

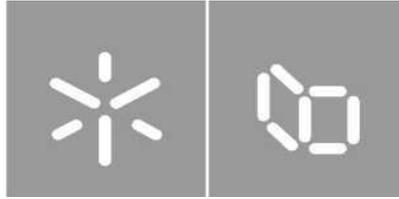
**Universidade do Minho**

Escola de Letras, Artes e Ciências Humanas

Mariana da Silva Miranda

**A Tradução e a Linguística de corpus: uma  
experiência de estágio na AP|Portugal**





**Universidade do Minho**

Escola de Letras, Artes e Ciências Humanas

Mariana da Silva Miranda

**A Tradução e a Linguística de corpus: uma  
experiência de estágio na AP|Portugal**

Relatório de Estágio

Mestrado em Tradução e Comunicação Multilíngue

Trabalho efetuado sob a orientação do

**Professor Doutor Fernando Gonçalves  
Ferreira Alves**

abril de 2024

## **DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS**

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

### ***Licença concedida aos utilizadores deste trabalho***



**Atribuição  
CC BY**

<https://creativecommons.org/licenses/by/4.0/>

## **Agradecimentos**

Quero começar por agradecer aos meus pais, os grandes responsáveis pelo meu percurso académico. Foram eles que desde sempre me incentivaram a estudar, sobretudo a estudar o que me fazia feliz. Sem eles, não chegaria ao lugar onde estou.

Ao meu Orientador, Professor Doutor Fernando Ferreira Alves, por todo o apoio, acompanhamento e conselhos ao longo do mestrado e durante este projeto. Demonstrou uma dedicação louvável, dentro e fora da sala de aula, permitindo que este trabalho alcançasse a sua melhor versão.

À AP|PORTUGAL, pela oportunidade de realização de um estágio curricular. Um obrigado em especial ao Eugénio Rodrigues, Ricardo Silva, Teresa Peixoto e todos os colegas estagiários, por todo o apoio dado ao longo do estágio e por terem proporcionado um ambiente que me permitisse crescer enquanto profissional e pessoa.

Ao meu avô.

À Renata e à Ângela. Foram dois dos meus pilares mais importantes nesta jornada. Por todas as horas passadas a me ouvirem, pelos conselhos infundáveis, pela cumplicidade, por serem as melhores amigas que alguém podia pedir.

À Helena Mendes, a principal responsável por ter desenvolvido um amor pelas línguas e pela literatura desde terna idade.

Ao Miguel, que me acompanhou desde a licenciatura até ao fim do mestrado.

Ao André, o meu afilhado, por todo o apoio e pela leveza que trouxe à minha vida.

Aos meus amigos que me acompanham desde o dia um da licenciatura.

Aos meus colegas da Tuna.

*“Never be ashamed of trying.*

*Effortlessness is a myth.”*

*- Taylor Swift*

## **DECLARAÇÃO DE INTEGRIDADE**

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

## **Resumo**

O presente relatório visa apresentar e analisar os projetos realizados e a experiência de estágio curricular com a duração de seis meses, realizado na empresa de serviços linguísticos AP | Portugal - Tech Language Solutions e inserido no Mestrado em Tradução e Comunicação Multilíngue da Universidade do Minho.

O tema central deste relatório, tendo como base as tipologias tradutórias trabalhadas, foca-se na análise multinível dos textos trabalhados enquanto tradutora, e a respetiva interligação que a tradução tem com a linguística de corpus, com especial enfoque na deteção de padrões linguísticos. O relatório focar-se-á, inicialmente, na contextualização teórica da Tradução, da Linguística de corpus e da forma como estas disciplinas se interligam e entrelaçam. Para além disso, conta com a descrição da instituição de acolhimento, explicitando a extensão e variedade de projetos e serviços prestados pela mesma, assim como uma apresentação das tarefas desempenhadas ao longo do estágio. Adicionalmente, serão analisados exemplos práticos dos trabalhos realizados, seguindo-se a análise multinível e apresentação dos corpora, terminando com o registo de conclusões acerca dos padrões e níveis de ocorrências da linguagem técnica em duas línguas distintas. Por fim, serão apresentadas as considerações finais relativamente à experiência obtida através do estágio curricular, refletindo sobre o conhecimento adquirido e as competências desenvolvidas.

## **Palavras-chave**

Linguística de corpus, Tradução, Tradução técnica

## **Abstract**

This report aims to present and analyse the projects conducted and the experience of the six-month curricular internship at the language services company AP|Portugal - Tech Language Solutions, within the scope of the Master's in Translation and Multilingual Communication at the University of Minho.

The scope of this report, having as a base the translation typologies worked on, focuses on the multi-level analysis of the texts worked on as a translator, and the respective interconnection that translation has with corpus linguistics, with a special focus on the detection of linguistic patterns.

The report will initially focus on the theoretical contextualisation of translation, corpus linguistics and how these subjects interconnect and help each other. It also includes a description of the host institution, explaining the extent and variety of projects and services it provides, as well as a presentation of the tasks fulfilled during the internship. In addition, some translation problems will be analysed, followed by presentation and the multilevel analysis of our corpora, and concluding about the patterns and levels of occurrences present on technical language. Finally, conclusions will be drawn regarding the experience gained during the internship, while reflecting on the knowledge acquired and the skills developed.

## **Keywords**

Corpus linguistics, Translation, Technical translation

## Índice

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS .....	ii
Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vi
Lista de abreviaturas .....	ix
Índice de Figuras.....	ix
Índice de Gráficos.....	x
Índice de Tabelas .....	x
Introdução.....	1
1. A Tradução e a Linguística de corpus .....	2
1.1 A Tradução .....	2
1.1.1 Tradução geral e tradução especializada .....	2
1.1.2 Tradução técnica .....	3
1.1.3 O papel da comunicação técnica para a tradução.....	5
1.2 A Linguística de corpus.....	8
1.2.1 Ferramentas de processamento de corpus.....	9
1.2.2 Anotação de corpora .....	9
1.2.3 Estudos de Tradução baseados em corpus.....	10
1.3 A Tradução e a Linguística de corpus.....	12
1.3.1 Características genéricas e específicas da comunicação e documentação técnica.....	12
2 A Entidade de acolhimento do estágio .....	15
2.1 A AP   Portugal.....	15
2.2 Constituição da empresa .....	17
2.3 Caracterização do estágio .....	17
3 Tarefas Realizadas .....	20
3.1 Diário de Bordo .....	20
3.2 Trabalho realizado.....	24
3.3 Línguas e domínios de trabalho .....	27
3.4 Ferramentas CAT utilizadas .....	29
3.5 Análise de questões associadas a projetos.....	33
3.5.1 Tradução literal.....	34

3.5.2	Questões terminológicas .....	36
3.5.3	Questões morfológicas .....	37
3.5.4	Questões pragmáticas.....	38
4	Estudo do corpus comparável.....	40
4.1	Workflow de criação do corpus .....	40
4.2	Tipologia textual   Macroestrutura .....	40
4.3	Tamanho do corpus .....	42
4.4	Escolha da ferramenta Sketch Engine.....	44
4.5	Análise linguística dos corpora.....	44
4.5.1	Análise a nível terminológico .....	45
4.5.2	Análise a nível morfossintático.....	51
4.5.3	Análise a nível semântico-enunciativo .....	53
4.6	Conclusões relativas aos corpora .....	60
5	Considerações finais .....	62
6	Bibliografia.....	64
7	Anexos.....	68
7.1	Anexo I: Organograma da AP   Portugal .....	68
7.2	Anexo II: Exemplo do relatório mensal.....	69
7.3	Anexo III: Tempo gasto, em horas, por tarefa e por mês.....	70
7.4	Anexo IV: Representação visual do workflow de criação dos corpora .....	71
7.5	Anexo V: Lista de POS tags para pesquisa nos corpora .....	71

## Lista de abreviaturas

CQL – Corpus Query Language
DeL – Densidade Lexical
DiL – Diversidade Lexical
DTP – Desktop publishing (em português, paginação eletrónica)
EN – inglês
LC – Língua de chegada
LCorpus – Linguística de corpus
LP – Língua de partida
PEMT – Post-Editing Machine Translation (em português, pós-edição de tradução automática)
POS – Part-of-speech
PSL – Prestador de Serviços Linguísticos
PT – português
TC – Texto de chegada
TP – Texto de partida

## Índice de Figuras

Figura 1 - Página web da AP Portugal.....	15
Figura 2 - Primeiras entradas de projetos de Tradução .....	21
Figura 3 - Primeiras entradas de projetos PEMT .....	21
Figura 4 - Editor do Smartcat .....	31
Figura 5 - Editor do memoQweb.....	32
Figura 6 - Editor do Smartling .....	32
Figura 7 - Tamanho do corpus PT .....	42
Figura 8 - Tamanho do corpus EN.....	42
Figura 9 - Tamanho real do corpus PT .....	42
Figura 10 - Tamanho real do corpus EN .....	43
Figura 11 - Monopalavras EN .....	45
Figura 12 - Monopalavras PT .....	46
Figura 13 - Multipalavras EN .....	46
Figura 14 - Multipalavras PT .....	47

## **Índice de Gráficos**

Gráfico 1 - Tempo gasto, em horas, por tarefa e por mês .....	24
Gráfico 2 - Diversidade de projetos desenvolvidos .....	25
Gráfico 3 - Domínios dos projetos de tradução .....	28
Gráfico 4 - Domínios dos projetos PEMT.....	29
Gráfico 5 - Ferramentas CAT utilizadas.....	30
Gráfico 6 - Ocorrências de classes de palavras no corpus EN .....	48
Gráfico 7 - Ocorrências de classes de palavras no corpus PT .....	48
Gráfico 8 - Colocações presentes no corpus PT .....	50
Gráfico 9 - Colocações presentes no corpus EN.....	50
Gráfico 10 - Hipotaxe vs. Parataxe no corpus PT.....	51
Gráfico 11 - Formas frásicas presentes no corpus EN.....	52
Gráfico 12 - Formas frásicas presentes no corpus PT .....	53
Gráfico 13 - Ocorrências de tipos verbais no corpus EN.....	54
Gráfico 14 - Ocorrências de modos verbais no corpus PT .....	55
Gráfico 15 - Tempos verbais presentes no corpus EN .....	55
Gráfico 16 - Ocorrências do modo Indicativo no corpus PT .....	56
Gráfico 17 - Verbos modais presentes no corpus EN .....	57
Gráfico 18 - Verbos modais presentes no corpus PT .....	58
Gráfico 19 - Verbos principais com valor modal no corpus PT .....	59
Gráfico 20 - Adjetivos com valor modal presentes no corpus PT.....	60

## **Índice de Tabelas**

Tabela 1 - Volume de trabalho realizado.....	26
Tabela 2 - Erros de tradução literal com tradução automática.....	34
Tabela 3 - Erros de tradução literal.....	35
Tabela 4 - Erros de terminologia .....	36
Tabela 5 - Erros morfológicos.....	37
Tabela 6 - Erros relativos a questões pragmáticas .....	38
Tabela 7 - Semelhanças estruturais entre o mesmo tipo de texto.....	41
Tabela 8 - Densidade lexical dos corpora.....	43
Tabela 9 - Diversidade lexical dos corpora .....	44
Tabela 10 - Colocações procuradas nos corpora .....	49

## **Introdução**

O presente relatório tem como objetivo a apresentação e descrição do estágio curricular realizado na AP|Portugal. Este estágio surge como conclusão do Mestrado em Tradução e Comunicação Multilíngue e permitiu aplicar a teoria e prática provenientes do contexto de sala de aula à realidade de uma empresa de serviços linguísticos, aprendendo, assim, a lidar com as dinâmicas de uma empresa, desde a aceitação de um projeto de tradução, a realização de paginação eletrónica para orçamentação e/ou tradução, a divisão de tarefas, a importância de prazos, a utilização de programas específicos para a tradução e edição, bem como o bom funcionamento e bom ambiente em contexto de equipa. Este relatório será realizado sob a ótica da tradução técnica e da linguística de corpus e pretende analisar como estas disciplinas se entrelaçam.

A informação encontra-se estruturada em capítulos. O primeiro é constituído pela presente introdução, que tem como objetivo enquadrar o relatório, definir os seus objetivos e apresentar a sua estrutura. De seguida, apresentamos o enquadramento teórico através da lente da tradução, especificando as diferenças entre a tradução técnica e a tradução especializada, e da linguística de corpus, sobretudo vendo a forma como esta se relaciona com a tradução. Neste enquadramento teórico, apresentamos ainda as características da comunicação técnica e a importância que a mesma tem para a tradução e para o seu público-alvo.

Posteriormente, além da apresentação da empresa, iremos descrever o modo como esta é constituída e as funções e tarefas que foram realizadas ao longo do estágio, particularmente tradução, pré-edição e pós-edição.

Por fim, todas as tarefas realizadas em contexto de estágio serão analisadas, no que diz respeito às tipologias textuais e pares linguísticos trabalhados, o número de palavras traduzidas, assim como as ferramentas utilizadas e exemplos de dificuldades de tradução. Partimos ainda para a análise multinível dos textos trabalhados, fazendo uma reflexão entre os dados obtidos e a pergunta de investigação a ser colocada.

O relatório encerra com as Considerações Finais, onde expomos uma síntese das principais ideias desenvolvidas neste relatório, a pertinência do estágio curricular enquanto complemento fulcral na formação de tradutores e a forma como os corpora podem ser uma ferramenta útil para os tradutores.

## **1. A Tradução e a Linguística de corpus**

Os Estudos da Tradução e a Linguística de corpus são duas disciplinas interligadas. Nos Estudos da Tradução, os corpora fornecem dados empíricos para pesquisa em tradução e auxiliam no desenvolvimento de ferramentas e recursos de tradução. Os corpora<sup>1</sup> podem ser utilizados para investigar o fenômeno tradutório – a equivalência tradutória, a influência do texto de partida (TP) no texto de chegada (TC), as estratégias de tradução e o impacto da tradução na variação linguística. Ao comparar os textos originais e as suas traduções num corpus, os investigadores e teóricos podem examinar escolhas tradutórias, identificar padrões e avaliar a qualidade das traduções. O que pretendemos explorar nesta secção é a relação existente entre a tradução, nomeadamente a tradução técnica, e a linguística de corpus, e a forma como estas disciplinas se autoajudam.

### **1.1 A Tradução**

A noção e definição de tradução têm vindo a sofrer uma constante evolução, desde sempre, mas especialmente nos últimos tempos. O termo tradução pode ainda ter vários significados, podendo referir-se ao campo de estudo, ao produto (o texto que foi traduzido) ou ao processo (o ato da produção da tradução) (Munday, 2016, p. 4). Quando pensamos em tradução, a primeira definição que surge é a dos pioneiros Vinay e Darbelnet (1958), que definem a tradução como a passagem de um texto de uma língua A para uma língua B para expressar a mesma realidade. A esta definição podemos juntar a de Nida e Taber (1982) que definem o processo de tradução em termos de reprodução, na língua de chegada, do equivalente natural mais próximo da mensagem na língua de partida a nível de significado e estilo. Já House (2014, p. 1) define a tradução como um processo cognitivo que ocorre na cabeça do tradutor, e ao mesmo tempo, uma prática social, interlinguística e intercultural e, desta forma, acaba por ser fortemente influenciada por uma variedade de fatores e condições extralinguísticos. Devido à pluralidade de definições de tradução, houve a necessidade de dividir a tradução em domínios. Assim, Hurtado Albir (2011) divide os tipos de tradução em tradução de textos especializados e tradução de textos não especializados. Já Gouadec (2007) divide a tradução em duas categorias, tradução geral e tradução especializada, divisão que será adotada neste trabalho.

#### **1.1.1 Tradução geral e tradução especializada**

Daniel Gouadec (2007) separa o conceito de tradução em duas categorias ou domínios: tradução geral e tradução especializada. A tradução geral refere-se à tradução de documentos e

---

<sup>1</sup> Um corpus (no plural, corpora) representa um conjunto de dados linguísticos coletado para ser utilizado como objeto de pesquisa linguística.

materiais que não se referem a nenhum domínio ou área específicos e que não implicam nenhum processo de tradução específico (Gouadec, 2007, p. 29). Em oposição, a tradução especializada pode ser definida como a tradução de materiais que se referem a uma área ou domínio específico, que se destinam a um público-alvo específico através da divulgação em canais específicos utilizados pelos especialistas dos mesmos, exigindo a utilização de procedimentos, ferramentas e protocolos especiais (p.30). Gouadec divide a tradução especializada em diferentes subdomínios, nos quais se inclui a tradução técnica, tema principal do nosso estudo. Também Hurtado Albir (2011) afirma que a tradução técnica pertence à tradução de textos especializados, distinguindo, assim, as noções de tradução de textos especializados de tradução especializada.

### **1.1.2 Tradução técnica**

A divisão em subdomínios da tradução especializada feita por Gouadec (2007) vem desmistificar a ideia da tradução especializada e da tradução técnica como sendo consideradas equivalentes. A tradução técnica é uma especialização por si só, incluindo a tradução de qualquer material pertencente a uma área específica de conhecimento, campo técnico ou tecnologia, desde que os materiais exijam conhecimento específico da área envolvida (Gouadec, 2007, p. 32). Também Byrne (2006) realça que é necessária esta distinção entre tradução especializada e técnica, destacando seis concepções erradas sobre a tradução técnica:

- A tradução técnica inclui economia, direito, negócios, entre outros domínios;
- A tradução técnica é uma questão de terminologia;
- O estilo não é importante na tradução técnica;
- A tradução técnica não é criativa, é apenas um processo de transferência reprodutiva;
- É necessário ser um perito num domínio altamente especializado e;
- A tradução técnica tem a ver com a transmissão de informação especializada.

