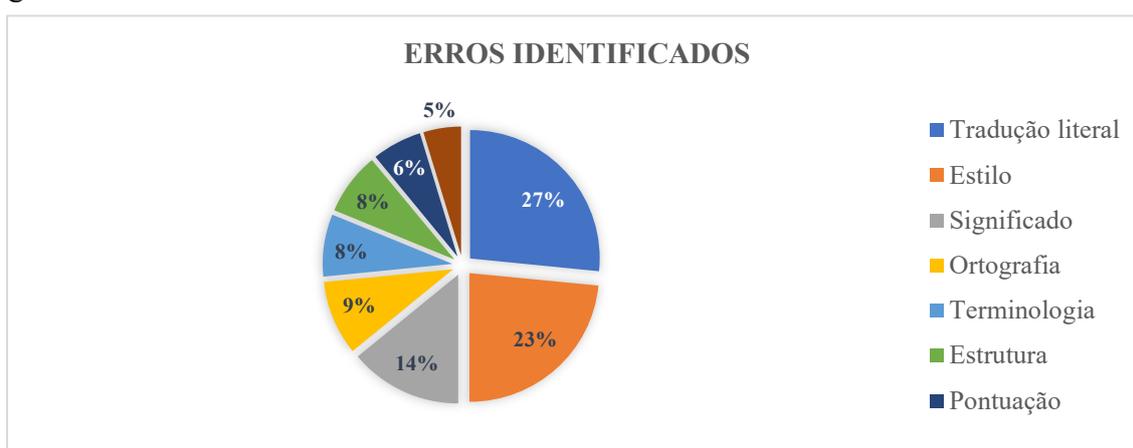


Gráfico 8 - Erros identificados no projeto

Depois da análise feita, foi possível reconhecer que o desempenho de qualidade do sistema de tradução automática não foi tão elevado em comparação com os outros tipos de textos. Isto implicava um trabalho mais meticuloso e era, por vezes, impossível aproveitar as sugestões

dadas pelo sistema. Durante o período de estágio, também se reparou que a tradução automática aplicada a este tipo de textos tinha tendência a traduzir as frases em português do Brasil, desde a conjugação dos verbos, à omissão de artigos definidos e à estrutura frásica.

A Tabela 24 apresenta um exemplo em que o desempenho da tradução automática obrigou a uma edição mais complexa, não sendo possível aproveitar a própria sugestão do sistema. O exemplo apresenta vários tipos de erros, mas a maioria provém da tradução literal que o sistema fez.

Alteração preferencial	Ortografia	Erro estrutural	Tradução incorreta	Tradução literal
Texto de partida	Sugestão de tradução automática		Proposta de tradução final	
<i>The co-founders have a natural choreography to their back-and-forth as they retrace their steps together over the years. Founded in 2010, they quickly built a reputation for their precise routines. "It started as this just-for-fun, meet-up-at-barbecues, and just perform-in-little-community-shows kind of thing," says.</i>	Os co-fundadores têm uma coreografia natural para os seus movimentos de costas e de frente enquanto retraçam os seus passos juntos ao longo dos anos. Fundada em 2010, rapidamente construíram uma reputação pelas suas rotinas precisas. "Começou como este tipo de reunião-para-diversão-para-churrascos-e apenas de performance-em-comunidade-mostra tipo de coisa ", diz.	Os cofundadores criam uma coreografia enquanto falam sobre os seus passos juntos ao longo dos anos. Fundado em 2010, este grupo rapidamente conquistou uma reputação com a precisão das suas coreografias . "Começou apenas por diversão, encontros durante churrascos e espetáculos em pequenas comunidades ", explica.		

Tabela 24 - Exemplo de tradução literal

O exemplo apresenta vários tipos de erros, mas a maioria provém da tradução literal que o sistema fez. É possível dizer que tal aconteceu devido à complexidade da frase: o facto de conter mais informação e ter mais do que uma frase pode ser um fator importante no desempenho do sistema de tradução. Além disso, também foi identificado um erro de ortografia, um erro de concordância e algumas traduções incorretas.

A Tabela 25 apresenta outros erros identificados no mesmo projeto. Embora o exemplo não contenha casos de tradução literal, a tradução continua a apresentar palavras que não representam a escrita natural exigida nem seguem as orientações do Cliente A.

Alteração preferencial	Erro estilístico	Pontuação	Significado	Tradução incorreta	Por traduzir
Texto de partida		Sugestão de tradução automática		Proposta de tradução final	
<p><i>“There are a lot of sports where you need to be a certain body type,” says. “The unique thing about dance is that whatever body you were born with, that is your tool. Maybe you’re skinny and maybe that doesn’t work for other sports, but when you start to dance, you realize, ‘I’m noodly, I could work with this!’</i></p>		<p>"Há muitos desportos onde precisa de ser um determinado tipo de carroçaria", afirma. "A única coisa sobre dança é que qualquer corpo com o qual you nasceu, essa é a sua ferramenta. Talvez you esteja skinny e talvez isso não funcione para outros Esportes, mas quando you começa a dançar, you percebe, “ Eu sou notavelmente, eu poderia trabalhar com isso!"</p>		<p>"Há muitos desportos em que é preciso ter um determinado tipo de corpo", afirma. "O que é peculiar na dança é que qualquer que seja o corpo com que nascestes, é o teu instrumento". Talvez sejas magro e isso não resulte para outros desportos, mas quando começas a dançar percebes: "Não sou magricela, serve perfeitamente!"</p>	

Tabela 25 - Outros exemplos

Além de o exemplo demonstrar um desempenho de pouca qualidade, o mais surpreendente foi a utilização do português do Brasil, algo que não foi encontrado nos restantes tipos de textos analisados. Reparou-se que quanto mais complexo fosse o segmento, em termos de tamanho e informação, mais havia a probabilidade de o sistema de tradução automática sugerir traduções com português do Brasil. O motivo para tal acontecer não foi explorado, mas seria interessante poder investigar este assunto no futuro.

Exemplo 2

O segundo documento continha mais segmentos com tradução automática neuronal que o primeiro documento analisado (ver Tabela 26).

	Nr. segmentos	Context match	100%	Fuzzy alta	Fuzzy baixa	NMT
Total	33	x	10	1	x	22

Tabela 26 - Dados do exemplo

À semelhança do primeiro exemplo, o número de segmentos com alguma correspondência (*fuzzy matches*, 100% e *contexto matches*) é reduzido quando comparado com o número de segmentos que contém palavras novas. Dos 22 segmentos que continham tradução automática neuronal, apenas um não sofreu qualquer alteração. Os restantes sofreram alterações preferenciais e de correção:

- um segmento não sofreu qualquer alteração;
- três segmentos sofreram correções de algum tipo de erro;
- 18 segmentos sofreram tanto alterações preferenciais como correções.

Tal como aconteceu no exemplo analisado anteriormente, a maioria dos segmentos sofreu alguma alteração. Observou-se que, mais uma vez, houve uma tendência para usar tradução literal, não traduzir certos termos, e acima de tudo, utilizar o português do Brasil na tradução. Foram identificados 35 erros de tradução, 27 erros estilísticos, 24 erros de tradução literal, 17 erros de significado, 16 erros estruturais, seis termos por traduzir, cinco erros de terminologia, três erros ortográficos e dois erros de pontuação.