O autor realça que não é por uma área ou domínio possuir terminologia única, específica ou especializada, que se trate de um texto técnico, salientando que a tradução técnica diz respeito a textos que se baseiam em conhecimento aplicado das ciências naturais (Byrne, 2006, p. 3). Também Zethsen (1999) havia abordado três dogmas presentes na tradução técnica: (1) o propósito da tradução técnica é transmitir informação factual; (2) o maior problema da tradução técnica é a terminologia e (3) não é necessária a utilização de técnicas de tradução desde que o tradutor esteja familiarizado com a terminologia e sintaxe do texto técnico. A autora revisita estes dogmas depois de definir tradução técnica e o papel do redator técnico afirmando que no caso (1) comunicar informação não é o único propósito da tradução técnica e que o foco constante na

função informativa acaba por enfraquecer as expectativas e consciência de outras funções textuais. Já para o dogma (2) realça que o verdadeiro desafio do tradutor técnico não é a terminologia específica da área, mas sim a combinação da terminologia específica com as características expressivas que o texto deve suportar. Por fim, em relação ao (3), o conhecimento de terminologia e convenções textuais não é suficiente para que um tradutor consiga realizar um bom trabalho.

Schubert (2010, p. 352) afirma que «technical translation is concerned with documents in languages for special purposes (LSP). LSP texts, at least in European texts, are generally characterized by specific features at the lexical, morphological, word-formational, syntactic and text-linguistic levels». Complementarmente, já Byrne (2006, p. 11) declara que o propósito da tradução técnica é:

present new technical information to a new audience, not to reproduce the source text, per se, or reflect its style or language. Technical translation is a communicative service provided in response to a very definite demand for technical information which is easily accessible (in terms of comprehensibility, clarity and speed of delivery).

De forma sintetizada, a tradução técnica tem como objetivo principal transmitir informação de uma forma clara e concisa, para que o leitor possa usá-la adequada e eficientemente. Para que o recetor da tradução a possa interpretar e consiga aplicar o resultado da tradução de forma simples, é necessário que o tradutor tenha em atenção a função que o texto deve possuir. Posto isto, a teoria que apoia esta afirmação é, precisamente, a *Skopostheorie*, desenvolvida por Hans Vermeer em finais de 1970, segundo a qual a regra fundamental na tomada de decisão de um tradutor é o *skopos*, ou seja, o propósito ou a função do texto, pois este é o princípio que determina o modo como o mesmo será traduzido. Assumindo que a função do autor é a comunicação, Reiss (em Reiss e Vermeer, 2013) categoriza os textos em três tipos distintos, de acordo com as funções da linguagem: informativo, expressivo e operativo (Reiss & Vermeer, 2013, p. 137). O tipo informativo corresponde ao texto produzido com a intenção de transmitir conhecimento, notícias, pontos de vista, entre outros (Reiss & Vermeer, 2013, p. 182), sendo as brochuras turísticas e os manuais de instruções exemplos deste tipo de texto. O tipo expressivo dá prioridade à equivalência a nível estético e à forma do conteúdo (Reiss & Vermeer, 2013, p. 142), do qual são exemplo as peças de teatro. Por último, temos o tipo operativo, que diz respeito à transmissão de sentimentos e emoções, com o intuito de persuadir e/ou convencer o recetor, fazendo utilização da função apelativa da linguagem (Reiss & Vermeer, 2013, p. 182).

Associado aos tipos de texto descritos acima, temos as funções da linguagem descritas por Juliane House. A linguista alemã destaca as funções mais influencias e foca-se em três funções principais: a referencial, a expressiva e a fática. A função referencial centra-se na transmissão de conhecimento, a função expressiva utiliza a linguagem para exprimir significados afetivos e a função fática tem como objetivo principal estabelecer, prolongar ou interromper a comunicação.

Ou seja, tendo em conta os tipos de texto propostos por Reiss e as funções da linguagem utilizadas no modelo de House podemos concluir que o texto técnico é um texto informativo com uma função referencial da linguagem e, por isso, a elaboração e tradução deste tipo de textos deve ser realizada tendo em conta o público-alvo e assumindo que a mensagem é passada de forma clara, permitindo a utilização da informação dessas produções de forma eficaz pelo recetor, o que realça a importância da comunicação técnica, que será abordada no tópico seguinte.

### **1.1.3 O papel da comunicação técnica para a tradução**

Para entendermos a comunicação técnica e a importância que esta representa para a tradução, primeiro temos de perceber como a mesma é processada e qual a sua origem, para que saibamos melhor como produzir a tradução de conteúdo técnico. Sempre que lemos um texto que contenha informação técnica, seja a explicar o funcionamento de ou como fazer algo e que ajude o recetor a compreender conceitos tecnológicos, estamos perante um produto da comunicação técnica. A Associação Portuguesa de Comunicação Técnica (APCOMTEC) realça tipos de textos onde a comunicação técnica e a linguagem técnica estão presentes: páginas de ajuda, guias do utilizador, manuais de instruções, entre outros. Assim, a comunicação técnica tem como objetivo informar o seu público-alvo sobre um determinado tópico, ajudar a executar um procedimento de forma clara e segura e/ou evitar que o recetor corra riscos de perigo ou utilize uma ferramenta ou dispositivo de forma incorreta (Byrne, 2014, p. 26). De forma a apoiar a definição de Byrne, a APCOMTEC, no seu artigo *Manual de Comunicação Técnica* explicita o seguinte:

A Comunicação Técnica está ligada à criação e manutenção de conteúdos técnicos usáveis e multimodais, relacionados com produtos, software ou serviços, que descrevem a sua instalação, configuração, utilização, resolução de problemas, etc., cumprindo com requisitos legais e destinados a um público específico, podendo ser incorporados no produto, software ou serviço, fornecidos na embalagem juntamente com o produto e/ou publicidade online.

Tendo em conta as seguintes posições, podemos afirmar que a comunicação técnica tem como objetivo principal informar o seu público-alvo. Byrne (2014) constata ainda que estes textos são redigidos por dois tipos de pessoas: profissionais técnicos e comunicadores técnicos. Os profissionais técnicos são os «peritos na matéria que desenvolvem os dados ou conhecimentos comunicados nos textos» e os comunicadores técnicos são definidos como um «profissional de comunicação a tempo inteiro, cuja função consiste em produzir e comunicar informações técnicas» (Byrne, 2014, p. 26, nossa tradução). Exemplos de profissionais técnicos são os engenheiros, técnicos ou especialistas que escrevem os seus próprios textos, enquanto os comunicadores técnicos podem ter um conhecimento mais limitado sobre o conteúdo que produzem, mas sabem como o expor de maneira que o utilizador comum o saiba utilizar e não tenha dificuldades em interpretar o documento, ou seja, um comunicador técnico sabe qual a melhor forma de comunicar com o público-alvo pretendido. O autor ainda refere que o processo de comunicação técnica é um processo colaborativo e complexo, que envolve redatores técnicos, ilustradores, editores, peritos em conteúdos, designers e tradutores (Byrne, 2006, p. 49).

De acordo com a informação apresentada acima, podemos referir o processo de tradução técnica como o transporte de todo o texto produzido pelo comunicador técnico para as línguas de chegada (LC), como concluído por Coelho (2023). Em paralelo às conclusões de características da tradução técnica, consideremos a abordagem de Schubert, em que o autor propõe o Modelo de Comunicação Especializada (Schubert citado em Schubert, 2010), onde se defende a tradução técnica como uma atividade levada a cabo por um agente em interação com outro agente. Esta atividade distingue-se em duas etapas: uma etapa 'interna' e uma etapa 'externa'. A etapa interna é um processo de tomada de decisão, ou seja, é necessária a atividade mental e cognitiva do tradutor para que este processo seja levado a bom-porto. Esta atividade cognitiva é bastante complexa, envolvendo a compreensão de um texto numa língua A e a produção do mesmo numa língua B. Já a etapa externa concentra-se em tudo o que consegue ser observado por uma entidade externa, todas as ações levadas a cabo pelo agente, incluindo a interação com outros.

Este Modelo pode ainda ser ligado às noções de agente e agência isto é, um agente é mais do que «produtores de texto, mediadores que modificam o texto, como os que produzem resumos, editores, revisores e tradutores, comissários e editoras» (Milton & Bandia citado em Petraru, 2015, p. 637, nossa tradução), ou seja, os agentes desempenham uma função intermediária entre um tradutor e o recetor do texto final, um agente de tradução é qualquer entidade envolvida num processo de inovação e troca cultural (Milton & Bandia, 2009). Ainda Simeoni (1995) define agente

como um 'sujeito', mas socializado. Falar de um agente de tradução sugere que a referência é uma 'voz', uma caneta, ou, atualmente, um computador, e esta referência está inextricavelmente ligada a redes de outros agentes sociais. Já a noção de agência é considerada como algo individual em relação ao tradutor como agente. Pereira (2023) utiliza a definição de agência de Kinnunen & Koskinen (2010) que definem este conceito como a «disposição e capacidade para agir», conceito que como Petraru (2015) menciona é considerado a nível individual em relação ao tradutor. Esta disposição para agir encontra-se associada a fatores internos de um agente, como a sua «consciência, reflexividade e intencionalidade» (Pereira, 2023, p. 7) , que podemos ligar à etapa interna do Modelo de Schubert mencionado acima. Em contraste, a capacidade para agir está relacionada com a etapa externa do Modelo, sendo esta relacionada com fatores externos ao agente, como a estrutura do ambiente em que se encontra a desenvolver o seu trabalho e as restrições com que, nele, se depara (Pereira, 2023, p. 7).

Portanto, podemos concluir que o Modelo de Schubert assenta nas noções de *agent* e *agency*, e, à semelhança dos LSP, é dividido em quatro dimensões: conteúdo técnico, forma linguística, meio técnico e processos de trabalho (Schubert, 2010, p. 350). O conteúdo técnico tem em conta a seleção do conteúdo (o que é dito), a sequenciação (como é organizado) e a estrutura (como encontrar o conteúdo). A forma linguística preocupa-se com as características dos textos em LSP a níveis léxico, morfológico, de formação de palavras, sintático e texto-linguístico. A dimensão do meio técnico contem a tipografia, o layout e o design web do(s) documento(s), assim como o uso de ilustrações. E por fim, os processos de trabalho centram-se na organização das tarefas a desempenhar e as ferramentas a serem utilizadas (Schubert, 2010).

Estabelecendo ligação entre o Modelo de Comunicação Especializada e a tradução de documentos em LSP, ou seja, fazendo uso de todas as dimensões que caracterizam um texto em LSP (níveis léxico, morfológico, formação de palavras, sintático e texto-linguísticos), cruzamo-nos com o estudo de corpora que permite a deteção de padrões na linguagem utilizada. Esta deteção de padrões é analisada sobretudo usando a dimensão da forma linguística dos textos em LSP, fazendo com que seja possível formular e verificar hipóteses sobre a padronização da linguagem da tradução técnica, através da criação e análise de corpora de textos pertencentes a este domínio da tradução, com o auxílio de ferramentas computacionais de análise de corpora, que iremos explorar nos capítulos seguintes.

## **1.2 A Linguística de corpus**

Como o próprio nome indica, a Linguística de corpus (LCorpus) representa uma área de estudo da linguística que se ocupa da coleta e análise de corpus. Para melhor entendermos esta disciplina, importa definir corpus. O Dicionário dos Termos Linguísticos define corpus como:

Conjunto finito de dados linguísticos tomados como ponto de partida da descrição linguística ou como meio de verificação de hipóteses sobre a língua a estudar, e que poderão eventualmente conduzir à elaboração de um modelo explicativo dessa língua. Segundo a investigação pretendida, trata-se de uma coleção de documentos, quer orais (registados e/ou transcritos), quer escritos, quer orais e escritos. As dimensões do corpus e o conjunto de enunciados característicos do fenómeno a estudar variam com o objetivo do investigador. Um corpus é dito exaustivo quando engloba todos os enunciados característicos. É dito seletivo quando só engloba uma parte desses enunciados. (Dicionário de Termos Linguísticos, s.d.)

Este ramo da linguística surge como necessidade dos linguistas se apoiarem em usos reais e textos autênticos para fazer generalizações ou esboçar teorias sobre o funcionamento linguístico. Partington (2011) sublinha que a linguística de corpus é utilizada para melhorar técnicas de tradução, ajudando a entender como é que a linguagem é estruturada e como é usada para comunicar significado, expressar avaliações e influenciar o comportamento e crenças dos seus interlocutores. Devido ao aumento da utilização e ligação da informática com o estudo linguístico, Laviosa (2010, p. 80) define um corpus como uma coleção de textos em forma eletrónica, sendo ainda classificados de acordo com seis critérios: quanto à sua finalidade (corpus de estudo, corpus de referência ou corpus de treino), quanto ao tempo (sincrónico ou diacrónico), quanto à linguagem (geral ou especializado), quanto à língua (monolíngue, bilingue ou multilingue), quanto ao tipo de texto (escrito, oral, misto ou multimodal) e quanto à anotação (anotado ou não anotado).

Posto isto, podemos afirmar que a linguística de corpus contemporânea se caracteriza pela coleta e análise de corpora eletrónicos com o auxílio de ferramentas computacionais.

### **1.2.1 Ferramentas de processamento de corpus**

As ferramentas computacionais de processamento de corpus são utilizadas na reorganização e extração de informações no corpus e para a observação e interpretação de dados, fornecendo, assim, novas perspectivas para a análise linguística e para a tradução.

Segundo Laviosa (2010), existem dois tipos específicos de ferramentas de processamento de corpus: concordanciadores e programas para listar palavras. Existe ainda um outro tipo de ferramenta computacional não mencionada pela autora, os etiquetadores. Os programas de listagem de palavras, ou *word listers*, calculam o número de *tokens* (palavras totais) e *types* (palavras únicas) presentes num corpus. Já os concordanciadores são programas que permitem a procura por palavras específicas num corpus, fornecendo listas exaustivas para as ocorrências da palavra em contexto, ou seja, constrói uma lista de ocorrências de determinada palavra ou frase. Os etiquetadores realizam análises automáticas do corpus e inserem etiquetas de ordem e/ou cariz morfossintático, sintático, semântico, prosódico e discursivo.

Este uso de concordanciadores permite evidenciar e quantificar padrões. Através desta evidência e quantificação de padrões, podemos constatar que a LCorpus faz uso de uma abordagem empirista e tem como central a noção de linguagem enquanto sistema probabilístico (Berber Sardinha, 2004, p. 30).

Desta forma, podemos assumir a linguagem como sendo padronizada, ou seja, existe uma correlação entre os traços linguísticos e os contextos situacionais do uso da linguagem. Esta padronização é evidenciada através de colocações, coligações ou estruturas repetidas, sendo os principais conceitos da padronização as colocações, as coligações e a prosódia semântica. Estes padrões linguísticos podem ser lexicais e léxico-gramaticais, apresentando regularidade e variação sistemática. A padronização é uma «regularidade expressa na recorrência sistemática de unidades coocorrentes de várias ordens (lexical, gramatical, sintática etc.)» (Berber Sardinha, 2004, p. 40). A padronização pode ser evidenciada na anotação de corpora.

### **1.2.2 Anotação de corpora**

A anotação de corpus é a prática de acrescentar informação linguística interpretativa a um corpus. Um tipo comum de anotação é a adição de tags que indicam a classe de palavras a que pertencem as palavras num texto, processo a que chamamos POS tagging (Leech, 2005). POS tagging, ou part-of-speech, é o tipo de anotação que utilizamos no nosso estudo.

Segundo Partington (2011), há duas principais formas de anotação: POS tagging e mark-up. Na anotação em POS, a cada elemento lexical presente no corpus é atribuída uma *tag*, que indica

o seu estatuto gramatical (nome, determinante, quantificador, etc.) no contexto em que aparece. Já mark-up é utilizado para uma variedade quase infinita de itens. Os editores podem querer indicar unidades estruturais de textos ou passagens de transação e interação, ou mesmo mudanças no tópico de discussão. Nos textos falados, podem querer acrescentar informações sobre os oradores, ou podem querer indicar a ocorrência de palavras estrangeiras, calão, nomes pessoais, nomes de lugares, datas ou quase tudo aquilo em que um analista possa estar interessado. (Partington, 2011, p. 37).

E é ao analisar padrões linguísticos e a construir uma representatividade dos elementos estudados, que nos deparamos com a introdução dos estudos de corpora nos Estudos de Tradução.

### **1.2.3 Estudos de Tradução baseados em corpus**

Mona Baker (1993) foi a primeira a introduzir os estudos de corpora nos Estudos da Tradução, com a sua publicação “Corpus Linguistics and Translation Studies: Implications and Applications”. Desde esta introdução que os corpora têm sido cada vez mais usados nos estudos descritivos da tradução, na formação de tradutores e na avaliação da qualidade de tradução. Baker introduz-nos a dois tipos de estudo baseados em corpora: os estudos descritivos e os estudos aplicados.

#### **1.2.3.1 Estudos descritivos baseados em corpora**

Os métodos de investigação baseados em corpus deram um impulso à procura dos chamados universais de tradução, tendo sido estes investigados com corpora paralelos monolíngues comparáveis e bilingues paralelos. O primeiro tipo de corpus divide-se em dois subcorpora: um de textos traduzidos e outro de textos não traduzidos, comparáveis a nível de tipo textual, tópico, distribuição temporal, distribuição a nível de autores e leitores. Este tipo de corpus é utilizado para explorar diferentes padrões linguísticos que ocorrem em textos *translational and non-translational* na cultura-alvo (Laviosa, 2010, pp. 83-84).

Já os corpora bilingue paralelos podem ser uni ou bidirecionais. Um corpus unidirecional (o nosso caso de estudo) contém dois subcorpora: um com textos numa língua A e outro com as suas traduções na língua B. um corpus bidirecional contém quatro subcorpora: um com textos numa língua A, um com as suas traduções numa língua B, um com textos na língua B e outro com as suas traduções na língua A (Laviosa, 2010, p. 84).