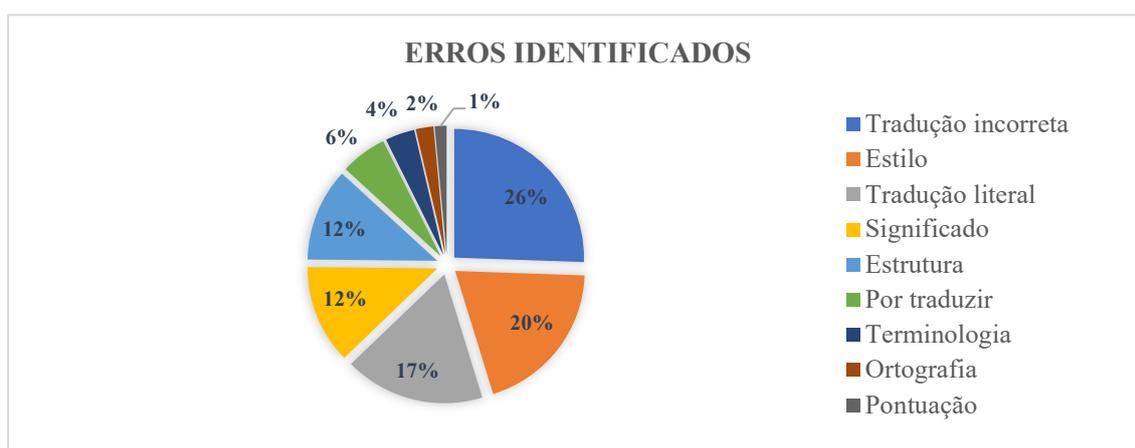


Gráfico 9 - Erros identificados no projeto

Dentro da categoria de tradução incorreta temos termos que foram traduzidos corretamente, segundo o dicionário, mas os termos utilizados não se adequaram ao contexto do documento. Já a maioria dos erros estilísticos identificados não cumpriu com a orientação do uso da linguagem informal do cliente. Além disso, os erros estruturais incluíram conjugações erradas

e problemas de concordância, tanto de registo como de género. A tradução apresentou uma quantidade considerável de erros graves, contando com 19 erros graves. Mais uma vez, a maioria deles deve-se à tradução literal feita pelo sistema, sugerindo um texto confuso.

Os exemplos da Tabela 27 apresentam alguns dos exemplos que demonstram os erros mais comuns e comprovam a baixa qualidade do sistema de tradução automática.

Alteração preferencial	Erro estilístico	Terminologia	Significado	Tradução incorreta	Por traduzir
Texto de partida	Sugestão de tradução automática			Proposta de tradução final	
<i>And because each foot stride during a run exerts a force about six times your body weight, under- or overpronation can prevent your body from optimally distributing that force, which can lead to a host of injuries.</i>	E porque cada batida do pé durante uma corrida exerce uma força aproximadamente seis vezes seu peso do corpo , sob-ou overpronation pode impedir seu corpo de distribuir optimally essa força, que pode conduzir a um anfitrião dos ferimentos			Além disso , dado que cada passada exerce uma força equivalente a cerca de seis vezes o teu peso corporal , a subpronação ou hiperpronação podem impedir o teu corpo de distribuir a força de forma ideal , o que pode resultar em muitas lesões .	
<i>To deal with plantar pain and help prevent full-blown fasciitis from striking in the future, take a look at your shoes, says. Because the plantar fascia supports the arch of your foot and absorbs shock when you walk and run, a stability shoe with a firmer midsole on the arch side of the foot and a lighter, softer foam on the outside can maximize shock absorption and counter overpronation, which can contribute to plantar fasciitis.</i>	Para lidar com a dor plantar e ajudar a evitar que a fascíte soprada de bater no futuro, dá uma vista de olhos às tuas sapatilhas, diz. Uma vez que a fâscia plantar suporta o arco do pé e absorve o choque quando caminha e corre , uma sapatilha de estabilidade com uma sola intermédia mais firme no lado do arco do pé e uma espuma mais leve e macia no exterior pode maximizar a absorção de choques e contra-pronação , que podem contribuir para fascíte plantar.			Para enfrentar a dor na planta do pé e ajudar a evitar a ocorrência da fascíte no futuro, dá uma vista de olhos às tuas sapatilhas, diz. Como a fâscia plantar suporta a planta do pé e absorve os impactos quando caminhas ou corres, umas sapatilhas de estabilidade com uma sola intermédia mais firme na zona da planta do pé e uma espuma mais leve e mais suave no exterior podem maximizar a absorção de impactos e ajudar na hiperpronação , que pode contribuir para a fascíte plantar.	

Tabela 27 - Erros identificados

Como podemos observar, os segmentos são maiores. Por esse motivo, acredita-se que esse é um dos fatores que possa ter provocado uma tradução de baixa qualidade.

Alteração preferencial	Erro estilístico	Tradução literal	Significado	Tradução incorreta	Por traduzir
Texto de partida		Sugestão de tradução automática		Proposta de tradução final	
<i>First and foremost: The moment you notice the telltale stabbing pain, address the inflammation. "This is the alarm phase," says. Ignore your barking dogs and you could end up sidelining yourself for weeks or months.</i>		Primeiro e acima de tudo: O momento em que você percebe a dor stabbing , tratar da inflamação. "Esta é a fase de alarme", diz. Ignore os seus cães que ladram e pode acabar por ficar sem parar durante semanas ou meses.		Acima de tudo: no momento em que sentires esta dor lancinante de alerta, trata da inflamação. "Esta é a fase de alarme", afirma. Se ignorares as dores, podes acabar por ter de te afastar durante semanas ou meses.	
<i>For extreme cases — your PT/sports doc should let you know if that's you — a night splint (don't laugh) can be effective. "When you're sleeping, your foot is kind of curled, so the tissues naturally will shorten," says. This is why plantar fasciitis tends to be most painful when you step out of bed in the morning. "A splint keeps the foot in a stretched position, so when you wake up, it hasn't shortened, and it's easier to put pressure on it," he explains.</i>		Para casos extremos o seu PT/documento desportivo deve informá-lo se é você uma farda noturna (não rir) pode ser eficaz. "Quando está a dormir, o pé está enrolado, para que os tecidos se encurtem naturalmente", afirma. É por isso que a fascíte plantar tende a ser mais dolorosa quando você sai da cama pela manhã. "UMA parte do pé mantém o pé esticado, por isso, quando acordar, não encurtou e é mais fácil exercer pressão", explica.		Para os casos extremos (o teu PT/médico desportivo deverá informar-te se esse é o teu caso), a utilização de talas de noite (não te rias) pode ser um tratamento eficaz. "Quando estás a dormir, o teu pé fica de certa forma enrolado, ou seja, os tecidos contraem naturalmente", afirma. É por este motivo que a fascíte plantar tende a ser pior quando te levantas da cama. "Uma tala mantém o pé numa posição alongada. Assim, quando acordas, não esteve contraído e é mais fácil fazer pressão sobre o mesmo", explica.	

Tabela 28 - Exemplos de erros identificados

A maioria dos erros aparenta ser resultado da tradução literal que o sistema faz. Isto acaba por provocar, não só erros de significado, mas também traduções incorretas de certos termos, produzindo assim uma tradução em que é preciso mais atenção e mais esforço durante a tarefa de edição.

4.3. Análise final

Apesar de todos os erros apresentados acima, a experiência de trabalhar com o sistema de tradução automática nas traduções dos projetos para o Cliente A foi, de uma forma geral, positiva. Se compararmos o desempenho do *SDL Machine Cloud* com outras plataformas *online* de tradução automática gratuitas, o sistema utilizado pela empresa é claramente mais especializado e consegue apresentar resultados mais concretos e corretos. No entanto, a última categoria de texto analisado não apresentou o mesmo desempenho.

Ao analisar todos os dados adquiridos, observou-se que as categorias de tradução literal são comuns a todos os tipos de textos e é geralmente a categoria que apresenta mais erros. No total, foram identificados 76 erros relacionados com tradução literal, seguidos de 54 erros relacionados com a tradução incorreta, 51 erros de estilo, 48 erros de terminologia e 42 erros de significado (ver Gráfico 12).

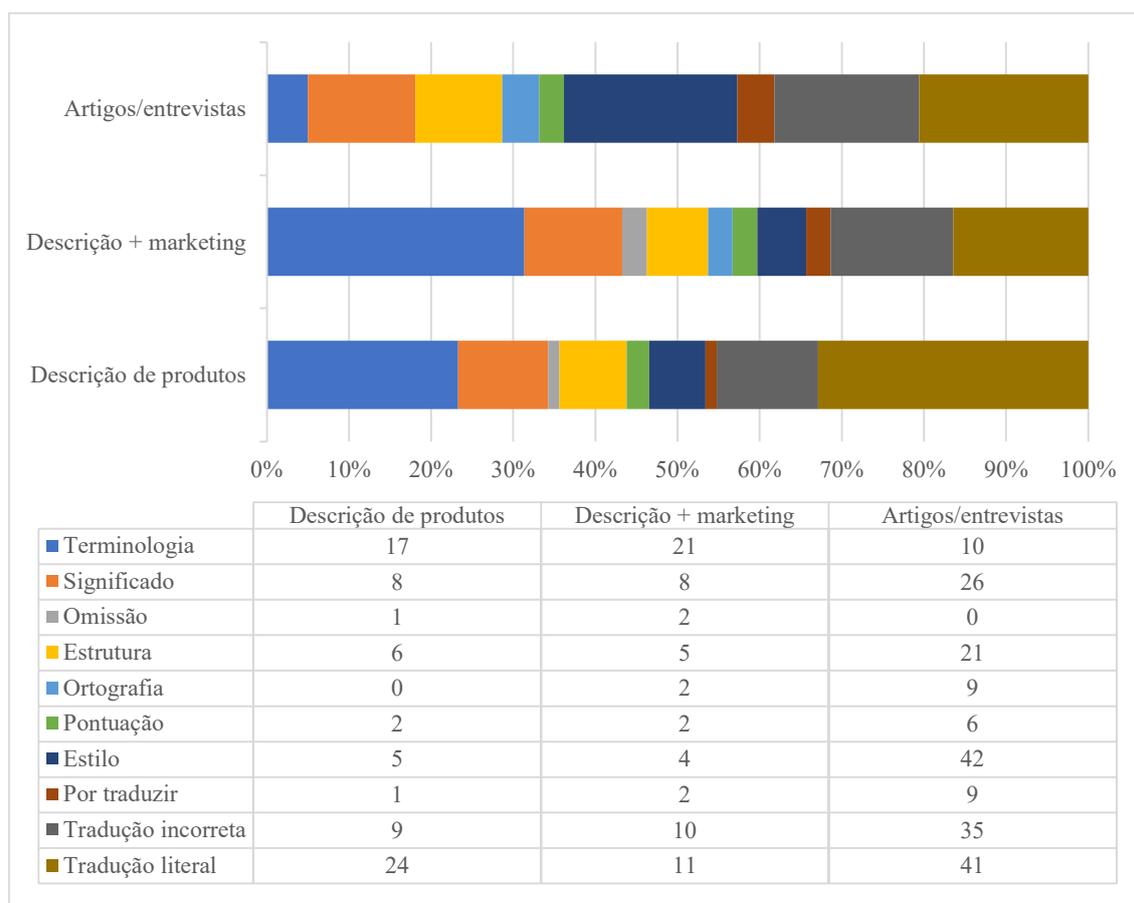


Gráfico 10 - Total de erros identificados

Apesar do total de erros identificados em todos os projetos, a tradução automática de cada tipo de texto parece ter um desempenho diferente, contendo diferentes categorias de erros. Por exemplo, enquanto o erro mais comum nos projetos “descrição de produtos” acontece devido à tradução literal, os projetos de artigos e entrevistas mostram uma maior quantidade de erros estilísticos. Isto permite averiguar que tipo de erros esperar durante a tarefa de edição de cada documento. No entanto, cada texto é um caso diferente e seria interessante explorar mais exemplos no futuro.

À semelhança do que Loureiro (2019, p.21) afirma, os resultados da tradução automática dependem de vários fatores, mas principalmente do tipo de texto em que o sistema está a ser utilizado. Embora, numa análise geral, os erros de tradução literal sejam consideravelmente mais, é possível ver que os textos que contêm segmentos maiores e mais complexos apresentam erros mais relacionados com as orientações do cliente e a fluência de escrita. Contrariamente, os tipos de textos que se focam mais na parte técnica do cliente apresentam erros relacionados com a terminologia.

Além disso, apesar de muitos erros serem de gravidade elevada, no geral, o sistema de tradução automática foi capaz de produzir textos que transmitissem a mensagem do texto. No entanto, é importante referir que, se os exemplos analisados fossem submetidos para avaliação, utilizando o formulário da RWS e adotando a avaliação *premium*, nenhum dos projetos conseguiria passar. Isto comprova que, embora as traduções fossem compreensíveis, sem a tarefa de pós-edição, não teriam a qualidade suficiente para serem aceites pelo cliente.

5. Da tradução automática à qualidade desejada

5.1. A edição das sugestões da tradução automática

Seguindo o processo de pós-edição demonstrado anteriormente²⁹ para a realização dos projetos durante o estágio, este capítulo mostra os passos necessários para tornar a tradução automática no produto final.

Durante o processo de edição, para além das memórias de tradução relativas ao cliente, era também adicionada uma memória de tradução temporária³⁰, criada no início do estágio (ver Figura 8). Além disso, era importante verificar se todas as memórias de tradução já incluídas no projeto não iriam ser atualizadas, uma vez que só os revisores as podiam atualizar com as traduções finais. Esta estratégia servia para não poluir as bases de dados das memórias de tradução. As principais memórias de tradução eram todas baseadas na nuvem. Isto permitia partilhar estas bases de dados em tempo real com outros tradutores (O'Hagan, 2013, p.504), e consequentemente, manter as memórias de tradução sempre atualizadas. Era especialmente útil quando dois projetos semelhantes eram feitos ao mesmo tempo, por exemplo um projeto em fase de tradução e outro em fase de revisão, pois era possível manter a consistência com a tradução final já revista.

²⁹ Consultar capítulo 1.3.2.

³⁰ Foi aconselhado criar uma memória de tradução própria para cada cliente para o tradutor ter sempre o seu trabalho guardado, de forma a prevenir qualquer problema.