Os corpora paralelos são utilizados no estudo de diferentes padrões linguísticos (o caso da prosódia semântica, das colocações e das coligações, por exemplo) nas traduções e nos seus textos de partida.

### 1.2.3.2 Estudos aplicados baseados em corpora

A LCorpus possuiu vários campos de aplicação: estudos léxicos e lexicografia, estudos gramaticais, variação e análise de gênero, ensino e aprendizagem de línguas e estudos da tradução. Em contexto de sala de aula, os corpora paralelos unidirecionais de especialidade são aplicados para se descobrir normas de comportamento tradutório a diferentes níveis de análise linguística, bem como para recuperar e examinar equivalentes terminológico, fraseológico, sintático e estilísticos (Laviosa, 2010, p. 84).

Os corpora comparáveis bilíngues e multilíngues de especialidade (sendo o caso do nosso estudo o primeiro) são habitualmente empregues para descobrirmos equivalentes funcionais tradutórios. Estes tipos de corpus são ainda utilizados como fontes de dados para melhorar a aquisição do conhecimento *subject-specific* (Laviosa, 2010, p. 84).

No que toca à avaliação da qualidade da tradução (TQA), aqui os corpora já são aplicados em estudos que apresentam critérios de avaliação verificáveis e que têm como objetivo estabelecer a que nível é que uma tradução pode ser considerada melhor do que outra (Bowker citado em Laviosa, 2010, p.84). Os dois tipos de corpus que são essenciais para o conjunto de ferramentas do avaliador da tradução: (1) um corpus paralelo unidirecional produzido por tradutores em formação e (2) um corpus de referência ou especialidade de textos comparáveis não traduzidos. Este último tipo de corpus fornece uma norma segundo a qual as traduções dos alunos são avaliadas de acordo com um conjunto de critérios: perceber o campo de atuação; exatidão léxica, terminológica, fraseológica e sintática; fluência e adequação estilísticas.

Este uso de corpora para efeitos de avaliação pode ser ligado com os estudos processuais. A “Encyclopedia of Case Study Research”, declara que os estudos processuais «envolvem o exame empírico de como e porquê alguns fenómenos significativos em evolução ocorrem ao longo do tempo» (Mills et al., 2010, p.736, nossa tradução). Ora, através do estudo do processo de desenvolvimento de uma certa disciplina, no nosso caso a tradução, podemos atingir resultados que satisfaçam o produto pretendido. É aqui que o uso de corpora na avaliação de tradução se mostra como útil na elevação do nível de objetividade em termos do número de erros identificados e corrigidos, na qualidade dos comentários fornecidos pelo avaliador, consistência na identificação de erros léxicos entre avaliadores e a utilidade dos comentários do avaliador para editar o *rascunho* da tradução. Esta abordagem prova ser vantajosa em contexto de sala de aula, uma vez que os dados do corpus proporcionam uma fonte de informação objetiva e concreta e os alunos recebem feedback em tempo real por parte do corpus e também por parte do avaliador (Laviosa, 2010).

### **1.3 A Tradução e a Linguística de corpus**

A utilização de corpora no âmbito dos Estudos da Tradução permitiu analisar e explorar os estilos de tradutores, a influência do inglês nas línguas europeias e na tradução, as características da dobragem (como esta difere entre línguas), entre outros aspetos de relevo (Laviosa, 2010, p. 84). Estas análises de corpora variados do produto e processo de tradução oferecem um excelente exemplo da relação entre a teoria da tradução e os estudos descritivos.

Uma das vantagens do uso de corpora nos Estudos da Tradução é a de podermos utilizar esta abordagem para investigar fenómenos linguísticos que sejam do interesse dos teóricos e também de futuros tradutores, como menciona Shen (2010). Além disso, estas abordagens permitem formular hipóteses sobre os padrões linguísticos presentes em determinados tipos de texto, uma vez que, tal como já realçamos, a linguagem é padronizada.

Para que isto aconteça, o investigador necessita de delinear passos básicos para levar a cabo a sua análise. Shen (2010, p. 182) realça etapas necessárias que adotaremos na nossa abordagem metodológica. O primeiro passo passa por colocar uma hipótese que se queira verificar. No nosso caso de estudo, a hipótese formulada é a seguinte:

*Será que a estrutura e as ocorrências aos níveis terminológico, morfossintático e semântico-enunciativo são semelhantes no texto de partida e no texto de chegada?*

O segundo passo passa por construir e estabelecer um corpus adequado, representativo de um tipo particular de produção e receção da linguagem. Esta fase encontra-se descrita na secção 4.1. do presente relatório. A última etapa é levar a cabo a investigação do nosso problema de investigação e contrastar os corpora e os dados dele provenientes, que podemos observar em maior detalhe na secção 4.5. do presente relatório. Como mencionado anteriormente, os corpora têm a capacidade de exemplificar com que frequência certos itens linguísticos ou estruturas linguísticas são utilizadas.

Deste modo, passemos a analisar as características da linguagem técnica e a sua relação da estilística com a tradução.

#### **1.3.1 Características genéricas e específicas da comunicação e documentação técnica**

Existem certas características-chave comuns à maioria, senão a todos os tipos de documentação técnica, e que se podem revelar ferramentas úteis ao contemplar a tarefa que estamos a desempenhar. Byrne (2014) faz referência a cinco características-chave, algumas delas, aliás, já mencionadas em etapas anteriores e que exploraremos agora.

**A comunicação técnica dirige-se a leitores específicos.** É de conhecimento geral que os documentos técnicos são elaborados como sendo ferramentas orientadas para tarefas que têm como público um grupo(s) particular(es) de pessoas. Alguns documentos podem ser lidos pela audiência pretendida ou pelo público em geral. Isto pode vir a exigir aos redatores técnicos a priorização destas audiências a quem o texto se dirige e a respetiva satisfação das suas necessidades. Um documento técnico deve ser concebido, desenhado e produzido tendo em conta as necessidades deste nicho específico.

**A comunicação técnica é uma ferramenta.** Os textos são ferramentas, ou seja, um meio para atingir um fim. O objetivo deste tipo de comunicação é transmitir a informação certa às pessoas certas, no formato correto, para que estas possam prosseguir com o seu trabalho.

**A comunicação técnica muitas vezes é produzida em equipa.** Os documentos técnicos são muitas vezes elaborados por uma equipa de redatores técnicos. O que poderá suceder ao ter várias pessoas a trabalhar no mesmo documento é acontecerem discrepâncias ao nível de estilo do documento. Isto pode acontecer devido ao diferente tipo de escrita de cada um dos redatores ou ao guia de estilo para determinado documento não ter sido seguido por parte da equipa. Para os tradutores, isto dificulta o processo de tradução, podendo, muitas vezes, o mesmo conceito ser referido através de diferentes termos ou chegarmos a um capítulo de um manual de instruções, por exemplo, e este deixar de fazer sentido. Da mesma forma que um conjunto de redatores técnicos pode elaborar um documento técnico, também uma equipa de tradutores se pode ocupar da tradução de um documento. Novamente, neste caso, levanta-se a problemática de inconsistências a nível de estilo, pois, à semelhança dos redatores técnicos, também os tradutores têm diferentes estilos de tradução. Estas inconsistências presentes no processo de tradução serão corrigidas e assinaladas, normalmente, por um editor ou um tradutor sénior. Estabelecendo a ligação teórica com o que acontece na prática, temos como exemplo o que acontecia na nossa entidade acolhedora do estágio, a AP|Portugal. Na empresa, havia a existência de guias de estilo de tradução para determinados clientes e/ou tipos de projetos, de forma que uma tradução colaborativa fluísse da melhor forma, ficando esta mais uniforme e acabando por diminuir, também, o trabalho do revisor.

**A comunicação técnica utiliza o design para facilitar o seu uso.** Fatores simples, como o tipo de letra e o tamanho, as cores e a quantidade de espaço em branco, podem fazer toda a diferença na eficácia com que o leitor absorve a informação que lhe está a ser transmitida.

**A comunicação técnica utiliza diversas tecnologias.** Espera-se que, na produção de documentação técnica, seja utilizada uma variedade de ferramentas e tecnologias distintas, conclusão à qual conseguimos chegar “a olho nu”, examinando, por exemplo, um manual de instruções, uma vez que estes possuem grafismos. Como consequência desta variedade de tecnologias e ferramentas a ser usadas no processo de redação, também o tradutor ficará exposto a esta diversidade de tecnologias. Esta variedade é uma das características da área.

A multiplicidade de informação contida no domínio da comunicação técnica molda toda a natureza dos documentos individuais e dá origem a uma gama de diferentes tipos de texto. Devido a esta pluralidade, há características que são transversais e específicas à documentação técnica. Das já mencionadas por Byrne (2014), realçaremos, em concreto, a linguagem e as referências.

O princípio base dos textos técnicos é a utilização de uma linguagem clara e direta, assim como manter a simplicidade textual, isto porque, muitas vezes, o recetor do texto técnico, quando está a ler o conteúdo a que se expôs, está a realizar outra tarefa ou a usar o texto técnico como ferramenta de auxílio. Ao mantermos a linguagem simplificada, reduzimos o risco de algo ser mal interpretado pelo leitor. A clareza e simplicidade dos textos técnicos podem ser aprimoradas através da utilização de informação declarativa e fornecimento de instruções claras, simples e em ordem cronológica, no caso dos manuais de instruções. Como mencionado anteriormente, o texto técnico não se resume apenas à terminologia. Aliás, a terminologia representa apenas uma pequena parte do total de palavras presentes num texto técnico. Newmark (citado em Byrne, 2014, p. 51) estimou que apenas 5-10% de um texto seria composto por terminologia da área.

Outro ponto importante são as referências. Um autor remete o leitor para informações contidas noutras partes do documento ou noutros documentos. Um exemplo comum são os manuais de apoio ao utilizador, como o de um software ou, um exemplo do quotidiano, o manual de um automóvel. No caso de documentos relativos a software, estas referências podem ser feitas ao nível de diagramas, itens de interface, menus ou documentos de acompanhamento.

Tendo por base as noções de texto técnico, as suas características, assim como as noções de tradução e linguística de corpus, e como o estudo de corpus pode ser benéfico nos Estudos da Tradução, iremos conduzir o nosso estudo apresentado na secção 4 do presente documento, com o intuito de percebermos se as características, a estrutura e as ocorrências aos níveis terminológico, morfossintático e semântico-enunciativo passam do texto fonte para o texto de chegada.

## 2 A Entidade de acolhimento do estágio

### 2.1 A AP | Portugal

A entidade escolhida para a realização do estágio curricular, que será descrito ao longo deste relatório, foi a AP|Portugal – Tech Language Solutions<sup>2</sup>, uma empresa portuguesa que nasceu em 1998, sendo a única empresa portuguesa certificada pelas Normas Internacionais da Qualidade ISO 17100<sup>3</sup> e ISO 18587<sup>4</sup> e, mais recentemente, ISO 23155<sup>5</sup>. Embora a tradução seja o seu principal foco de atividade, a AP|Portugal garante uma vasta gama de serviços, como interpretação, transcrição, legendagem, locução, transcrição e, mais recentemente, serviços tecnológicos de localização de websites, software e apps, gestão de conteúdos multilingues e criação de websites multilingues, entre outros.

Conta também com posicionamento nos mercados nacional e internacional, tendo sede em Vila Nova de Gaia e escritórios em Lisboa. De forma a apoiar os países e as variantes da língua portuguesa, a empresa opera também no mercado brasileiro e angolano, com a AP|Brasil e a AP|Angola.

**Figura 1**

*Página web da AP|Portugal*



A AP|Portugal associa as normas ISO com a metodologia Kaizen, uma metodologia que assenta numa melhoria contínua que envolve todos os membros de uma organização. Esta fusão é vista como uma mais-valia, trazendo diversos ganhos para as empresas, como a

<sup>2</sup> As informações presentes neste capítulo têm como fonte o website oficial da empresa (<https://go.apportugal.com/>) assim como documentos internos disponibilizados no início do estágio.

<sup>3</sup> A ISO 17100 tem como objetivo definir requisitos de qualidade e certificar os serviços de tradução durante as diferentes fases do processo de tradução.

<sup>4</sup> A ISO 18587 tem o objetivo de definir requisitos de qualidade para o processo de pós-edição completa e humana da tradução automática, do seu produto e das competências dos pós-editores.

<sup>5</sup> A ISO 23155 tem como objetivo definir requisitos gerais e recomendações para a prestação de projetos de interpretação de conferência.

consciencialização dos colaboradores para serem parte ativa da mudança e «que as pessoas podem (...) melhorar o desempenho das suas funções, ajudando à melhoria do desempenho da empresa, mas sem deixar de atender às necessidades individuais» (AP | Portugal, 2020).

Além da preocupação da empresa com a prestação de serviços de qualidade certificados juntamente com a metodologia Kaizen, a AP encontra-se ligada institucionalmente ao ApoioXXI, um centro de ação psicopedagógica (de onde derivam as iniciais “AP”), comprometendo-se a contribuir para o desenvolvimento social das crianças, permitindo que aquelas com necessidades educativas especiais usufruam de apoio técnico especializado, nas áreas da psicologia, terapia da fala, psicomotricidade e ensino especial.<sup>6</sup>

Como organismo de tradução e forte aliado do associativismo, a AP | Portugal destaca-se por ser membro de diversas associações nacionais e internacionais de tradutores e empresas de tradução, de entre as quais destacamos a GALA (Globalization and Localization Association), a ELIA (European Language Industry Association), a ALC (Association of Language Companies), a ATC (Association of Translation Companies), a LEXIS (Comunidade Internacional de Profissionais em Serviços Linguísticos), a EUATC (European Union of Associations of Translation Companies), a APET (Associação Portuguesa de Empresas de Tradução) e a APTRAD (Associação Portuguesa de Tradutores e Intérpretes).

Adicionalmente às inúmeras associações e empresas de tradução da qual a AP | Portugal faz parte, a empresa é ainda a representante oficial do programa Wordbee em Portugal, uma CAT tool (software assistente de tradução online) que permite traduzir, rever, criar memórias de tradução e organizar documentos. Sendo uma ferramenta online, não requer instalação e é de bastante fácil utilização. A AP disponibiliza também quer para os seus colaboradores, formandos e estagiários ou para a população em geral, tutoriais em vídeo e ações de formação, sejam estas presenciais nos seus escritórios ou através da plataforma Tech Training Center<sup>7</sup>.

Sendo uma empresa multidisciplinar e atuando em diversas áreas dos serviços linguísticos, a empresa conta com o apoio de diversos colaboradores internos e externos, o que permite um leque bastante alargado de oferta de pares linguísticos, contemplando as línguas mais comuns como o inglês, espanhol, alemão e francês, assim como línguas menos comuns, como o turco, chinês, ucraniano, russo, entre outros. Aliado à variedade de pares linguísticos, a empresa atua

---

<sup>6</sup> Informações disponíveis no website do centro ApoioXXI (<https://www.apoioxxi.com/>)

<sup>7</sup> Página web do centro de formação online da empresa: <https://formacao.apportugal.com/>

também em diferentes setores da tradução, destacando os campos empresarial, político, económico, técnico, científico e jurídico.

## **2.2 Constituição da empresa**

Tendo em conta o número de serviços que a AP|Portugal oferece, a empresa está dividida em vários departamentos com a finalidade de apoiar todos os seus colaboradores internos e externos, em regime *freelancer*. Quanto à sua constituição, a empresa apresenta os seguintes departamentos:

- DIRI (Departamento Informativo e de Relações Internacionais): responsável pelas relações B2B (*business to business*) tanto no mercado nacional como internacional;
- CATTI (Centro de Apoio aos Tradutores, Transcritores e Intérpretes): responsável pela gestão de projetos e pela gestão de qualidade, sendo o departamento que estabelece a ponte entre os clientes e os prestadores de serviços linguísticos, prestando apoio na terminologia e dúvidas/dificuldades que estes possam ter;
- PACQ (Departamento de Paginação e Controlo de Qualidade): responsável pela paginação, pela realização de DTPs (*Desktop Publishing*, em português, paginação eletrónica) e pelo controlo de qualidade de todos os projetos realizados pela empresa;
- Departamento Jurídico: responsável pelos textos jurídicos, apostilhas e certificações;
- DAF (Departamento Administrativo e Financeiro): responsável pela parte financeira da empresa;
- MARCOM (Departamento de Marketing e Comunicação): responsável por todas as questões relacionadas com o marketing, como a manutenção da página e das redes sociais da empresa.

Além destes departamentos mencionados supra, destaco a existência do AP|Academy, inserido no departamento de Recursos Humanos. O AP|Academy é composto pelos estagiários da empresa tendo, durante os seis meses de estágio, contado sempre com 10 a 13 elementos. Podemos observar em maior pormenor a organização e estruturação da empresa no organograma apresentado no Anexo I.

## **2.3 Caraterização do estágio**

O estágio em análise teve uma duração de seis meses, com início a 1 de fevereiro e fim a 31 de julho. Este teve lugar de segunda a sexta-feira, das 9h00 às 18h00, e foi realizado em regime remoto.

Apesar de o estágio ter decorrido 100% em regime remoto, a comunicação com os restantes membros da equipa não era afetada, uma vez que a empresa faz uso da plataforma Workplace, uma ferramenta de comunicação empresarial que combina o envio de mensagens instantâneas, realização de videoconferências, compartilhamento de notícias e criação de grupos de trabalho<sup>8</sup>. Era através desta plataforma que toda a empresa e os estagiários ficavam a par das últimas notícias em relação à empresa e que comunicavam com os seus mentores, os linguistas seniores, os restantes estagiários e qualquer outro membro interno da empresa. Além disso, todos os dias, perto da hora de início do dia de trabalho, os responsáveis pelo departamento do Academy, Ricardo Silva e Teresa Peixoto, perguntavam aos estagiários quais os planos para aquele dia. Cada estagiário informava que tipo de tarefas iria desempenhar ou se estava à espera de que um projeto lhe fosse atribuído.