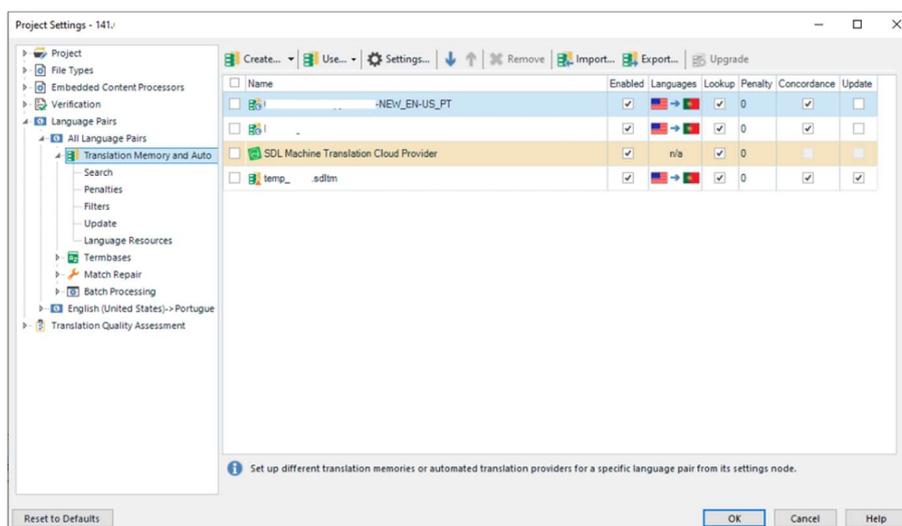


Figura 8 - Memórias de tradução de um projeto

Geralmente, os projetos já apresentavam a tradução automática aplicada, ou seja, era apenas preciso começar a editar o texto. O *SDL Machine Translation* estava programado para inserir automaticamente uma sugestão de tradução nos segmentos para serem editados (Zaretskaya, *et al.*, 2015, p.78). Porém, isto só acontecia quando o sistema de memória de tradução não conseguia sugerir uma correspondência de 100% ou um *fuzzy match* para o mesmo segmento (SDL, s.d.-d). A Figura 9³¹ demonstra como são apresentadas as sugestões de tradução automática.

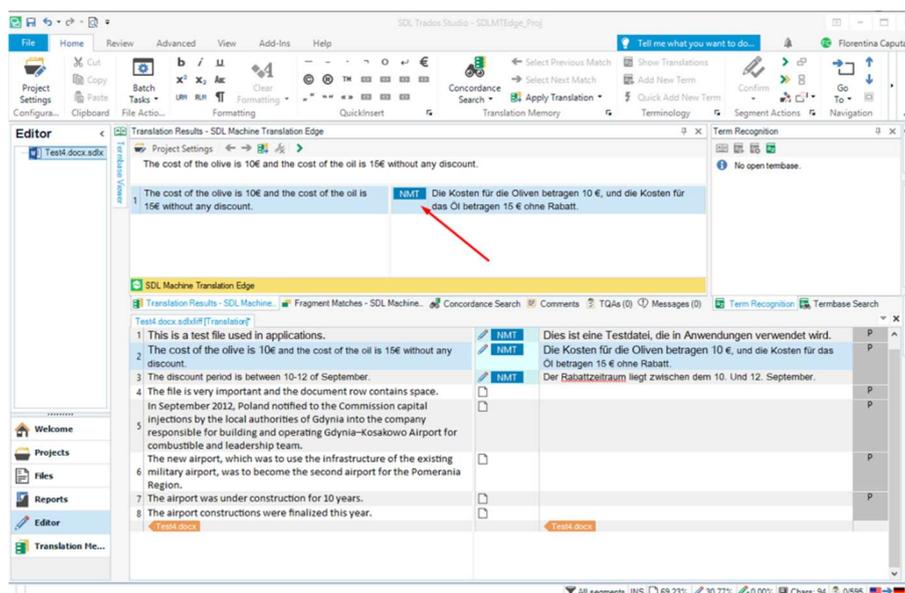


Figura 9 - Exemplo de tradução automática neuronal aplicada

³¹ <https://community.sdl.com/product-groups/translationproductivity/w/customer-experience/3450/sdl-machine-translation-edge-provider-former-sdl-ets>

Consoante a leitura de cada segmento, eram feitas as alterações necessárias, tendo sempre em mente as orientações dadas pela RWS e pelo Cliente A. Sempre que surgia um termo ou expressão que poderiam estar guardados nas memórias de tradução, era feita a pesquisa na *concordance*. Se o que procurava se encontrava nas memórias de tradução, bastava copiar e colar. No entanto, se o termo ou expressão procurados não existissem nas memórias de tradução e não soubesse qual a melhor forma para os traduzir, recorria a programas externos, como páginas de dicionários e tradutores para encontrar a melhor tradução possível. À medida que os segmentos eram editados, estes eram confirmados. No final, era feita uma verificação de qualidade do trabalho realizado, utilizando o *Xbench*³² ou a funcionalidade do *Trados Studio* e revendo aquilo que foi escrito. Todos os erros encontrados eram editados e considerava-se a tradução completa.

Depois disso, o projeto era entregue ao revisor. O papel do revisor passava por rever aquilo que o tradutor escreveu e produziu, e modificar o que fosse necessário para que o texto tivesse a qualidade desejada. Estas mudanças incluíam correções de potenciais erros por parte do tradutor e mudanças preferenciais. No final, todos os projetos eram entregues ao cliente.

³² Consultar capítulo 1.4.6.

Conclusão

Tal como Flanagan, *et al.*, (2010) afirma: “The line between TM and Machine Translation (MT – the automatic translation of text) is now becoming blurred” (p.190). Atualmente, muitas empresas de tradução apostam na combinação destas duas ferramentas para aproveitarem todas as vantagens que oferecem, modificando o mundo da tradução e obrigando o tradutor a adaptar o seu papel.

A RWS não é exceção, tirando o máximo partido destas tecnologias para realizar os seus projetos, principalmente técnicos. Como tal, a empresa teve de adotar certas medidas para tornar o ambiente de trabalho mais prático e produtivo, desde aproveitar as sugestões do sistema de tradução automática até utilizar a funcionalidade de pesquisa de *concordance*.

Embora o sistema de tradução automática neuronal seja o mais avançado até à data, este continua a apresentar problemas que obrigam o tradutor a intervir, uma vez que os resultados não correspondem muitas vezes à qualidade desejável. Para além disso, a qualidade desejável está sempre dependente das indicações, não só da empresa, mas também do cliente cujo projeto está a ser traduzido/pós-editado.

Assim, foi realizado um estudo de caso com o intuito de analisar o desempenho do sistema de tradução automática utilizado e compará-lo com os padrões de qualidade definidos pela empresa e o cliente. No geral, os resultados revelaram que, em todos os tipos de texto analisados, as sugestões não apresentam a qualidade considerada publicável, obrigando à tarefa de pós-edição. Observou-se que as categorias de erros mais comuns foram a tradução literal, tradução incorreta, estilo, terminologia e significado. No entanto, estas categorias divergiram muito consoante o tipo de texto, provando assim que o desempenho do sistema de tradução está dependente do tipo de conteúdo e do tipo de registo utilizado. Além disso, foram apresentados resultados de tradução automática mais satisfatórios quando havia uma repetição do conteúdo, ou seja, as sugestões do *SDL Machine Translation* foram, no geral, capazes de transmitir a mensagem original. Em relação ao esforço requerido por cada projeto, chegou-se à conclusão de que todos os tipos de texto exigem diferentes estratégias de edição por parte do tradutor. Ainda assim, é possível notar que os projetos de artigos e entrevistas foram os que obrigaram a fugir às recomendações de pós-edição da RWS, visto serem textos para os quais o sistema utilizado sugeriu traduções mais literais e menos fluentes.

Apesar dos resultados, seria interessante explorar um pouco mais este tema, uma vez que foi apenas possível analisar um número pequeno de exemplos. Por esse motivo, é impossível retirar conclusões concretas, mas é possível ter uma ideia daquilo que o sistema de tradução é

capaz. Todavia, a experiência de pós-edição durante o estágio foi bastante positiva e foi possível aprender mais sobre estas novas tecnologias e novas formas de tradução.

Bibliografia

- Allen, J. (2003). Post-editing. In H. Somer (Ed.), *Computers and translation: A translator's guide* (pp. 297-317). Paris, França: Benjamins Translation Library. doi:<https://doi.org/10.1075/btl.35>
- Alotaibi, H. M. (2020). Computer-assisted translation tools: An evaluation of their usability among arab translators. *Applied Sciences*, 10(18), 1-13. doi:<https://doi.org/10.3390/app10186295>
- American Translators Association. (2021). *Explanation of error categories*. Acedido em 2 de agosto de 2021, em: <https://www.atanet.org/certification/how-the-exam-is-graded/error-categories/>
- ApSIC, S.L. (2021). *Xbench.net*. Acedido em 13 junho de 2021, em: <https://www.xbench.net/>
- Arenas, A. G. (2008). Productivity and quality in the post-editing of outputs from translation memories and machine translation. *The International Journal of Localisation*, 7(1), 11-21. Acedido em 18 março de 2021, em: <https://core.ac.uk/download/pdf/228161879.pdf>
- Arenas, A. G., & Moorkens, J. (2019). Machine translation and post-editing training as part of a master's programme. *The Journal of Specialised Translation*(31), 217-238. Acedido em 4 de abril de 2021, em https://www.jostrans.org/issue31/art_guerberof.pdf
- Aziz, W., Sousa, S. C., & Specia, L. (2012). PET: A tool for post-editing and assessing machine translation. *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 3982–3987. Acedido em 24 de março de 2021, em: http://www.lrec-conf.org/proceedings/lrec2012/pdf/985_Paper.pdf
- Bennett, P. (2021) *The medical translator's toolkit: a personal reflection on DeepL Pro* [Unpublished paper]. Acedido em 25 de junho de 2021
- Beßler, P. (2021). *Post-editing and the evolution of translators*. Acedido em 11 de outubro de 2021, em: <https://www.rws.com/blog/what-is-post-editing/>
- Bowker, L. (2005). Productivity vs quality? A pilot study on the impact of translation memory systems. *Localisation Focus*, 13-20.
- Bundgaard, K., & Christensen, T. P. (2019). Is the concordance feature the new black? A workplace study of translators' interaction with translation resources while post-editing TM and MT matches. *The Journal of Specialised Translation*(31), 14-37. Acedido em 2 de abril de 2021, em: https://www.jostrans.org/issue31/art_bundgaard.pdf
- Career Framework Library. (2019). Acedido em 15 de maio de 2021, em: <https://sdl365.sharepoint.com/hr/Career%20Framework%20Library/Forms/Language%20Delivery%20%20Operations.aspx>
- CERTT. (2012). Glossary of translatio tool types. *Collection of electronic resources in translation technologies*. Acedido em 20 de março de 2021, em http://aix1.uottawa.ca/~certt/Glossary%20of%20translation%20tool%20types_FINAL.pdf
- Clarke, C. (2019). *Career Framework Library* . Acedido em 15 de maio de 2021, em: <https://sdl365.sharepoint.com/hr/Career%20Framework%20Library/P2%20Translation%20Intern.pdf>
- Davies, D. (2021). *Smart new features and capabilities added to SDL Machine Translation*. Acedido em 27 de setembro de 2021, em: <https://www.businesswire.com/news/home/20210224005746/en/Smart-New-Features-and-Capabilities-Added-to-SDL-Machine-Translation>
- DeepL. (2020). *Another breakthrough in AI translation quality*. Acedido em 25 de setembro de 2021, em: <https://www.deepl.com/en/blog/20200206>
- Doherty, S., & Gaspari, F. (2013). Understanding and implementing effective translation quality evaluation techniques. *7th Framework Programme of the European Commission*. Acedido em 21 de setembro de 2021, em: <https://www.qt21.eu/launchpad/sites/default/files/QTLP%20GALA%20Webinar%203.pdf>

- Flanagan, M., O'Hagan, M., & O'Brien, S. (2010). Keeping an eye on the UI design of Translation Memory: How do translators use the "Concordance" feature? *ECCE '10: Proceedings of the 28th Annual European Conference on Cognitive Ergonomics*, (pp. 187-190). Delft, Holanda. doi:10.1145/1962300.1962338
- Forcada, M. L. (2010). Machine Translation today. In Y. Gambier, & L. v. Doorslaer (Eds.), *Handbook of Translation Studies* (Vol. 1, pp. 215-223). John Benjamins Publishing Company. doi:<https://doi.org/10.1075/hts.1.mac1>
- Freitas, C., & Liu, Y. (2017). Exploring the differences between human and Machine Translation. *WWU Honors Program Senior Projects*. Acedido em 13 de agosto de 2021, em https://cedar.wvu.edu/cgi/viewcontent.cgi?article=1060&context=wwu_honors
- Garcia, I. (2009). Beyond translation memory: Computers and the professional translator. *The Journal of Specialised Translation*, 12(12), 199-214
- Garcia, I. (2014). Computer-aided translation. In C. Sin-wai, *The Routledge Encyclopedia of Translation Technology* (pp. 68-87). Reino Unido: Routledge. doi:<https://www.routledgehandbooks.com/doi/10.4324/9781315749129.ch3>
- Gaspari, F., Toral, A., Naskar, S. K., Groves, D., & Way, A. (2014). Perception vs reality: Measuring machine translation post-editing productivity. (S. O'Brien, M. Simard, & L. Specia, Eds.) *Proceedings of the Third Workshop on Post-editing Technology and Practice (WPTP-3)*, pp. 60-72. Acedido em 18 de março de 2021, em: http://www.amtaweb.org/AMTA2014Proceedings/AMTA2014Proceedings_PEWWorkshop_final.pdf
- Ghislandi, M. (2018). *SDL Trados Studio 2019 Has Arrived*. Acedido em 13 de setembro de 2021, em: SDL Trados Studio: <https://www.trados.com/blog/sdl-trados-studio-2019-is-here.html>
- Han, B. (2020). Translation, from pen-and-paper to computer-assisted tools (CAT Tools) and machine translation (MT). *Proceedings*, 63(56). doi:<https://doi.org/10.3390/proceedings2020063056>
- Heyn, M. (1996). Integrating machine translation into translation memory systems. *EAMT Machine Translation Workshop*, pp. 113-126. Acedido em 31 de março de 2021, em: <https://www.aclweb.org/anthology/1996.eamt-1.12.pdf>
- Hutchins, J. (1998). The origins of the translator's workstation. *Machine Translation*, 13(4), pp. 287-307. doi:<https://doi.org/10.1023/A:1008123410206>
- Jia, Y., Carl, M., & Wang, X. (2019). How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study. *The Journal of Specialised Translation*(31), 60-86. Acedido em 4 de abril de 2021, em: https://www.jostrans.org/issue31/art_jia.pdf
- Kartsaklis, D., & Kanavos, P. (2010). Integrating machine translation with translation memory: A practical approach. In V. Zhechev (Ed.), *Proceedings of 2nd Joint EM+/CNGL Workshop "Bringing MT to the User: Research on Integrating MT in the Translation Industry"*, (pp. 11-20). Denver, EUA
- Koglin, A., & Cunha, R. (2019). Investigating the post-editing effort associated with machine-translated metaphors: a process-driven analysis. *The Journal of Specialised Translation*(31), 38-59. Acedido em 3 de abril de 2021, em: https://www.jostrans.org/issue31/art_koglin.pdf
- Kurniawati, L. A., Rahajeng, D. T., Kristanto, B., & Kastuhandani, F. C. (2016). Introducing SDL Trados to beginning translators. *Indonesian Journal of English Language Studies*, 2(1), 91-105. doi:10.24071/IJELS.V2I1.353
- LeBlanc, M. (2013). Translators on translation memory (TM). Results of an ethnographic study in three translation services and agencies. *The International Journal for Translation & Interpreting Research*, 5(2), 1-13. doi: ti.105202.2013.a01