Semanalmente, e através da plataforma Zoom, decorria uma reunião de acompanhamento em grupo de todos os mentorandos e mentores, onde eram relatadas as semanas de cada um, que tarefas tinham sido realizadas e eram expostos os principais problemas e dificuldades encontrados por cada um, recebendo ainda conselhos e apoio por parte dos mentores bem como dos outros colegas estagiários. Estas reuniões provaram ser um espaço de partilha de experiências, de reforço do espírito de equipa e confiança nos restantes colaboradores e que permitia o relacionamento interpessoal apesar das limitações impostas pelo regime remoto.

Antes de o estagiário iniciar formalmente as suas tarefas, as primeiras semanas de estágio foram dedicadas à formação exigida e fornecida pela empresa, composta por diversos módulos e temáticas. A maioria das formações era acompanhada de suporte em vídeo e também em material escrito, sendo a maioria das formações ministradas por colaboradores internos da empresa. No total, foram realizadas sete formações internas, de entre as quais destaco a formação sobre Gestão de Projetos e Tradução<sup>9</sup>, que contava com vinte módulos, dando especial atenção aos módulos relativos às diferentes fases de produção, de controlo de qualidade e introdução às normas ISO 17100 e ISO 18587, de formação em ferramentas CAT e dos erros mais comuns em tradução. Após a conclusão destas formações, foi atribuído um projeto de tradução interna, isto é, traduções de materiais da empresa e não de terceiros, de forma a nos familiarizarmos com a ferramenta utilizada pela AP | Portugal e, em simultâneo, de treinar e esclarecer dúvidas sobre terminologia e sobre a própria ferramenta – o Wordbee.

---

<sup>8</sup> Informação disponível no website da ferramenta Workplace (<https://www.workplace.com/>)

<sup>9</sup> Formação disponível no Tech Training Center da AP | Portugal (<https://formacao.apportugal.com/courses/gestao-projetos-traducao>)

No fim deste projeto interno, deu-se início à realização de projetos externos, requeridos por clientes da empresa, que variavam consoante o tipo de projeto, tipologia textual, prazo de entrega e também ferramenta de trabalho.

Ao longo dos seis meses de estágio foram realizados diversos tipos de projetos, entre os quais traduções, releituras (para familiarização de temáticas e construção de glossários), pré e pós-edições de tradução automática, transcrições e DTP, projetos estes que serão apresentados e descritos em detalhe no capítulo seguinte.

Porém, antes de avançarmos para o capítulo seguinte, importa sublinhar as diferenças entre a noção de releitura e revisão, do ponto da empresa e do ponto da estagiária. Uma revisão trata de analisar comparativamente o TC e o TP, de forma a verificar que o sentido do texto original foi respeitado; já uma releitura analisa apenas o TC tendo em conta determinadas categorias – gramática, ortografia, pontuação, fraseologia e coerência. Estas são as definições que podemos encontrar no website da AP|Portugal<sup>10</sup>. A propósito, podemos ainda comparar estas definições com as propostas apresentada na norma ISO 17100:2015. A norma propõe, assim, três definições para descrever o processo de revisão: *revision* (revisão), *review* (revisão monolíngue) e *proofreading* (releitura). A noção de revisão presente no website da empresa em que o nosso estágio foi efetuado corresponde, de facto, à de revisão presente na norma, assim como a de releitura. Já a norma apresenta uma etapa entre revisão e releitura, designada *review* ou revisão monolíngue, que consiste na revisão do conteúdo da língua de chegada relativamente à sua adequação à finalidade acordada. No entanto, para a estagiária, pessoalmente, a função de releitura envolverá não só os parâmetros descritos anteriormente, mas é também uma forma de o tradutor conseguir criar glossários e memórias de tradução mais atualizadas e ainda de familiarização com temáticas com as quais não trabalhou diretamente.

---

<sup>10</sup> Definições retiradas e adaptadas do website da AP|Portugal (disponível em: <https://www.apportugal.com/contactos/perguntas-frequentes/55-pt/>)

### 3 Tarefas Realizadas

O estágio revelou-se um período de aprendizagem e trabalho bastante versátil, apesar de ter realizado maioritariamente projetos de tradução, pós-edição de tradução automática (doravante designado como PEMT), releitura e transcrição, assim como tarefas pontuais de DTP e legendagem.

Neste capítulo, serão expostas e exploradas com mais detalhe as tarefas desenvolvidas ao longo do estágio, as ferramentas utilizadas ao longo do estágio, as línguas de trabalho e os temas, assim como dificuldades encontradas e correções feitas pelos revisores internos.

#### 3.1 Diário de Bordo

Durante os seis meses de estágio foi mantido um diário de bordo como ferramenta de apoio, com o objetivo de registar e acompanhar as tarefas desenvolvidas diariamente, assim como as suas principais características, de modo a poder refletir sobre a variedade e quantidade de trabalhos e projetos no final do estágio. Este diário foi criado para uso pessoal, embora a empresa pedisse o preenchimento de um relatório mensal, onde deveriam constar todos os projetos e atividades realizados ao longo do mês, assim como a ferramenta CAT utilizada, o número de palavras e o tempo investido em cada tarefa (ver exemplo no [Anexo II](#)). O formato do diário de bordo foi inspirado no formato do relatório mensal da empresa, facilitando, assim, o registo e uniformidade de informação.

A ferramenta mais útil para este registo foi um documento Google Sheets, onde foi criada uma tabela que continha o número de projeto, a data de entrega, o domínio temático, o par linguístico, a ferramenta CAT utilizada, a duração da tarefa e o número de palavras, tudo isto preenchido de acordo com os dados fornecidos pelo Wordbee. Esta categorização apresentada no Wordbee, feita pelo gestor de projeto, seguindo um modelo simplificado de *translation brief*, permitia logo, de forma imediata, um contacto inicial e primeira impressão do que seria o conteúdo do projeto e, ao mesmo tempo, estipular uma possível duração de realização de acordo com o número de palavras apresentado, assim como a correspondência de palavras já presente em memórias de tradução. Para melhor filtragem, os tipos de tarefas foram divididos por folhas de cálculo diferentes, de forma a permitir uma melhor visualização dos dados. Abaixo, nas figuras 2 e 3, podemos visualizar as primeiras entradas do diário de bordo dos projetos de tradução e dos projetos de PEMT.

**Figura 3***Primeiras entradas de projetos PEMT*

	A	B	D	E	F	G	H
1	<b>PEMT</b>						
2	<b>Nº</b>	<b>Data</b>	<b>Tema</b>	<b>Par linguístico</b>	<b>CAT tool</b>	<b>Duração da tarefa</b>	<b>Nº de Palavras</b>
3	1	31/03/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:17:00	104
4	2	31/03/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:08:00	156
5	3	10/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:31:00	153
6	4	10/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:08:00	83
7	5	11/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:20:00	305
8	6	12/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:16:00	413
9	7	12/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:08:00	63
10	8	13/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:03:00	30
11	9	13/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:02:00	21
12	10	17/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:09:00	128
13	11	18/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:01:00	2
14	12	18/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:06:00	187
15	13	18/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:03:00	66
16	14	19/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:08:00	176
17	15	19/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:07:00	27
18	16	28/04/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:26:00	350
19	17	02/05/2023	Finance/Cryptocurrency	EN>PT	Smartling	0:02:00	8

**Figura 2***Primeiras entradas de projetos de Tradução*

	A	B	D	E	F	G	H
1	<b>Tradução</b>						
2	<b>Nº</b>	<b>Data</b>	<b>Tema</b>	<b>Par linguístico</b>	<b>CAT tool</b>	<b>Duração da tarefa</b>	<b>Nº de Palavras</b>
3	1	13/02/2023	Artigos de Blog	EN>PT	Wordbee	5:36:00	4365
4	2	16/02/2023	Technology > Kitchen Appliances/Recipes	EN > PT	memoQ	1:14:00	364
5	3	16/02/2023	Industry & Technology	EN > PT	Memsources	0:15:00	31
6	4	17/02/2023	Computers & Internet	EN > PT	Smartcat	0:07:00	97
7	5	17/02/2023	Business	EN > PT	Smartcat	0:31:00	397
8	6	17/02/2023	Computers & Internet	EN > PT	Smartcat	0:24:00	217
9	7	20/02/2023	Technology > Kitchen Appliances/Recipes	EN > PT	memoQ	0:10:00	56
10	8	20/02/2023	Technology > Kitchen Appliances/Recipes	EN > PT	memoQ	0:04:00	20
11	9	20/02/2023	Technology > Kitchen Appliances/Recipes	EN > PT	memoQ	0:20:00	228
12	10	20/02/2023	Computers & Internet	EN > PT	Smartcat	1:09:00	804
13	11	22/02/2023	Industry & Technology	EN > PT	Smartcat	0:26:00	275
14	12	22/02/2023	Industry & Technology	EN > PT	Smartcat	0:16:00	424
15	13	22/02/2023	Industry & Technology	EN > PT	Smartcat	0:12:00	430
16	14	23/02/2023	Industry & Technology	EN > PT	Smartcat	0:19:00	881

Este diário de bordo pretendia seguir o modelo de *translation brief*, um conjunto de instruções que acompanha um pedido de tradução que descreve o modo como o texto deve ser tratado. Para o tradutor, o *translation brief* providencia informações importantes sobre o contexto no qual o texto deve ser tratado. É também uma ferramenta essencial para qualquer profissional de tradução enquadrar o seu trabalho, pois nele encontramos a função, o público-alvo, a data de receção, o local de receção, o meio e o motivo. Nord (1997/2008) atribuiu maior importância ao *translation*

*brief* para os tradutores em formação, uma vez que os tradutores profissionais com mais experiência já sabem o que cada cliente exigirá consoante determinados tipos de texto. Cada tarefa de tradução deve, então, ser acompanhada por um resumo que define as condições sobre as quais o texto de chegada deve atuar, qual o seu propósito final e através de que meio é que será difundido. Isto advém da ideia de que a situação comunicativa determina as características verbais e não-verbais de um texto. Assim, podemos assumir que a descrição de fatores situacionais define a *posição* na qual o texto se deve enquadrar, isto aplicando-se ao TP e ao TC. Existe, assim, um conjunto de fatores extra e intratextuais que fazem com que uma tradução tome um significado específico ou mesmo diferente daquele que o tradutor pensaria inicialmente.

Os fatores extratextuais incluem o emissor/ produtor textual, o recetor (quem será o público-alvo), o meio ou suporte material, a intenção do emissor ao produzir este texto, o local de publicação do texto, o tempo (momento de publicação, desfasamento entre a publicação do TP e do TC) e o motivo que levou o autor a produzir o texto. Do conjunto destes setes fatores, resulta a função do texto, que Nord (1997/2008) divide em quatro categorias principais: referencial, expressiva, apelativa e fática. Adicionalmente a estes fatores, encontramos os fatores intratextuais: o tema, o conteúdo, os pressupostos pragmáticos e situacionais, a estruturação do texto, os elementos não verbais (e.g. imagens), o léxico, a sintaxe e os elementos suprasegmentais.

Já a tradução simultânea pode ser considerada uma exceção em relação à diferença ao nível do local, tempo, motivo e propósito da comunicação, mas mesmo nestas situações temos de considerar uma diferença relacionada ao conhecimento ligado à cultura, experiência ou suscetibilidade do público-alvo.

No entanto, na AP | Portugal, quem realizava o *translation brief* eram os gestores de projeto e não os tradutores, fossem estes estagiários ou colaboradores da empresa. Sendo o gestor de projetos a realizar o *translation brief*, este encontrar-se-á mais bem equipado para providenciar informações sobre o projeto ao tradutor, ou seja, no caso da tradução técnica (maioria da tradução realizada no estágio), o facto de ser o gestor de projetos a tratar do *translation brief* permite que o gestor consiga escolher mais facilmente o tradutor que se adequa à temática a ser tratada, visto que os tradutores se especializam mais numa área técnica do que noutra. Sendo uma ferramenta gerida pelo gestor de projetos, torna ainda mais facilmente a que a informação do *brief* seja facilmente partilhada posteriormente com o revisor.

Contudo, a elaboração de um *brief* deveria ser considerada um processo colaborativo entre o tradutor e o gestor de projetos. Podemos dividir funções entre estas duas entidades. O gestor de

projetos é o que interage diretamente com o cliente, estando a par do âmbito do projeto, os objetivos, o público-alvo e outras especificações do cliente; é quem lida com as conversações de orçamentação, cronogramas e logística geral do projeto; e quem está responsável por garantir que as questões legais e éticas, como acordos de confidencialidade, são cumpridas e abordadas. Já o tradutor é o responsável pelos aspetos linguísticos do briefing, conseguindo assegurar que todos os detalhes sobre o TP são representados com precisão; podem ainda fornecer informações sobre glossários e bases terminológicas que possam auxiliar o processo de tradução e, também, fornecer informações sobre potenciais desafios linguísticos. Em última análise, uma colaboração eficaz entre o tradutor e o gestor de projetos é fundamental para produzir um *brief* abrangente e preciso. Uma comunicação clara, uma compreensão partilhada dos objetivos do projeto e o aproveitamento da experiência de ambas as funções contribuem para o sucesso do projeto de tradução.

Por conseguinte, não se verificando esta colaboração no processo de elaboração do *brief*, este era preenchido pelo gestor, fazendo com que as informações que chegam aos tradutores já viessem filtradas naquilo que é considerado o mais importante para o decorrer do projeto (número de palavras, domínio temático e instruções). As instruções para cada projeto raramente vinham diretamente do cliente; porém, quando tal acontecia, consistiam maioritariamente na adequação da tradução a uma plataforma específica (por exemplo, diferentes formas de tratamento a serem utilizadas em diferentes plataformas) ou a utilização específica de uma forma de tratamento quando nos referíamos a uma pessoa durante a tradução (os pronomes, por exemplo). Estas nuances poderão ser observadas mais à frente.

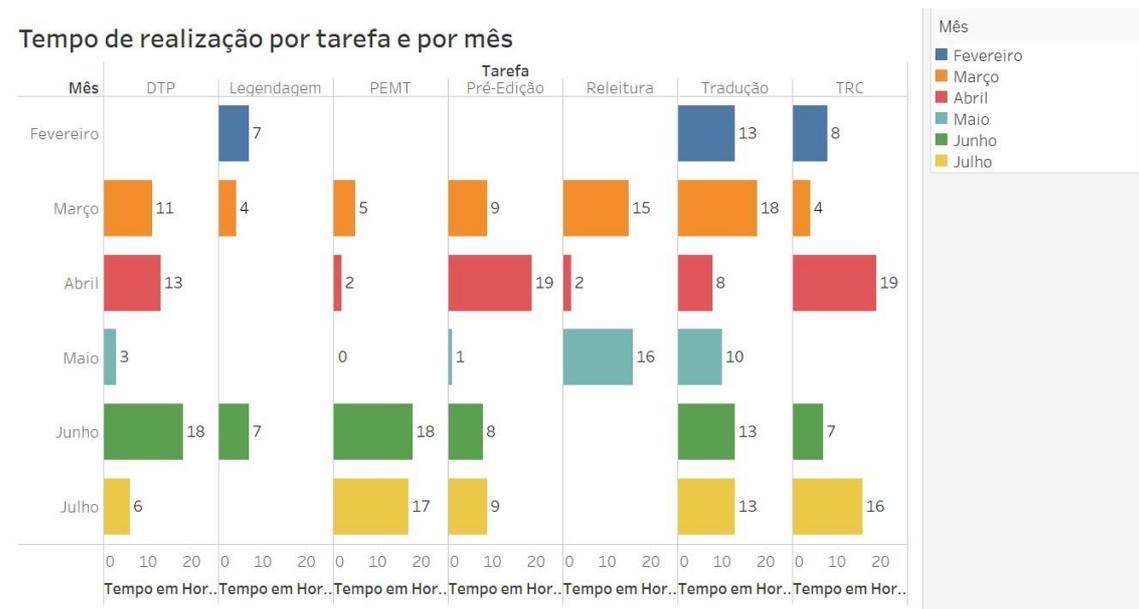
Todavia, o modelo de diário de bordo por nós criado teve em consideração grande parte dos elementos constituintes de um *translation brief*, considerando as informações limitadas concedidas, algo que se mostrou benéfico, permitindo, assim, uma melhor organização de um diário de bordo e contabilização de tarefas para o desenvolvimento do nosso relatório.

### 3.2 Trabalho realizado

As ferramentas mais úteis para o registo e contagem do trabalho realizado foram os relatórios mensais da empresa e o diário de bordo criado. À semelhança do diário de bordo mencionado no ponto anterior, o relatório mensal da empresa continha diversas folhas, cada uma correspondente a uma diferente tarefa realizada com o objetivo de reportar ao mentor no final do mês para que este pudesse ver tudo aquilo que foi feito. Quando era recebida uma tarefa, esta era logo registada no relatório, de modo a mantê-lo atualizado, sendo a coluna relativa ao tempo preenchida no fim da tarefa. Com o preenchimento do relatório, é possível elaborarmos um gráfico que ilustre o número de horas gasto em cada tarefa ao longo do estágio, por cada mês de trabalho, conforme observado abaixo, no Gráfico 1. Para uma visualização em tamanho maior deste gráfico, será possível consultar o [Anexo III](#).