- Liu, X., Lai, H., Wong, D. F., & Chao, L. S. (2020). Norm-based curriculum learning for neural machine translation. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. doi:10.18653/v1/2020.acl-main.41
- Loureiro, A. S. (2019). Linguistic challenges of translating, revising, post-editing and editing academic publications. pp. 6-43. Acedido em 23 de março de 2021, em: <http://hdl.handle.net/10400.22/15142>
- Moorkens, J., & Way, A. (2019). Post-editing neural machine translation versus translation memory segments. *Machine Translation*(33), 1-29. doi:10.1007/s10590-019-09232-x
- O'Brien, S., & Moorkens, J. (2014). Towards intelligent post-editing interfaces. *Proceedings of the XXth FIT World Congress*. Alemanha: BDU Fachverlag. Acedido em 18 de março de 2021, em: https://www.researchgate.net/publication/275039972_Towards_Intelligent_Post-Editing_Interfaces
- O'Hagan, M. (2013). The impact of new technologies on translation studies: A technological turn? In *The Routledge Handbook of Translation Studies* (pp. 503-518). Londres e Nova Iorque: Taylor & Francis Group. doi:10.4324/9780203102893.ch37
- Okpor, M. D. (2014). Machine translation approaches: issues and challenges. *International Journal of Computer Science Issues*, 11(2), 159-165. Acedido em 28 de agosto de 2021, em: <https://www.ijcsi.org/papers/IJCSI-11-5-2-159-165.pdf>
- Polo, L. R. (2013). Managing the translation workflow with a computer assisted translation tool: SDL Trados 2011. *Language Value*, 5(1), pp. 161-174. doi:https://doi.org/10.6035/LanguageV.2013.5.10
- Prates, M. O., Avelar, P. H., & Lamb, L. C. (2019). Assessing gender bias in machine translation: a case study with Google Translate. *Neural Computing and Applications*, 32, pp. 6363-6381. doi:https://doi.org/10.1007/s00521-019-04144-6
- QT21. (2014). *Guide to selecting MQM issues for the MT Evaluation Metric*. Acedido em 25 de julho de 2021, em: <https://www.qt21.eu/downloads/annotatorsGuidelines-2014-06-11.pdf>
- QTLaunchPad. (2015). *Multidimensional Quality Metrics (MQM) definition*. Acedido em 18 de julho de 2021, em: <https://www.qt21.eu/mqm-definition/definition-2015-12-30.html>
- Reinke, U. (2018). State of the art in Translation Memory Technology. (G. Rehm, D. Stein, F. Sasaki, & A. Witt, Eds.) *Language technologies for a multilingual Europe: TC3 III*, pp. 55-84. doi:10.5281/zenodo.1291947
- RWS. (2020). *Acquisition of SDL by RWS enhances capabilities in Language Services and Technology*. Acedido em 15 de abril de 2021, em: <https://www.rws.com/news/acquisition-of-sdl-by-rws-enhances-capabilities-in-language-services-and-technology/#>
- RWS. (n.d.). *SDL Machine Translation Cloud (former BeGlobal)*. Acedido em 5 de setembro de 2021, em: <https://community.sdl.com/product-groups/translationproductivity/w/customer-experience/3356/sdl-machine-translation-cloud-former-beglobal>
- RWS Trados. (2020). *Trados GroupShare - Translation Project Collaboration*. Acedido em 15 de abril de 2021, em: <https://www.trados.com/products/groupshare/>
- RWS Trados. (n.d.). *What is machine translation?* Acedido em 15 de abril de 2021, em: <https://www.trados.com/solutions/machine-translation/>
- RWS. (n.d.). *What is a Translation Memory?* Acedido em 7 de outubro de 2021, em: <https://www.trados.com/solutions/translation-memory/>
- SAE International. (2016). *Translation Quality Metric (STABILIZED Aug 2016) J2450_201608*. Acedido em 20 de setembro de 2021, em: https://www.sae.org/standards/content/j2450_201608/

- Schjoldager, A., & Christensen, T. P. (2017). Translation-Memory (TM) research: What do we know and how do we know it? *HERMES - Journal of Language and Communication in Business*, 23(44), pp. 89-101. doi:10.7146/hjlc.v23i44.97268
- SDL AppStore. (2021). *SDLXLIFF Compare*. Acedido em 13 de junho de 2021, em: <https://appstore.sdl.com/language/app/sdlxliff-compare/89/>
- SDL Limited. (2018). *Translation 101 - What is a CAT tool?* Acedido em 15 de abril de 2021, em: <https://www.trados.com/solutions/cat-tools/translation-101-what-is-a-cat-tool.html>
- SDL plc. (n.d.). *About SDL*. Acedido em 15 de abril de 2021, em: <https://www.sdl.com/about/>
- SDL. (n.d.). *Quality Assessment Models*. Acedido em 25 de setembro de 2021, em: <https://docs.rws.com/791627/788067/sdl-multitrans/quality-assessment-models>
- SDL. (n.d.). *SDL BeGlobal*. Acedido em 6 de setembro de 2021, em: <https://docs.sdl.com/783545/576252/sdl-trados-studio/sdl-beglobal>
- SDL. (n.d.). *SDL Documentation*. Acedido em 13 de junho de 2021, em: <https://docs.sdl.com/785465/787392/sdl-multitrans/setting-timesheet-categories>
- SDL. (n.d.). *SDL Language Cloud Machine Translation*. Acedido em 17 de julho de 2021, em: <https://docs.sdl.com/783545/636537/sdl-trados-studio/sdl-language-cloud-machine-translation>
- SDL. (n.d.). *Translating with Machine Translation (MT) - SDL Trados Studio 2019*. Acedido em 16 de abril de 2021, em: <https://docs.sdl.com/783545/537149/sdl-trados-studio/translating-with-machine-translation--mt->
- Stein, D. (2013). Machine translation: Past, present and future. (G. Rehm, F. Sasaki, D. Stein, & A. Witt, Eds.) *Language technologies for a multilingual Europe: TC3 III*, pp. 5-17. doi:10.5281/zenodo.1291924
- thebigword. (n.d.). *J2450 Translation Quality Metric*. Acedido em 20 de setembro de 2021, em: <https://www.translationdirectory.com/article581.htm>
- Trados. (n.d.). *What is Translation Alignment?* Acedido em 18 de outubro de 2021, em: <https://www.trados.com/solutions/translation-alignment/>
- Tripathi, S., & Sarkhel, J. K. (2010). Approaches to machine translation. *Annals of Library and Information Studies*, 57, pp. 388-393. Acedido em 18 de julho de 2021, em: <http://nopr.niscair.res.in/bitstream/123456789/11057/4/ALIS%2057%284%29%20388-393.pdf>
- Vardaro, J., Schaeffer, M., & Hansen-Schirra, S. (2019). Translation quality and error recognition in professional neural machine translation post-editing. *Informatics 2019*, 6(3). doi:https://doi.org/10.3390/informatics6030041
- Vela, M., Pal, S., Zampieri, M., Naskar, S. K., & Genabith, J. v. (2019). Improving CAT tools in the translation workflow: New approaches and evaluation. *Proceedings of Machine Translation Summit XVII Volume 2: Translator, Project and User Tracks*, 8-15. Acedido em 30 de março de 2021, em: <https://www.aclweb.org/anthology/W19-6702.pdf>
- Xu, J., Crego, J., & Senellart, J. (2020). Boosting neural machine translation with similar translations. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 1580-1590). França: Association for Computational Linguistics. Acedido em 22 de julho de 2021, em: <https://aclanthology.org/volumes/2020.acl-main/>
- Zaretskaya, A. (2015). The use of machine translation among professional translators. *EXPERT Scientific and Technological Workshop*. Malaga, Espanha. Acedido em 25 de agosto de 2021, em: https://www.researchgate.net/publication/283667234_The_Use_of_Machine_Translation_among_Professional_Translators/citations
- Zaretskaya, A., Pastor, G. C., & Seghiri, M. (2015). Integration of machine translation in CAT tools: State of the art, evaluation and user attitudes. *SKASE Journal of Translation and Interpretation*,

8(1), 76-88. Acedido em 4 de abril 2021, em:
[https://www.researchgate.net/publication/283667119_Integration_of_Machine_Translation_in
CAT_Tools_State_of_the_Art_Evaluation_and_User_Attitudes](https://www.researchgate.net/publication/283667119_Integration_of_Machine_Translation_in_CAT_Tools_State_of_the_Art_Evaluation_and_User_Attitudes)

Zhou, X., & Ma, X. (2021). Research on computer-aided translation. *Scholars International Journal of Linguistics and Literature*, 7(4), pp. 213-215. doi:10.36348/sijll.2021.v04i07.003

ANEXOS

Anexo A – Padrão de avaliação de qualidade J2450 (presente no formulário de avaliação da RWS)

QRF-121 J2450 Errors & Weighting Definitions



This tab contains the detailed information & guidelines on error categories and weightings assigned to specific errors according to J2450 Quality Evaluation Standard. Select category from the list below to see definitions with examples.

CATEGORIES – DEFINITIONS, EXAMPLES & WEIGHTINGS

- A. Wrong Term Error (WT)
- B. Wrong Meaning (WM)
- C. Omission Error (OM)
- D. Structural Error (SE)
- E. Misspelling Error (SP)
- F. Punctuation Error (PE)
- G. Style Error (ST)
- H. Miscellaneous Error (ME)
- I. Legacy Data Error
- J. Source Text Error
- K. Formatting / non-linguistic

SEVERITIES - DEFINITIONS

- Serious Error
- Critical Error
- Repeated Error

For more guidance on error categorization, please go [Meta Rules](#).

[Frequently Asked Questions](#)

A. Wrong Term (WT):

This category cannot be defined until the concept of 'term' has been defined. We define a term as any:

- Single word;
- Multi-word phrase used as a single term;
- Abbreviation;
- Acronym;
- Number or numeral;
- Proper name, including trade names, brand names, registered trademarks, place names;
- Personal names.

Thus, 'automotive' is a single word; 'powertrain control module' is a multi-word term; 'Hz' is an abbreviation; 'ABS' is an acronym; '3.14' is a numeral; 'three' is a number, and 'Society of Automotive Engineers' is a name.

All are terms, as are function words such as prepositions, conjunctions, determiners and articles.

Definition

Given the definition of term, we define a 'wrong term' to be any target language term that:

- a. violates a client term glossary;
- b. is in clear conflict with de facto standard translation(s) of the source language term in the subject domain;
- c. is inconsistent with other translations of the source language term in the same document or type of document unless the context for the source language term justifies the use of a different target language term, for example due to ambiguity of the source language term.

Examples

1. An example of (b) would be 'Parking Brake' translated as "Parkbremse" (aircraft term) in German rather than "Handbremse" (automotive domain).

In Dutch, the equivalent error would be the use of 'parkeerrem' (aircraft) rather than 'handrem' (automotive).

2. Clause (c) is important because it captures the error of inconsistency of terminology, which is a frequent error.

An example of (c) would be a translation of English 'accelerator' into French that sometimes uses 'accélérateur' and sometimes 'pédale des gaz'.

In German, an inconsistency might be to translate 'odometer' sometimes as "Odometer" or sometimes as "Kilometerzähler".

Weighting

Critical weight: 100 error points

Serious weight: 5 error points

Minor weight: 2 error points

Preferential improvement: 0 error points

[Back >](#)

B. Wrong Meaning (WM):

Definition

A term or expression that denotes:

- a. a concept in the target language that is clearly and significantly different from the concept denoted by the source language term, resulting in an incorrect meaning in the target text.
- b. contradicts the meaning of the original.

Examples

1. Taking English as the target language, an example of (a) would be 'Install the new hub nut in order to drive the shaft' instead of 'Install the new hub nut to the drive shaft.'

2. An example of (b) would be 'Do not engage the brake' instead of 'Engage the brake.'

Weighting

Critical weight: 100 error points

Serious weight: 5 error points

Minor weight: 2 error points

Preferential improvement: 0 error points

[Back >](#)

C. Omission (OM):

Definition

An error of omission has occurred if:

- a. a continuous block of text or smaller units of text like individual words and phrases in the source language have no counterpart in the target language text and, as a result, the semantics of the source text is absent in the translation; in other words, source text that should appear as translated in the target text has not been translated;
- b. a graphic that contains source language text has been deleted from the target language deliverable.

Examples

By a 'continuous block of text' is meant anything from a word to a phrase to a sentence.

The corresponding missing text in the target language may or may not need to be continuous.

1. For example, a verb in English may translate to a so-called separable verb in German (where the verb is supposed to appear first, followed later in the sentence by a preposition) and that preposition is then missing in the German. For example, the German 'ablesen' is such a verb that is a translation of the English 'to read off (instruments)', where the verb 'lesen' would normally appear, then be followed by the direct object, and finally by the preposition 'ab'.

In Spanish, it would be an omission to translate "Remove the wheelnuts and the wheel" as "Desmonte la rueda" (Remove the wheel).
2. Part (b) of the definition implies that if the missing graphic does NOT contain any source language text, then there is NO occurrence of an error with respect to SAE J2450.

This would still be, of course, an error and a problem that needs to be rectified by the translation supplier.

But the point is that SAE J2450 is limited to linguistic errors, and does not therefore encompass all potential problems with a translation deliverable.