**Gráfico 1**

*Tempo gasto, em horas, por tarefa e por mês*

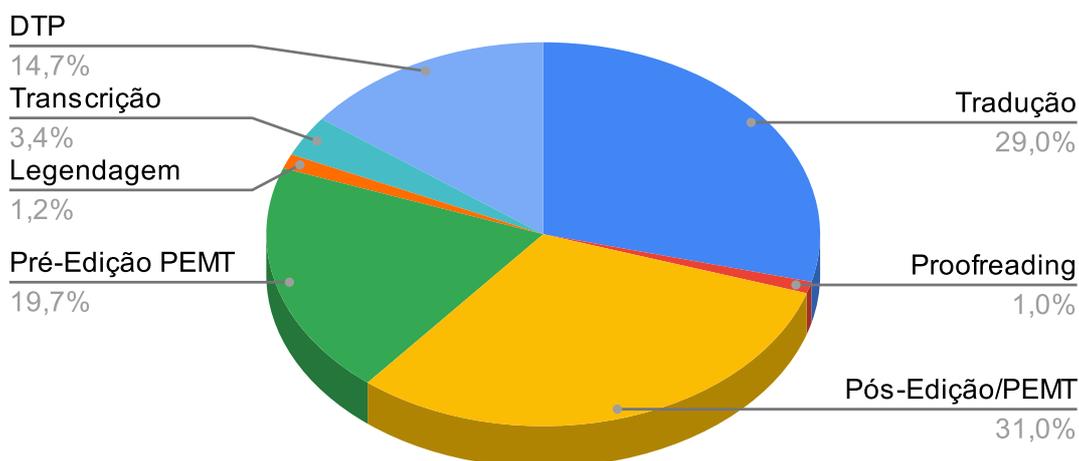


Através da combinação dos dados dos relatórios mensais com a estrutura do diário de bordo, foi possível verificar que durante os seis meses de estágio foram realizados 407 projetos diferentes. No gráfico 2, podemos observar a percentagem de projetos diferentes desenvolvidos, constatando que a maior percentagem de tarefas desenvolvidas foi a de Pós-Edição/PEMT, com 126 projetos realizados, seguido de tradução, com 118 projetos realizados. De Pré-Edição PEMT contamos com 80 projetos desenvolvidos, 60 projetos de DTP, 14 projetos de Transcrição, 5 projetos de Legendagem e 4 projetos de Releitura.

### Gráfico 2

*Diversidade de projetos desenvolvidos*

## Diversidade de projetos desenvolvidos



Para expor em maior detalhe o volume de tarefas realizadas dentro das já mencionadas no gráfico acima, registou-se a produtividade de cada um dos seis meses de estágio. Para os projetos de tradução, foram contabilizadas todas as palavras traduzidas, independentemente da ferramenta utilizada e do tipo de tarefa realizada. Para os projetos de transcrição e legendagem, foram contabilizadas o total de palavras transcritas e legendadas, assim como a soma da duração dos ficheiros áudio, em horas, realizados durante o mês. Já para os projetos de DTP, foram contabilizados os números de páginas tratados ao longo do mês.

**Tabela 1***Volume de trabalho realizado*

Mês	Tradução		PEMT		Transcrição	Legendagem	DTP ( <i>Desktop Publishing</i> )	
	Tradução	Releitura	Pré-Edição	Pós-Edição			Orçamentação	Tradução
Fevereiro	10 153	—	—	—	33 726 (8:26:14)	5 022 (1:52:49)	—	—
Março	33 624	26 525	20 635	2 422	52 361 (5:40:08)	5 617 (1:21:21)	582	80
Abril	8 731	4 739	35 882	2 004	37 028 (4:38:15)	—	424	88
Maio	31 133	18 934	1 377	26 348	—	—	117	219
Junho	8 842	—	12 686	18 156	9 192 (1:00:44)	5 873 (0:46:15)	24	495
Julho	24 428	—	13 746	59 879	21 517 (2:01:48)	—	205	433

Como se pode evidenciar, o mês de fevereiro foi o mês que registou menor produtividade, especialmente por ter sido o primeiro mês de estágio e a maior parte de o mês ter sido dedicada a formação. Quanto ao mês mais produtivo, foi o mês de março, onde se registou maior nível de tarefas realizadas e de diferentes tipos de tarefa. Este volume de projetos e diferentes tipologias permitiu o desenvolvimento e aprimoramento de novas competências e a aquisição de novos conhecimentos na ótica do tradutor, tais como competência cultural, competência em pesquisa, aquisição e processamento da informação e competência técnica. Ao adquirir e reforçar o conjunto de competências de tradutora proveniente do contexto da formação académica, era mais fácil prever que tipo de dificuldades poderia encontrar em determinado tipo textual, como os textos técnicos, que são, naturalmente, caracterizados por uma terminologia específica e complexa, muitas vezes exigindo conhecimento técnico e especializado em diferentes áreas, o seu conteúdo (onde prevalecem os produtos ou serviços técnicos), o meio pelo qual será difundido e os processos de trabalho, onde a prática deste tipo de tradução exige o recurso a instrumentos de controlo terminológico, guias de estilo, normas e manuais. A prática adotada no contexto desta tipologia textual permitiu que fosse estabelecido um fluxo de trabalho facilitador da tarefa, fluxo esse que passava pela procura online de textos paralelos da mesma empresa, produto e/ou temática, criação e consulta de glossários de uso interno partilhados entre todos os colaboradores da empresa, consulta de traduções prévias do mesmo produtor e ainda consulta das bases terminológicas disponibilizadas pelo cliente (se estas fossem fornecidas) e online, como a IATE e o Eur-Lex, por exemplo.

### 3.3 Línguas e domínios de trabalho

No que diz respeito às combinações linguísticas trabalhadas, apesar de a estagiária possuir competências para tradução de e para língua inglesa e língua francesa, o par linguístico trabalhado durante os seis meses de estágio foi o de EN > PT europeu.

Já no que toca aos domínios trabalhados, por intermédio do Wordbee era possível verificar a que domínio e/ou tipologia pertencia o projeto em questão antes de iniciar a tradução. Através desta catalogação, foi facilitado o registo das tipologias trabalhadas, permitindo ainda, ao longo do estágio, ter uma noção mais clara de quais as tipologias mais exploradas pela estagiária, bem como as tipologias mais trabalhadas pela empresa. A variedade de domínios explorados possibilitou um considerável desenvolvimento das competências tradutórias da estagiária, competências estas explícitas na secção 3.1.3 da norma ISO 17100, pela qual a empresa se rege.

- Competência tradutória: consiste na capacidade de traduzir conteúdo de acordo com o objetivo do projeto de tradução e do acordo cliente-PSL;
- Competência linguística e textual na língua de partida (LP) e na língua de chegada (LC): a capacidade de compreender o texto de partida, a fluência na língua-alvo e o conhecimento das respetivas convenções textuais. Inclui ainda a capacidade de aplicar tal conhecimento ao traduzir ou produzir conteúdos na língua de chegada;
- Competência cultural: aptidão de utilizar informação referente a padrões comportamentais, terminologia atual, sistemas de valores e *locales* que caracterizam as culturas de chegada e de partida;
- Competência em pesquisa, aquisição e processamento da informação: capacidade de adquirir conhecimento especializado e linguístico adicional necessário para perceber o conteúdo da LP e para produzir conteúdo na língua-alvo. Esta competência requer experiência com ferramentas de pesquisa e capacidade de desenvolver estratégias para o uso eficiente de fontes de informação;
- Competência de domínio: capacidade de compreender o conteúdo produzido na LP e reproduzi-lo na LC usando estilo e terminologia próprios;
- Competência técnica: o conhecimento, capacidades e competências necessárias para realizar as tarefas técnicas no processo de tradução ao empregar recursos técnicos incluindo ferramentas e sistemas informáticos que apoiem todo o processo de tradução.

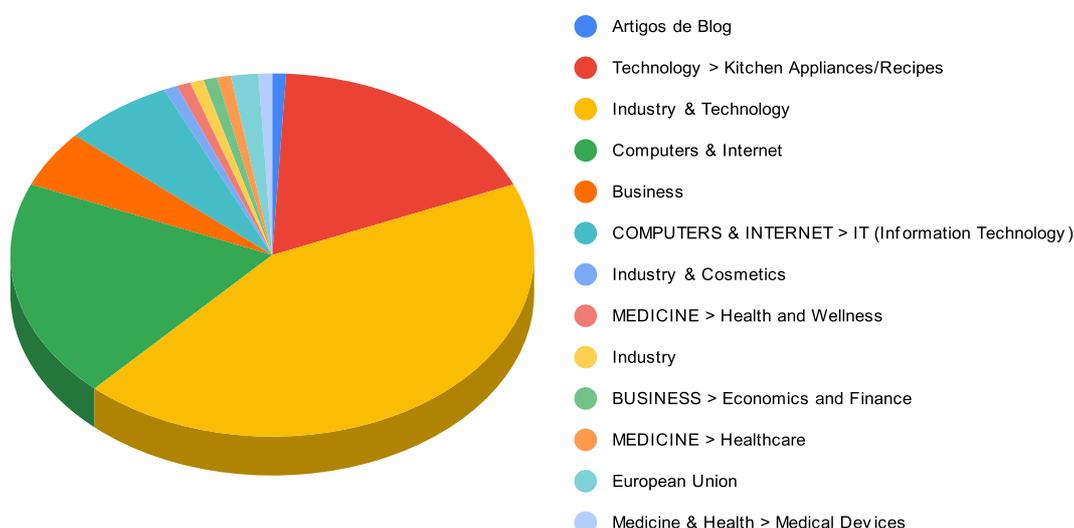
O EMT (Mestrado Europeu em Tradução) destaca ainda a competência tecnológica do tradutor. Esta «inclui todos os conhecimentos e competências utilizados para implementar tecnologias de tradução presentes e futuras no âmbito do processo de tradução. Inclui também conhecimentos básicos de tecnologias de tradução automática e a capacidade de implementar a tradução automática de acordo com as potenciais necessidades tecnológicas.» (Sousa, 2023, p. 7).

### Gráfico 3

*Domínios dos projetos de tradução*

## Domínios

Projetos de tradução



Como podemos ver na imagem acima, o domínio mais trabalhado em projetos de tradução é o da “Indústria e Tecnologia”, domínio que abrangia tradução técnica, localização de software<sup>11</sup> e tradução de manuais de apoio.

<sup>11</sup> A localização pode ser definida como a adaptação linguística e cultural de um produto digital para um determinado *locale* de um mercado estrangeiro. (Schäler, 2010)

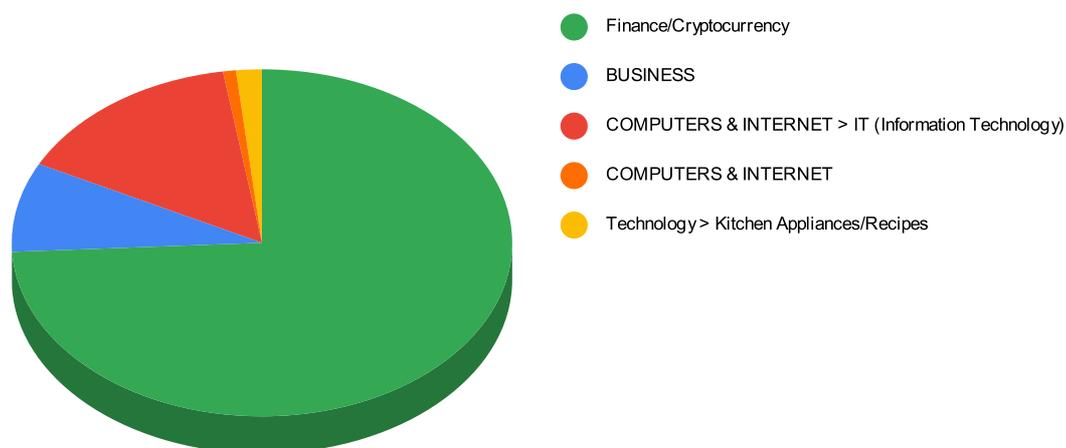
Já nos projetos PEMT, verifica-se que o domínio mais presente é o de “Finanças/Criptomoedas”<sup>12</sup> seguido do domínio de “Computadores & Internet”. Estes domínios passavam pela localização de publicidade ou artigos de blogue e FAQs, assim como manuais de utilização de software.

#### Gráfico 4

*Domínios dos projetos PEMT*

## Domínios

Projetos PEMT



### 3.4 Ferramentas CAT utilizadas

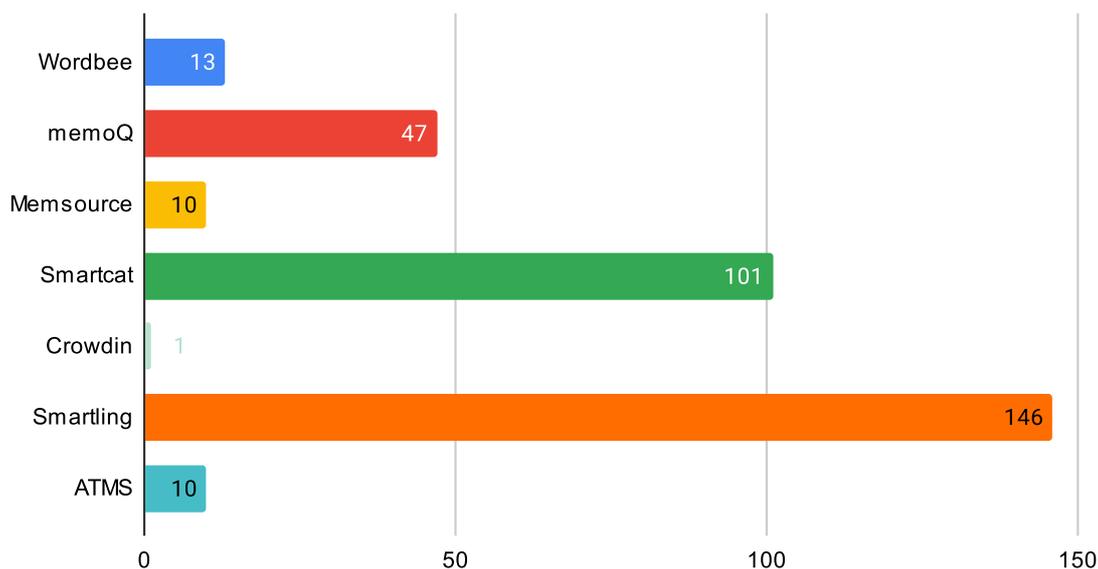
Relativamente às ferramentas CAT utilizadas ao longo do estágio, no total, foram empregues sete ferramentas diferentes para a realização de projetos. A escolha da ferramenta tanto dependia da preferência do cliente como da própria empresa, que normalmente possuía ferramentas específicas para um determinado cliente.

<sup>12</sup> Esta categorização de domínios foi retirada da plataforma Wordbee, onde o gestor de projetos inseria o domínio e/ou subdomínio da tradução a realizar.

## Gráfico 5

Ferramentas CAT utilizadas

### CAT tools utilizadas



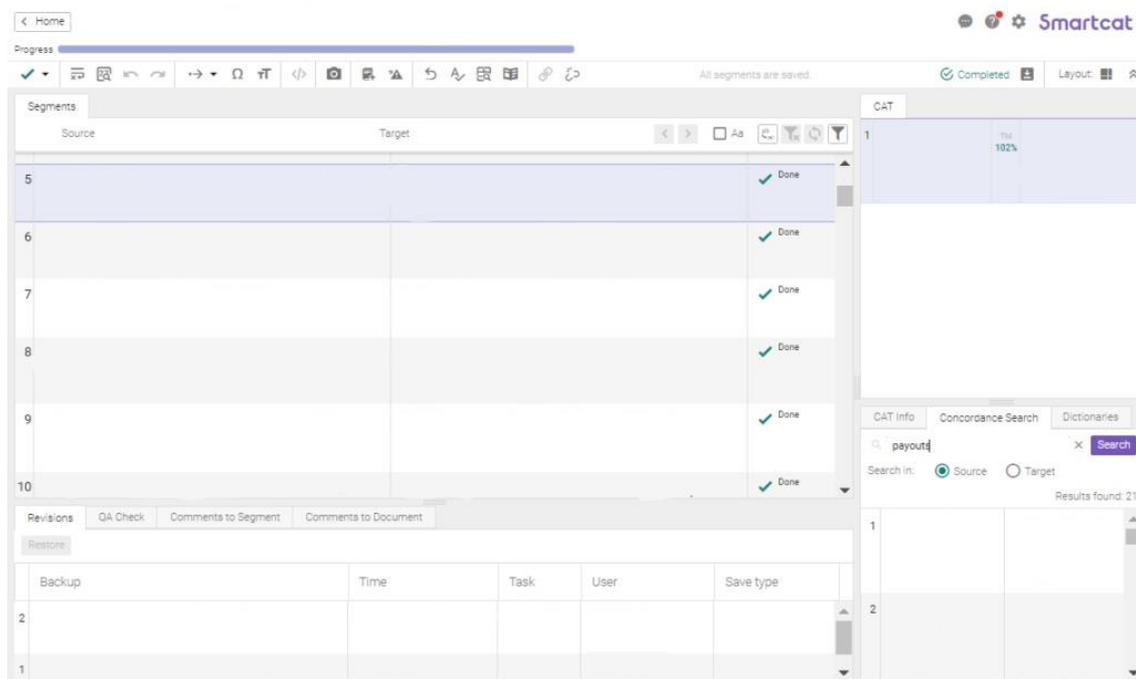
Houve uma maior utilização da ferramenta Smartling, seguida da ferramenta Smartcat e do memoQ. Esta diversidade de ferramentas exploradas foi extremamente benéfica, visto que, até este ponto, apenas duas das ferramentas acima tinham sido exploradas em contexto académico (memoQ e Smartcat). Uma característica comum a todas as ferramentas é o facto de estas serem ferramentas baseadas na nuvem, o que facilita o trabalho do tradutor, tendo este toda a documentação que precisa numa única página e acessível a partir de qualquer local com ligação à internet. Seguidamente, apresentamos uma breve descrição de cada uma das ferramentas, assim como da ferramenta da empresa, o Wordbee.