Note that *omission* does not imply that the source and target language words must be in a 1:1 correspondence.

Weighting

<i>Critical weight:</i>	<i>100 error points</i>
<i>Serious weight:</i>	<i>4 error points</i>
<i>Minor weight:</i>	<i>2 error points</i>
<i>Preferential improvement:</i>	<i>0 error points</i>

[Back >](#)

D. Structural Error (SE):**Definition**

A structural error comprises the following instances of syntactic errors, incorrect word structure, or agreement errors:

- a. a source term is assigned the wrong part of speech in its target language counterpart;
- b. the target text contains an incorrect phrase structure, e.g., a relative clause when a verb phrase is needed;
- c. the target language words are correct, but in the wrong linear order according to the syntactic rules of the target language;
- d. an otherwise correct target language word (or term) is expressed in an incorrect morphological form, e.g. case, gender, number, tense, prefix, suffix, infix, or any other inflection;
- e. two or more target language words disagree in any form of inflection as would be required by the grammatical rules of that language, e.g. feminine noun + adjective.

Examples

1. To show an example of a clause (a) error, the English sentence 'The throttle valve connects to the accelerator pedal' could be mistranslated into French as 'Le papillon des gaz connecte à la pédale d'accélérateur.' In this example, the French verb 'connecter' is transitive and should take a direct object, not a prepositional phrase complement.

2. Clause (c) could occur, for example, in Romance languages if an adjective that was supposed to follow its noun were instead placed in front of the noun and this placement was not driven by reasons of style or emphasis. For example, the English "Check gauge lamp" could be mistranslated into Portuguese as "Luz dos indicadores de verificação" instead of "Luz de verificação dos indicadores".

An English to German example of this error would be "This car was built in Germany" translated into German as "Dieses Auto wurde

3. Note that clause (d) implies that an article of the wrong gender/number/case would be tagged in this category, and not as a Wrong Term error.

An example of (d) could be the use of the wrong Dutch article 'De (rather than: het) gekozen voertuig wordt niet ondersteund'.

An example would be English 'The vehicle selected is not supported' translated to 'Gewählte Fahrzeug wird nicht unterstützt' instead of the correct German 'Gewähltes Fahrzeug wird nicht unterstützt'.

4. An example of (e) would be German 'Der Motor läuft' translated to English 'The motors is running' or 'The motor are running' rather than 'The motor is running.'

In Spanish, "fog lights" should be translated as "luces antiniebla" rather than "luces antinieblas".

Another example of (e) would be the English 'the following steps' translated into French as 'les étapes suivants' rather than 'les étapes suivantes.'

The intent of part (e) of this rule is to assure that even though more than one word may be wrong, there is only one occurrence of the error, rather than one for each word with the incorrect inflection. This is to recognize that the error is a single error in the mind of the translator and that this error simply manifests itself in several words.

Weighting

<i>Critical weight:</i>	<i>100 error points</i>
<i>Serious weight:</i>	<i>4 error points</i>
<i>Minor weight:</i>	<i>2 error points</i>
<i>Preferential improvement:</i>	<i>0 error points</i>

[Back >](#)

E. Misspelling (SP):**Definition**

A misspelling has occurred if a target language term:

- a. violates the spelling as stated in a client glossary;
- b. violates the accepted norms for spelling in the target language;
- c. is written in an incorrect or inappropriate writing system for the target language.

Clause (a) of course assumes that the client glossary is correct.

Examples

1. An example in German of a misspelling could be in the compound "Stromkreis Kraftstoffreigabe Bereich" for "Fuel Enable Circuit Range" where the sequence of 3 f's should have been reduced to just 2 consecutive f's. The English "Screw on" could be misspelled in Portuguese as "Atarraxar" instead of "Atarraxar".
2. Clause (c) is intended to flag as an error words in target languages such as Japanese that use more than one writing system, and the word in question was written with the wrong writing system, even though it may have been spelled correctly in that (wrong) system. For example in Japanese if a foreign term, normally written in Katakana, were instead written in Hiragana, albeit with the correct syllabic symbols in Hiragana for that word's pronunciation.

Weighting

Critical weight:	100 error points
Serious weight:	3 error points
Minor weight:	1 error point
Preferential Improvement:	0 error points

[Back >](#)

F. Punctuation Error (PE):**Definition**

The target language text contains an error according to the punctuation rules for that language.

Examples

1. For example, in German "Am 2 Zylinder" for "On the 2nd cylinder" there should be a period following the "2" to indicate an ordinal number.
2. A serious punctuation error could involve a missing decimal point or comma according to the target language, e.g., "116" instead of "11.6" or "11,6".
3. In Dutch and other languages, there could be a missing concatenation mark for implied compounds, such as the translation of "particle separators and filters" as "partikelafscheiders en filters" instead of "partikelafscheiders en -filters".

Weighting

Critical weight:	100 error points
Serious weight:	2 error points
Minor weight:	1 error point
Preferential Improvement:	0 error points

[Back >](#)

G. Style Error (ST):**Definition**

A style error is applicable if:

- a. the register or tone has not been adapted appropriately for the target audience;
- b. specified style guidelines have not been followed as required.

Examples

1. An example of a style error according to clause (a) would be if it is appropriate to adopt formal language to address the target audience, such as "Vous" in French or "Sie" in German, but the translation uses the informal level of address ("Tu" or "Du").
2. Another example is use of slang when it is inappropriate for the target audience, or conversely not using slang when it is genuinely appropriate for the context as may be the case in video games.

Weighting

Critical weight:	100 error points
Serious weight:	2 error points
Minor weight:	1 error point
Preferential Improvement:	0 error points

[Back >](#)

H. Miscellaneous Error (ME):

Definition

Any linguistic error related to the target language text which is not clearly attributable to the other categories listed previously should be classified as a miscellaneous error.

This rule recognizes that not all errors fall neatly into the other seven error categories, and that some evaluators may proofread a translation and recognize it as clearly wrong, yet not be able to say precisely what it is that is wrong.

Examples

Examples of miscellaneous errors include (but are not limited to):

1. A literal translation of an idiom, e.g., the English idiom "to push the envelope" meaning "to perform beyond existing boundaries."
2. A superfluous addition or "extra" block of text that has no relevance to the source text.
3. An extra space, introduced by the translator, that is redundant in the target language.
4. Failure to adapt country standards (time and date formats, units of measurement, currency, number formats, sorting order etc.).
5. Inadequate localization of scenarios, examples and images appearing in the source text that are inappropriate in their source state for the target audience.

An example of an error not easily categorized in any preceding category could be the French "l'usage" ("the usage") appearing as "l usage". Here a space appears after the apostrophe, which could have occurred as a typo, for instance.

Superfluous additional text could occur where the translator includes two synonyms in a text, perhaps intending to check with the client for the preferred term, only to forget to do so, leaving both terms in the delivered translation. Thus, in a translation to French the English term "relay box" may show up as "boîte de relais / boîte à relais".

Generally the process of translation is performed as part of the process of localization, which makes a product culturally appropriate for the target audience as well as linguistically correct.

With this in mind, a translator may be perfectly justified in adding text to a translation in order to explain some cultural reference that may not be familiar to the target audience of the translation.

Such additions should not be regarded as superfluous, and no error should be marked.

Weighting

Critical weight:	100 error points
Serious weight:	3 error points
Minor weight:	1 error point
Preferential improvement:	0 error points

[Back >](#)