- Smartcat<sup>13</sup>: plataforma com uma interface bastante simples, o que facilita a sua utilização. Está integrada com inteligência artificial, o que permite agilizar a precisão e velocidade das traduções; a colaboração em tempo real permite que os linguistas, revisores e clientes trabalhem de uma forma eficaz e coordenada; facilita a gestão de projetos com os seus recursos intuitivos, de acompanhamento de tarefas, prazos e orçamentos; possui uma biblioteca de memórias de tradução onde armazena e reutiliza traduções anteriores de forma a economizar tempo e mantendo a consistência das traduções. Possui ainda uma vertente *marketplace* para a procura de linguistas e/ou clientes.

<sup>13</sup> As informações presentes neste ponto foram retiradas do website do Smartcat (<https://www.smartcat.com/why-smartcat/>)

**Figura 4**

*Editor do Smartcat*

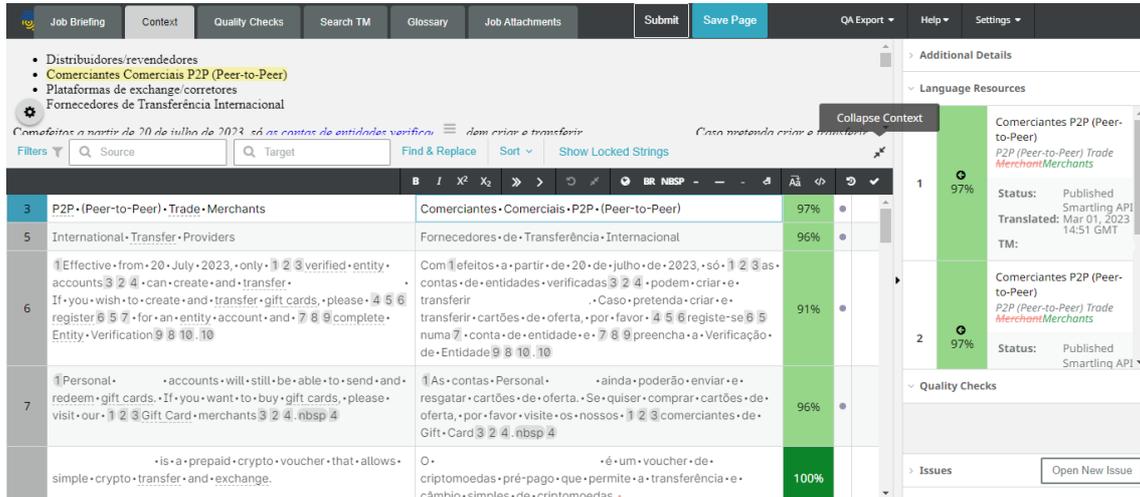


- Smartling<sup>14</sup>: plataforma com uma interface bastante simples e intuitiva, conseguindo aceder a todos os parâmetros a partir de uma única página. Tal como o Smartcat, está integrada com inteligência artificial, o que permite agilizar a precisão e velocidade das traduções. Permite ainda que o linguista consiga ver o seu produto final na própria página web antes de enviar ao cliente, algo que as outras ferramentas não disponibilizam. Possui uma biblioteca de memórias de tradução em tempo real, o que permite sincronizar e partilhar as memórias com os clientes instantaneamente. Duas das grandes vantagens são as verificações de qualidade automáticas, que aparecem à medida que o trabalho decorre, e a fácil manipulação de tags através de *machine learning*.

<sup>14</sup> As informações apresentadas foram retiradas do website do Smartling (<https://www.smartling.com/software/CAT-tool/>).

Figura 6

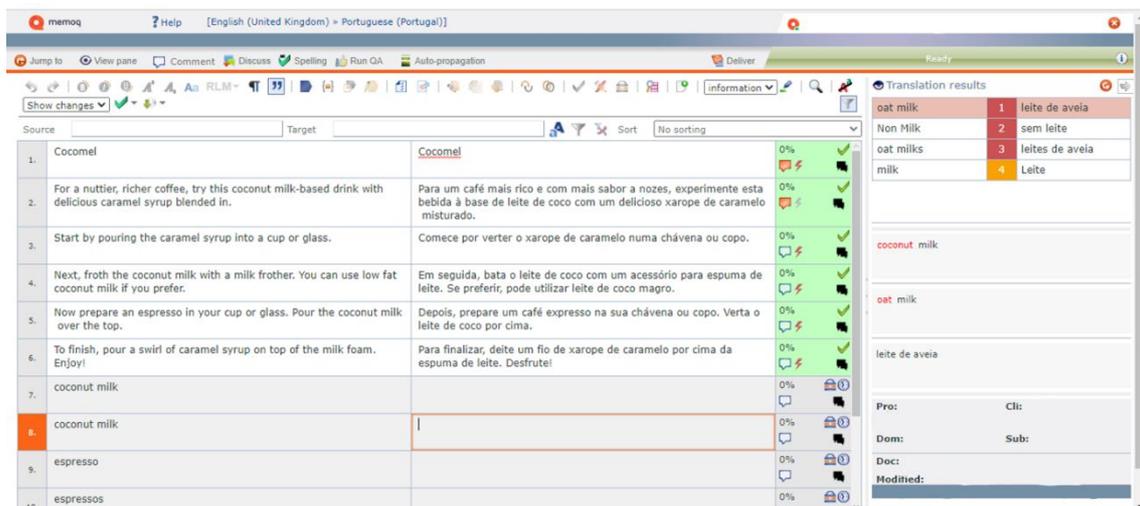
Editor do Smartling



• memoQ<sup>15</sup>: ambiente de tradução baseado no navegador, o que elimina a necessidade de o tradutor ter de instalar esta ferramenta. Esta versão do memoQ está apenas disponível para clientes que possuam um servidor ou licença na nuvem, tendo cada cliente um website específico para si. O ambiente de trabalho é bastante semelhante à versão desktop, pelo que foi de fácil adaptação. Uma vez criado o projeto no servidor do cliente, através de credenciais providenciadas ao PSL, é possível aceder ao projeto, traduzir, gerir projetos e utilizadores, podendo ainda criar novos projetos. Semelhante ao Smartling, possui uma biblioteca de memórias de tradução atualizada onde é possível consultar traduções passadas para o mesmo cliente.

Figura 5

Editor do memoQweb



<sup>15</sup> As informações foram retiradas do website do memoQ (<https://www.memoq.com/extensions#memoqweb>).

- Wordbee<sup>16</sup>: o Wordbee é a CAT tool onde todos os projetos são apresentados e recebidos, tendo sido esta a ferramenta de eleição da empresa. Sendo assim, todas as tarefas aparecem no *feed* pessoal de cada funcionário, mesmo que a tarefa a desempenhar não seja feita neste software. À semelhança das ferramentas mencionadas anteriormente, o Wordbee é uma ferramenta inteiramente online, permitindo que o tradutor possa trabalhar a partir de qualquer local. Dentro das ferramentas exploradas, esta é a mais completa, pois permite traduzir, rever, criar memórias de tradução e organizar documentos online, e permite uma gestão de projetos otimizada, pois é possível observarmos a referência do projeto; o tipo de projeto a realizar (tradução, pré-edição, pós-edição, revisão, etc.); língua de partida e língua de chegada; prazo de entrega; orientações (sítio onde geralmente indicam o link para a ferramenta CAT a ser utilizada e instruções do cliente, se aplicável); o domínio (tipo de texto) e material de referência. Esta gestão permite também uma maior organização do diário de bordo da estagiária, como pudemos ver anteriormente.

Esta diversidade de ferramentas permitiu que as competências na ótica do tradutor, nomeadamente a competência tecnológica, fossem exploradas ao seu máximo, tornando, assim, o tradutor mais versátil e mais qualificado para projetos futuros, uma vez que o leque de ferramentas CAT que domina é mais vasto, o que fará com que o seu valor e qualificações se destaquem no mercado de trabalho.

### **3.5 Análise de questões associadas a projetos**

Tendo trabalhado em diversas traduções e pós-edições PEMT com os mais variados temas, foram surgindo dificuldades ao longo do estágio. Nesta secção, serão apresentadas e analisadas algumas das questões mais recorrentes através de tabelas que contêm o texto de partida, o texto de chegada, criado por tradução automática (quando se tratar de um projeto PEMT), o texto de chegada produzido pela estagiária e a revisão realizada pela empresa. Estas questões serão divididas em diferentes categorias, consoante o problema observado: tradução literal, questões terminológicas, questões morfológicas e questões pragmáticas. A categorização destes problemas de tradução decorre das propostas de estratégias de tradução apresentadas por teóricos como Vinay e Darbelnet e também da divisão desenvolvida por Valente (2021), segundo estratégias propostas por Nord (1991). O facto de conseguirmos ter acesso à versão do revisor e às suas

---

<sup>16</sup> As informações da ferramenta foram retiradas do website da AP|Portugal (<https://www.apportugal.com/software-de-traducao/software-de-traducao-online/>).

correções revelou ser uma mais-valia ao longo do estágio, permitindo o reconhecimento de erros e impedindo a repetição dos mesmos.

### 3.5.1 Tradução literal

Para que o recetor de uma tradução tenha o mesmo tipo de experiência que um leitor da LP, muitas vezes o tradutor vê-se a recorrer a uma tradução o mais literal possível. A tradução literal é uma tradução que segue palavra por palavra o TP e que procura manter o seu significado (Molina & Hurtado Albir, 2002). Este tipo de tradução é mais comum quando as traduções são efetuadas entre línguas da mesma família e sobretudo da mesma cultura, sendo muitas vezes um bom aliado do tradutor.

Contudo, no exemplo que se segue (área da culinária), a tradução literal realizada pela tradução automática não se revelou benéfica para o tradutor. A solução passou por pesquisar, numa primeira fase, num glossário da empresa sobre culinária e na respetiva memória de tradução associada ao projeto os termos *flat white* e *milk frother*. Não encontrando solução para o termo *flat white*, procedeu-se à consulta de fontes de referência em português, ou seja, em espaços onde esta bebida à base de café pode ser habitualmente adquirida.

**Tabela 2**

*Erros de tradução literal com tradução automática*

<b>Texto de Partida</b>	<b>Tradução automática</b>	<b>Texto de Chegada<sup>17</sup></b>	<b>Comentário</b>
<b>Flat White</b>	<b>Branco Liso</b>	<b>Flat White</b>	Neste caso, a tradução automática traduziu literalmente o nome da bebida.
Next, froth the milk with a <b>milk frother</b> . <b>Tap the jug a few times on a hard surface</b> to remove large air bubbles before pouring the	De seguida, bata o leite com o <b>batedor de leite</b> . <b>Bata o jarro algumas vezes numa superfície dura</b> para remover as	De seguida, bata o leite com um <b>acessório para espuma de leite</b> . <b>Bata com o fundo do jarro algumas vezes numa</b>	Embora o termo <i>batedor de leite</i> exista, o mais correto e presente na memória de tradução era <i>acessório para espuma de leite</i> .

<sup>17</sup> Na Tabela 2, a coluna do Texto de Chegada realizado pela estagiária não sofreu alterações na fase de revisão.

milk slowly into the glass.	bolhas de ar grandes antes de verter o leite lentamente para o copo.	<b>superfície dura</b> para remover as bolhas de ar antes de verter o leite lentamente para o copo.	No segundo caso, para melhor compreensão do recetor da tradução, incluímos a indicação do fundo do jarro.
-----------------------------	--	---	---

Seguimos com mais dois exemplos, desta vez relacionados com Indústria e Tecnologia. Neste caso, em particular, verificamos que a tradução literal não deveria ter sido a estratégia adotada pela estagiária. Na primeira instância, verifica-se que deveríamos ter optado pela transposição, pois na revisão, verificamos mudanças ao nível das categorias gramaticais. Já no segundo exemplo, tendo em conta o comentário deixado pela revisora e visível na tabela abaixo, deveríamos ter optado pela utilização da modulação ou mesmo da expansão, de forma a tornar o texto mais natural e claro para o seu recetor.

**Tabela 3**

*Erros de tradução literal*

<b>Texto de Partida</b>	<b>Texto de Chegada</b>	<b>Revisão da empresa</b>	<b>Comentário</b>
<b>There's nothing to turn on</b> - you can work away happily knowing your project is safely backed up.	<b>Não há nada para ativar</b> - pode trabalhar com a certeza de que o seu projeto tem uma cópia de segurança segura.	<b>Não precisa de ativar qualquer opção</b> - pode trabalhar com a certeza de que o seu projeto tem uma cópia de segurança segura.	
Better tools like [X] make you better	Melhores ferramentas como o [X] fazem-no melhor	Melhore as suas competências com as melhores ferramentas, como o [X]	Podemos fugir um pouco ao original para tornar o texto mais natural e para facilitar a sua leitura e compreensão

### 3.5.2 Questões terminológicas

Ao longo do estágio, foram traduzidos textos da área da tecnologia, indústria, tecnologias da informação, finanças/criptomoedas, entre outras. Assim, evidenciamos que a linguagem especializada foi um aspeto bastante trabalhado e que levantou dúvidas e alguns problemas. Para a melhor tradução destes termos, recorreu-se a diversas pesquisas, consultando dicionários e *corpora* eletrónicos, artigos e ainda linguistas internos, de forma a garantir que a tradução alcançava a sua forma mais “perfeita”.

**Tabela 4**

*Erros de terminologia*

Texto de Partida	Texto de Chegada	Revisão da empresa	Comentário
1. Log in to your [X] account and click <b>[Wallets]</b> - <b>[Fiat and Spot]</b> - <b>[Withdraw]</b> .	1. Inicia sessão na tua conta [X] e clica em <b>[Carteiras]</b> - <b>[À Vista e Fiduciária]</b> - <b>[Levantar]</b> .	1. Inicia sessão na tua conta [X] e clica em <b>[Carteiras]</b> - <b>[Moeda fiduciária e À vista]</b> - <b>[Levantamento]</b> .	Correção do termo “Fiat”.
<b>min grid qty = max(minQty, minNotional/grid_1ower_limit)</b>	<b>qtd min grid = máx(qtd min, minNotional/limite_inferior_do_grid)</b>	<b>min grid qty = max(minQty, minNotional/grid_1ower_limit)</b>	Neste exemplo, ocorreu tradução das fórmulas de cálculo quando estas não deveriam ser traduzidas.
What is <b>Futures grid trading</b> ?	O que é o <b>trading de grid de futuros</b> ?	O que é o <b>grid trading de futuros</b> ?	Em traduções anteriores, este termo foi traduzido como “trading de grid”. Contudo, o glossário do cliente foi atualizado e o termo “grid trading” passou a ser mantido em

			inglês, algo em que a estagiária não reparou.
--	--	--	---

### 3.5.3 Questões morfológicas

A morfologia é a disciplina que descreve e analisa a estrutura interna das palavras e os processos morfológicos de variação e formação de palavras. (Morfologia - Dicionário Terminológico, 2023). Para analisar as palavras, a morfologia estuda detalhadamente a sua estrutura e as suas partes – radicais, prefixos, sufixos – e as relações com outras palavras. Este levantamento de questões morfológicas leva ao estudo dos procedimentos técnicos de tradução, especialmente a transposição, uma vez que este procedimento envolve a ocorrência de uma mudança gramatical na tradução. Neste aspeto não foram sentidas dificuldades, mas ocorreram pequenos problemas a nível morfológico ao longo do estágio, como a capitalização de expressões e se esta deve ser mantida na LC e ainda os plurais de siglas, questões estas que surgiam com frequência e levaram a mais pesquisa e perguntas a linguistas internos da AP | Portugal.

**Tabela 5**

*Erros morfológicos*

<b>Texto de Partida</b>	<b>Texto de Chegada</b>	<b>Revisão da empresa</b>	<b>Comentário</b>
<b>Customer Data Privacy</b> (2023)	<b>Privacidade de Dados dos clientes</b> (2023)	<b>Privacidade dos dados dos clientes</b> (2023)	Em português, evita-se o uso excessivo de maiúsculas, algo que não acontece com tanta frequência em inglês.
With this bot, traders can open an initial position (long or short) based on their analysis and simultaneously place buy-limit and sell-limit orders at	Com este bot, os traders podem abrir uma posição inicial (longa ou curta) com base na sua análise e, ao mesmo tempo, colocar ordens com limite de compra e	Com este bot, os traders podem abrir uma posição inicial (longa ou curta) com base na sua análise e, ao mesmo tempo, colocar ordens com limite de compra e	Apesar da grafia “pré-determinados” ser bastante comum, esta não existe em português, sendo a forma correta “predeterminado”.

<b>predetermined</b> intervals to capitalize on market volatility and ranging conditions.	ordens com limite de venda em intervalos <b>pré-determinados</b> para capitalizar a volatilidade do mercado e as condições de variação.	ordens com limite de venda em intervalos <b>predeterminados</b> para capitalizar a volatilidade do mercado e as condições de variação.	
Unite <b>PDFs</b> , images, spreadsheets, and more	Una <b>PDF</b> , imagens, folhas de cálculo e muito mais	Una <b>PDFs</b> , imagens, folhas de cálculo e muito mais	Ainda que as siglas em português não tenham plural, o cliente pede especificamente que neste caso se mantenha “PDFs”.

### 3.5.4 Questões pragmáticas

A pragmática pode ser definida como a análise das relações existentes entre as formas linguísticas, os interlocutores do texto e o contexto da comunicação, ou seja, os fatores condicionantes e determinantes do uso da linguagem. Muito resumidamente, a pragmática analisa o funcionamento significativo e comunicativo da linguagem e a intenção comunicativa de quem neles fala ou escreve (Pragmática - Dicionário Terminológico, 2023).

**Tabela 6**

*Erros relativos a questões pragmáticas*

<b>Texto de Partida</b>	<b>Texto de Chegada</b>	<b>Revisão da empresa</b>	<b>Comentário</b>
<b>Create</b> with [X]	<b>Crie</b> com o [X]	<b>Criar</b> com o [X]	Este segmento trata-se de um CTA (call-to-action), que normalmente assume a forma de um botão, cujo objetivo é levar o utilizador a realizar

			uma ação. Em português, os CTA geralmente vão sempre no infinitivo, não conjugando, assim, os verbos.
<b>Change</b> the network	<b>Mude</b> a rede	<b>Mudar</b> a rede	Nestes segmentos, como se tratava de títulos de ações e não de situações descritivas, o verbo permanece no infinitivo e não no imperativo.
<b>Enter Wi-Fi</b> password	<b>Insira</b> a palavra-passe	<b>Inserir</b> a palavra-passe <b>do Wi-Fi</b>	
<b>Select</b> [X]	<b>Selecione</b> [X]	<b>Selecionar</b> [X]	
<b>Combine</b> files and <b>organize</b> them with [X].	<b>Combina e organiza</b> ficheiros com o [X].	<b>Combine e organize</b> ficheiros com o [X].	Correção da forma de tratamento para um tom mais formal.
<b>Learn</b> from creators in real time	<b>Aprender</b> com os criadores em tempo real	<b>Aprenda</b> com os criadores em tempo real	
<b>Travel the world with</b> [X].	<b>Viaja o mundo com o</b> [X].	<b>Viaje pelo mundo com o</b> [X].	
<b>Scale your art with</b> [X]	<b>Escala a tua arte com o</b> [X]	<b>Escale a sua arte com o</b> [X]	
[X] <b>lets</b> you edit, create, sign, and share PDFs — and impress potential employers with your work.	O [X] <b>deixa-o</b> editar, criar, assinar e partilhar PDF — e impressione potenciais empregadores com o seu trabalho.	O [X] <b>permite-lhe</b> editar, criar, assinar e partilhar PDF — e impressionar potenciais empregadores com o seu trabalho.	

## **4 Estudo do corpus comparável**

Nesta secção, começaremos por ilustrar todo o processo de recolha e limpeza do corpus, começando por uma breve apresentação da área de especialidade, descrição e delimitação do corpus e, por fim, uma análise linguística multinível, o cálculo do seu tamanho, da sua densidade lexical e da diversidade lexical.

### **4.1 Workflow de criação do corpus**

À luz dos domínios explorados ao longo do estágio, este corpus é constituído por 209 textos dos projetos realizados. O objetivo é analisar como a linguística de corpus se interliga com a tradução técnica e quais são os desafios mais levantados por tal. Ou seja, perceber que classes de palavras e estilo de linguagem é o mais utilizado nesta área de especialidade.

Como mencionado supra, o nosso corpus foi criado a partir de 209 traduções realizadas ao longo do estágio dos diferentes domínios abrangidos pela tradução técnica. Seguiram-se as seguintes etapas:

- i. Extração dos textos de partida e traduções
- ii. Junção dos TP num único ficheiro .docx
- iii. Junção dos TC num único ficheiro .docx
- iv. Limpeza dos corpora (eliminação de tags e links)
- v. Upload de cada um dos documentos para a plataforma Sketch Engine

De forma a auxiliar a visualização mental deste workflow, no [Anexo IV](#) encontramos o diagrama que o representa.

### **4.2 Tipologia textual | Macroestrutura**

A análise de um texto depende sempre da perspetiva pela qual o iremos analisar. Enquanto estudantes de tradução, somos encorajados a analisar um texto tendo por base os níveis macro (aprofundamento do conhecimento das características de diferentes tipos de texto) e microestrutural (mecanismos linguísticos de coesão textual), para melhor entender como proceder à tradução de determinado produto linguístico.

Recorrendo ao Dicionário Terminológico (s.d.) adotado pelo ministério da Educação, podemos definir texto como “Os textos, para além das propriedades fundamentais da textualidade, apresentam estruturas verbais peculiares, semânticas e formais, e marcas pragmáticas que possibilitam a sua classificação em tipos ou géneros. As características dos tipos ou géneros constituem indicadores importantes para a produção e para a interpretação dos textos.”

Posto isto, os termos microestrutura e macroestrutura referem-se a perspetivas sobre determinadas estruturas ou sistemas. Através da etimologia dos prefixos (macro-, «grande» e micro- «pequeno»), podemos afirmar, segundo Demirgüneş (2017), que a macroestrutura de um texto é «a portrayal that addresses to the main sections of the text in the context of plot, narrative program, narrator», ou seja, permite identificar quem produziu o texto, quem é o destinatário(s), com que intuito foi produzido, onde e quando foi produzido, a identificação das secções presentes no texto e a relação entre elas. Infra, podemos observar uma comparação estrutural do mesmo texto e como este é semelhante entre ambas as línguas.

**Tabela 7**

*Semelhanças estruturais entre o mesmo tipo de texto*

<b>English (United Kingdom)</b>	<b>Portuguese (Portugal)</b>
Affogato	Affogato
Did someone say Italian dessert? With just four ingredients, this nutty and sweet espresso-based drink is as delicious as it is quick to make.	Alguém falou em sobremesa italiana? Com apenas quatro ingredientes, esta bebida doce à base de café expresso e com sabor a avelã é tão deliciosa quanto rápida de fazer.
Start by preparing an espresso.	Comece por preparar um café expresso.
Mix in the hazelnut liqueur with the espresso.	Misture o licor de avelã com o café expresso.
Next, place the vanilla ice cream into a coffee glass.	Em seguida, coloque o gelado de baunilha num copo de café.
Pour the coffee mixture over the ice cream and top with crushed hazelnuts	Verta a mistura de café sobre o gelado e cubra com as avelãs picadas
espresso	
espressos	
vanilla ice cream	
vanilla ice cream	
crushed hazelnut	
crushed hazelnuts	
hazelnut liqueur	
hazelnut liqueurs	

Por outro lado, a microestrutura de um texto é o conjunto formado pelas frases que integram a superfície textual, é aqui que se procede a uma análise linguística do texto, debruçando-se sobre os aspetos lexical, sintático e semântico-enunciativo. Em relação à microestrutura do nosso corpus, poderemos observá-la mais minuciosamente na secção 4.5. deste relatório.

### 4.3 Tamanho do corpus

Para a análise do tamanho dos corpora, recorreu-se a duas ferramentas – WordCounter<sup>18</sup> e Sketch Engine<sup>19</sup>. O WordCounter foi utilizado de forma a contabilizar as palavras totais e as palavras únicas. Já o Sketch Engine utilizou-se para verificar o tamanho total do corpus, assim como para realizar a sua análise linguística, a qual poderemos observar na secção 4.5 deste relatório.

**Figura 8**

*Tamanho do corpus EN*

Details	
Words	164,833
Unique Words	16,203
Characters	1,037,622
Sentences	8,258
Paragraphs	28,255
Reading Level ?	College Graduate
Reading Time ?	9 hrs 59 mins
Speaking Time ?	15 hrs 16 mins

**Figura 7**

*Tamanho do corpus PT*

Details	
Words	181,946
Unique Words	18,125
Characters	1,176,837
Sentences	8,110
Paragraphs	28,699
Reading Level ?	College Graduate
Reading Time ?	11 hrs 2 mins
Speaking Time ?	16 hrs 51 mins

**Figura 9**

*Tamanho real do corpus PT*

corpus PT		
MANAGE CORPUS		
MANAGE SUBCORPORA		
COMPARE CORPORA		
TEXT TYPE ANALYSIS		
<b>GENERAL INFO</b> Language: Portuguese CORPUS DESCRIPTION & BIBLIOGRAPHY TAGSET WORD SKETCH GRAMMAR TERM GRAMMAR	<b>COUNTS</b> Tokens 209,897 Words 176,386 Sentences 8,531 Documents 1	<b>TEXT TYPES</b> TEXT TYPE ANALYSIS <doc> (2) 1 File ID, doc.id 1 File name, doc.filename 1 <g> (0) 27,951 <s> (0) 8,531

<sup>18</sup> WordCounter: <https://wordcounter.net/>

<sup>19</sup> Sketch Engine: <https://sketchengine.eu/>

**Figura 10**

*Tamanho real do corpus EN*



Como podemos observar, o corpus português mostra-se superior ao corpus inglês por, aproximadamente, 10.38%. Isto acontece, pois, o português é uma língua mais complexa e com mais classes gramaticais do que o inglês. Posto isto, vamos apresentar a densidade lexical dos corpora.

A densidade lexical (DeL) classifica-se como o número de palavras lexicais (ou palavras de conteúdo) que um corpus apresenta, dividido pelo número de tokens, ou palavras totais. As palavras lexicais são as que dão ao texto um significado e fornecem informação sobre o que é o texto. As palavras lexicais a serem consideradas são os nomes, verbos, advérbios e adjetivos. Embora existam outras categorias de palavras, como as preposições, apenas as mencionadas supra nos permitem obter informações sobre o que está a ser comunicado. Infra, podemos observar a tabela de cálculo da DeL dos corpora. Para a realização deste cálculo, recorreu-se ao método de Ure (Santos et al, 2018), que consiste na divisão dos itens lexicais pelo total de palavras presentes no texto.

**Tabela 8**

*Densidade lexical dos corpora*

Corpus PT	Corpus EN
$\frac{94\,968}{181\,946} \times 100 = \sim 52,20\%$	$\frac{78\,259}{164\,833} \times 100 = \sim 47,48\%$

Como podemos verificar, o corpus português, além de apresentar maior DeL, apresenta também maior número de tokens e maior número de palavras lexicais. Tendo por base estes dados, podemos concluir que o corpus português possui uma maior DeL e maior percentagem de nomes, o que permite afirmar que a língua portuguesa descreve em maior pormenor o conteúdo dos nossos corpora do que o corpus inglês.

A diversidade lexical (DiL) indica o desenvolvimento linguístico associado à quantificação da variação de palavras utilizadas em determinada produção escrita, ou seja, quanto maior a variação de palavras, maior a diversidade. Isto traduz-se matematicamente para o quociente do número de *types* (palavras únicas) pelo número de tokens (palavras totais). Para a obtenção deste resultado, é apenas necessário multiplicar o resultado da divisão por 100. Abaixo, segue a expressão que determina a diversidade lexical de cada *corpus*.

**Tabela 9**

*Diversidade lexical dos corpora*

<b>Corpus PT</b>	<b>Corpus EN</b>
$\frac{18\ 125}{181\ 946} \times 100 = \sim 9,96\%$	$\frac{16\ 203}{164\ 833} \times 100 = \sim 9,83\%$

Embora a diferença entre a DiL dos corpora seja mínima, o corpus português apresenta um desenvolvimento linguístico maior o que pode sugerir que a língua portuguesa, no que toca ao domínio técnico, tem maior competência lexical por parte dos redatores e tradutores.

#### **4.4 Escolha da ferramenta Sketch Engine**

A ferramenta escolhida para analisar os nossos corpora foi o Sketch Engine, uma ferramenta bastante intuitiva para o utilizador que se destaca por ser um software de gestão de corpora e de análise textual, agindo como concordanciador. Um concordanciador é um programa que permite que o utilizador procure por palavras específicas no corpus, fornecendo listas exaustivas para as ocorrências da palavra em contexto, ou seja, é uma ferramenta computacional utilizada para reorganização e extrações de informações no corpus e para observação e interpretação de dados providenciando novas perspetivas para a análise linguística e para a tradução.

O Sketch Engine é uma ferramenta online paga, o que pode causar entrave a alguns utilizadores, mas também conta com um período experimental de um mês assim como protocolos com inúmeras instituições de ensino superior. Este concordanciador conta com a integração de um etiquetador morfossintático e também a integração de corpus de referência fundamental para elaborar listas de candidatos a termo. A plataforma disponibiliza ainda tagsets *part-of-speech* (POS) para o utilizador consultar e poder combinar de forma a obter os resultados que pretende.

#### **4.5 Análise linguística dos corpora**

Para a análise dos nossos corpora foi utilizada a ferramenta Sketch Engine recorrendo ao uso de Corpus Query Language (CQL). O uso de CQL permite pesquisar por padrões gramaticais complexos no corpus, tendo assim o utilizador mais controlo sobre os padrões que pesquisa.

Existem três grandes vantagens no uso de CQL: **a)** a pesquisa por construções gramaticais, **b)** a pesquisa por *gapped constructions* (por exemplo, a pesquisa da combinação “more [x] than” que gerará resultados como “more important than”) e **c)** a pesquisa por palavras que contêm uma *string* particular de caracteres (por exemplo, a pesquisa por palavras que terminam em *-ing*).

#### 4.5.1 Análise a nível terminológico

A análise a nível terminológico dividiu-se em três secções principais: extração da terminologia dos nossos corpora, análise a nível de classes de palavras e análise de colocações. Para analisar a terminologia, procedeu-se a uma extração terminológica dos termos mais recorrentes dos corpora. Primeiramente, serão apresentadas, sob forma de nuvem de palavras, as monopalavras mais recorrentes dos corpora. Para efeitos de privacidade de dados, foram eliminadas da lista palavras que mencionem entidades, pessoas e produtos. De forma a ser de mais fácil visualização, as monopalavras dos dois corpora serão expostas com recurso a uma nuvem de palavras, criada no website WordArt<sup>20</sup>. Infra, podemos observar as duas nuvens de palavras.

**Figura 11**

*Monopalavras EN*



<sup>20</sup> WordArt: <https://wordart.com/create>

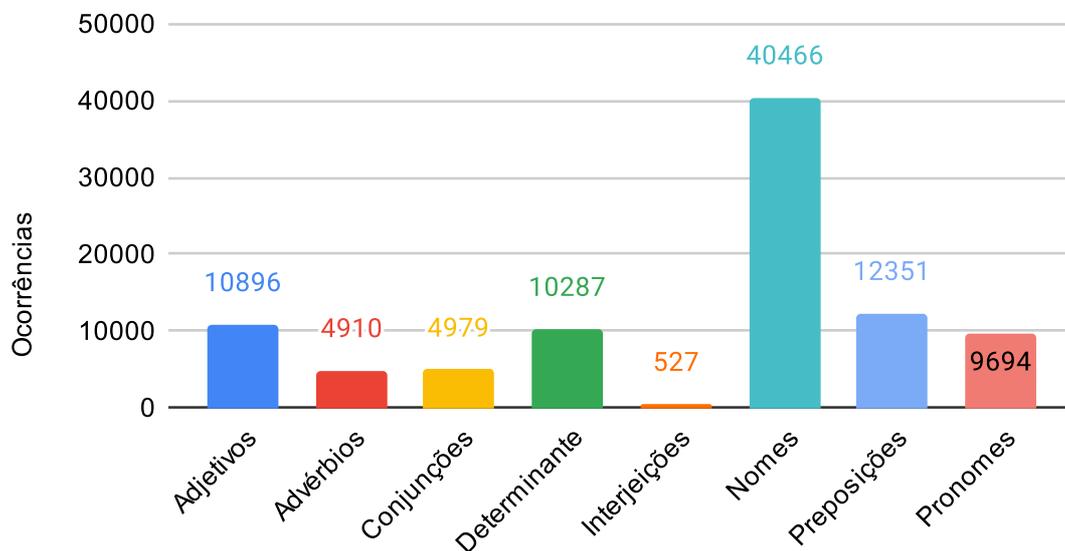




**Gráfico 6**

*Ocorrências de classes de palavras no corpus EN*

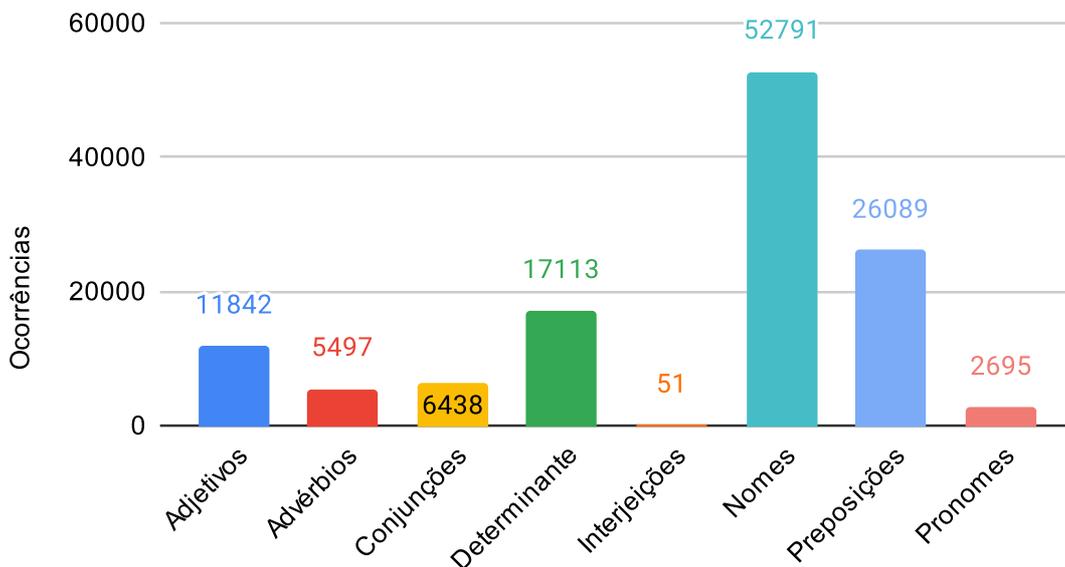
### Classes de palavras



**Gráfico 7**

*Ocorrências de classes de palavras no corpus PT*

### Classes de palavras



Como podemos verificar pelos gráficos acima, a distribuição de ocorrências por classe de palavras encontra-se bastante uniforme. Observamos que, em ambos os corpora, a classe dos nomes é a mais populosa, com 52 791 ocorrências no corpus português e 40 466 no corpus

inglês. Seguem-se as classes das preposições e dos determinantes no corpus português, com 26 089 e 17 113 ocorrências, respetivamente; e no corpus inglês seguem-se as preposições e os adjetivos, com 12 351 e 10 896 ocorrências, respetivamente. A discrepância mais notória entre os corpora é a classe das interjeições, que conta com 51 ocorrências no corpus português e com 527 ocorrências no corpus inglês.

Outro aspeto a considerar na deteção de padrões e na análise de corpora são as colocações. As colocações podem ser brevemente caracterizadas como «(...) a multiword unit or formulaic language that occurs with very high frequency and semantically or syntactically forms a meaningful unit» (Moon citado em Phoocharoensil, 2021, p. 241), ou seja, expressões linguísticas compostas por duas unidades lexicais. Já Iriarte (2000) afirma que as colocações podem ser divididas entre duas formas:

- a) combinações frequentes, prováveis, preferenciais ou usuais de palavras (nomeadamente substantivo + adjetivo e verbo + substantivo); e
- b) combinações de palavras aparentemente livres, geradas a partir das regras da língua, mas onde actua qualquer tipo de restrição lexical determinada pela norma. (pp. 175-176)

Posto isto, para a nossa análise, consideramos as seguintes colocações para os nossos corpora:

**Tabela 10**

*Colocações procuradas nos corpora*

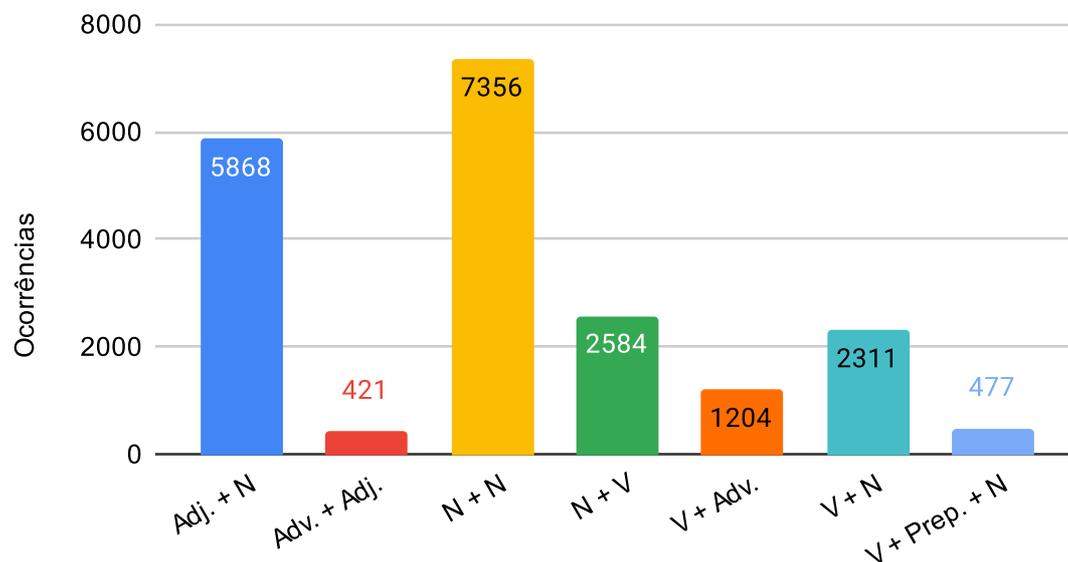
<b>Corpus PT</b>	<b>Corpus EN</b>
Adjetivo + Preposição + Nome	Adjetivo + Nome
Adjetivo + Advérbio	Advérbio + Adjetivo
Nome + Adjetivo	Nome + Nome
Nome + Preposição + Nome	Nome + Verbo
Verbo + Nome	Verbo + Advérbio
Nome + Verbo	Verbo + Nome
Verbo + Advérbio	Verbo + Preposição + Nome
Verbo + Preposição + Nome	

Abaixo, apresentamos a frequência e número de ocorrências de cada colocação em cada um dos corpora.

**Gráfico 9**

*Colocações presentes no corpus EN*

## Colocações

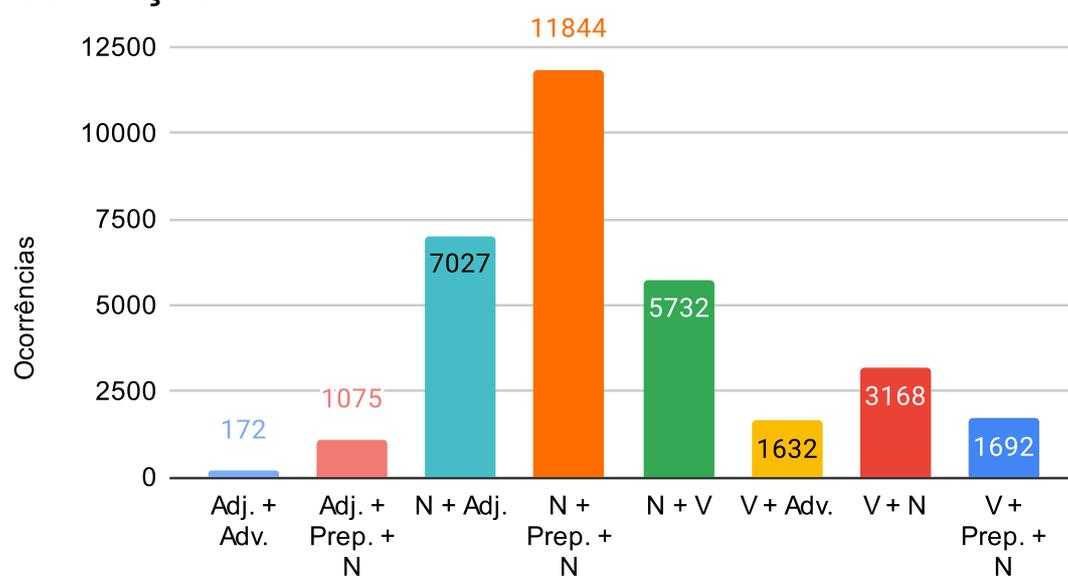


Como podemos observar, a colocação mais recorrente em inglês é a conjugação nome + nome, com 7 356 ocorrências, seguida de adjetivo + nome com 5 868 ocorrências e nome + verbo com 2 584 ocorrências.

**Gráfico 8**

*Colocações presentes no corpus PT*

## Colocações



Já no corpus português, a colocação mais recorrente é a combinatória nome + preposição + nome, com 11 844 ocorrências, seguido de nome + adjetivo com 7 027 ocorrências. Semelhante a inglês, a terceira combinação mais frequente é nome + verbo, com 5 732 ocorrências.

Do ponto de vista de análise de dados, podemos concluir que as colocações podem ser interpretadas como previsões empíricas acerca da vizinhança de uma palavra ou classe específica. Como observamos através de ambos os gráficos, em inglês é maior a probabilidade de um nome ser seguido de outro nome e, em português, um nome ser seguido de uma preposição e nome. No entanto, temos de ter em atenção que as colocações são imprevisíveis e, de domínio para domínio, obedecem a regras de utilização diferentes, tendo em conta o domínio de conhecimento do corpus. Assim, o número de ocorrências e a probabilidade de combinações no nosso corpus (área técnica) difere daquele da área jurídica, por exemplo.

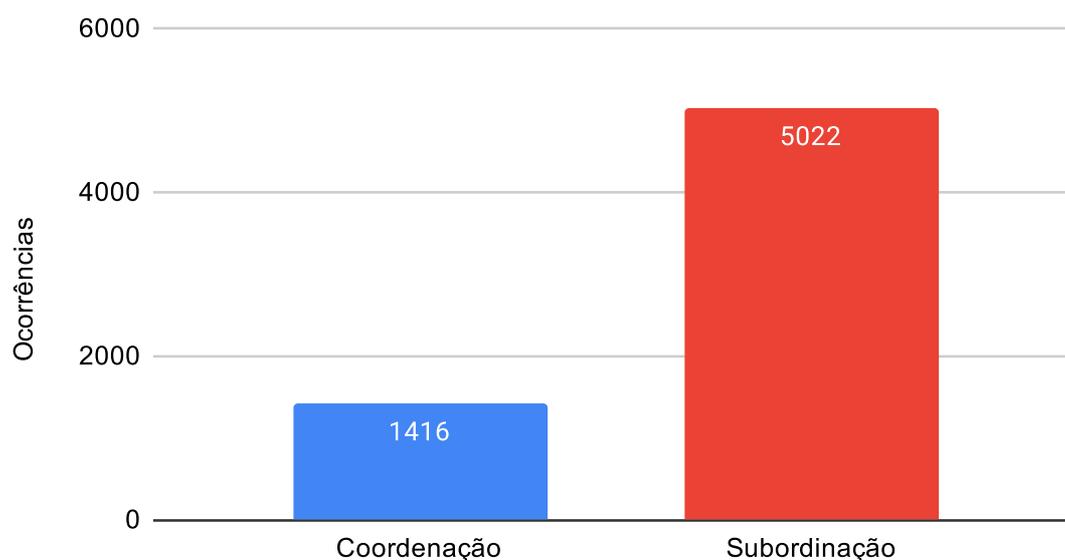
#### 4.5.2 Análise a nível morfossintático

Ao nível sintático, destacamos a parataxe, hipotaxe e a forma frásica. A parataxe caracteriza-se como o processo de articulação de frases por coordenação. Já a hipotaxe, é o processo de subordinação que ordena duas orações, estabelecendo a dependência de uma relativamente à outra. Posto isto, passemos a verificar a ocorrência de coordenação e subordinação.

#### Gráfico 10

*Hipotaxe vs. Parataxe no corpus PT*

### Hipotaxe vs. Parataxe



Através do Gráfico 10, conseguimos afirmar que a subordinação é predominante no corpus PT, contando com 5 022 ocorrências, em contraste com as 1 416 ocorrências de coordenação. Já no

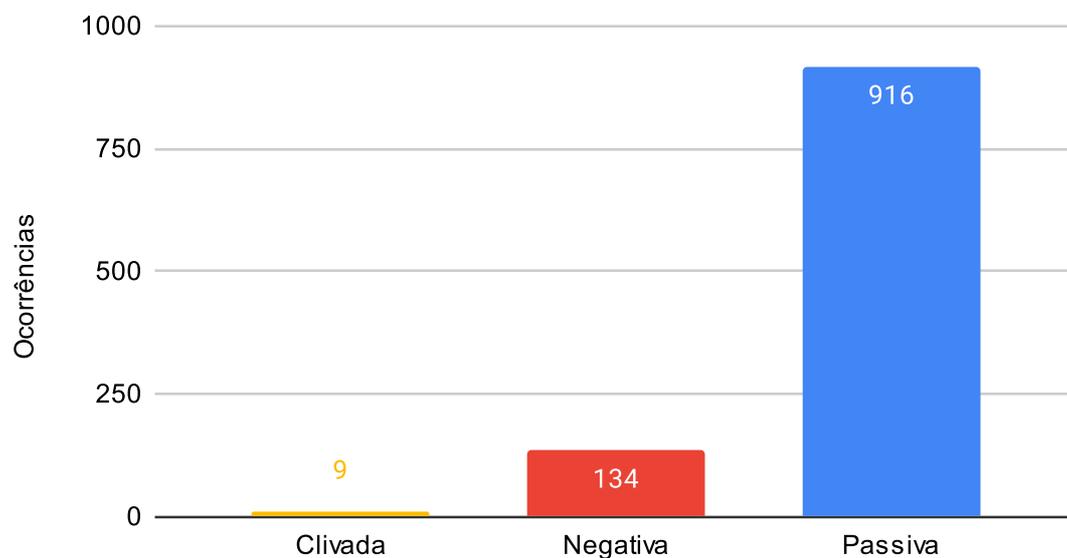
corpus EN, apenas foram verificadas ocorrências de coordenação, contando com 4 979 ocorrências.

A nível de formas frásicas, foram explorados três tipos: estrutura clivada, forma negativa e forma passiva. A estrutura clivada é uma construção que tem como objetivo colocar em destaque um constituinte da frase, sendo sempre construídas com o verbo SER e com um elemento qu- (wh-, no inglês) ou complementador (Bolrinha, 2017). A forma negativa é algo que exprime uma negação, tendo sido procurado nos corpora todas as ocorrências com a palavra “não”, e a forma passiva é uma construção frásica onde o sujeito é aquele que sofre a ação e não quem a pratica.

### Gráfico 11

*Formas frásicas presentes no corpus EN*

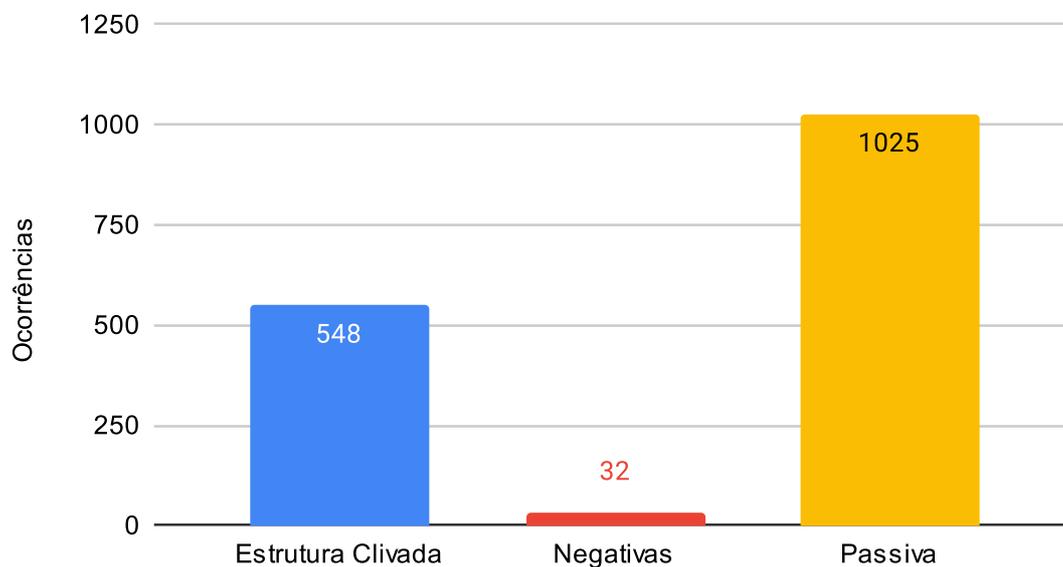
## Formas frásicas



**Gráfico 12**

*Formas frásicas presentes no corpus PT*

## Formas frásicas



Uma das diferenças mais notórias ao analisarmos comparativamente os gráficos 11 e 12 é a discrepância na utilização de estruturas clivadas, que no corpus EN conta com 9 ocorrências e no corpus PT conta com 548 ocorrências. Outra diferença bastante acentuada é a presença de frases negativas, que no corpus inglês conta com 134 ocorrências contra 32 ocorrências no corpus português. Por fim, também notamos uma diferença a nível da frase passiva, com 916 ocorrências no corpus EN e 1025 no corpus PT, o que se traduz num aumento de 11,90%

### **4.5.3 Análise a nível semântico-enunciativo**

A análise a nível semântico-enunciativo passará pela análise de todos os componentes verbais presentes nos corpora. Antes de nos debruçarmos nas categorias escolhidas, é importante definirmos o que caracteriza um verbo. Um verbo é um constituinte frásico que, normalmente, assinala eventos e ações, e que flexiona em tempo, aspeto, modo, pessoa, número e género, e que constitui o elemento principal do grupo verbal.

Para o corpus EN, baseamo-nos na lista de tipos de verbos providenciada no SIL Glossary of Linguistic Terms<sup>21</sup> e procuramos as seguintes categorias no nosso corpus: *defective verbs*; *finite*

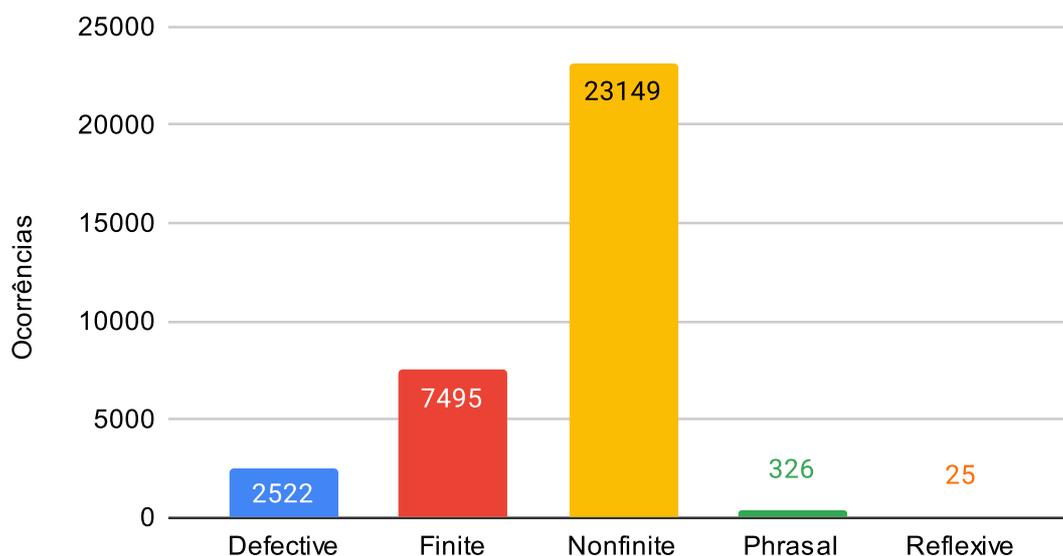
<sup>21</sup> Website do SIL Glossary of Linguistic Terms. Definição de verbo e apresentação dos tipos de verbos: SIL International. (2023). *What is a Verb (Linguistics)*. Obtido de SIL Glossary of Linguistic Terms: <https://glossary.sil.org/term/verb-linguistics>

*verbs*; *impersonal verbs*, e *nonfinite verbs*. Já para o corpus PT, exploraram-se os seguintes modos verbais: modo indicativo, modo conjuntivo, modo imperativo, gerúndio e infinitivo.

### Gráfico 13

Ocorrências de tipos verbais no corpus EN

## Tipos Verbais



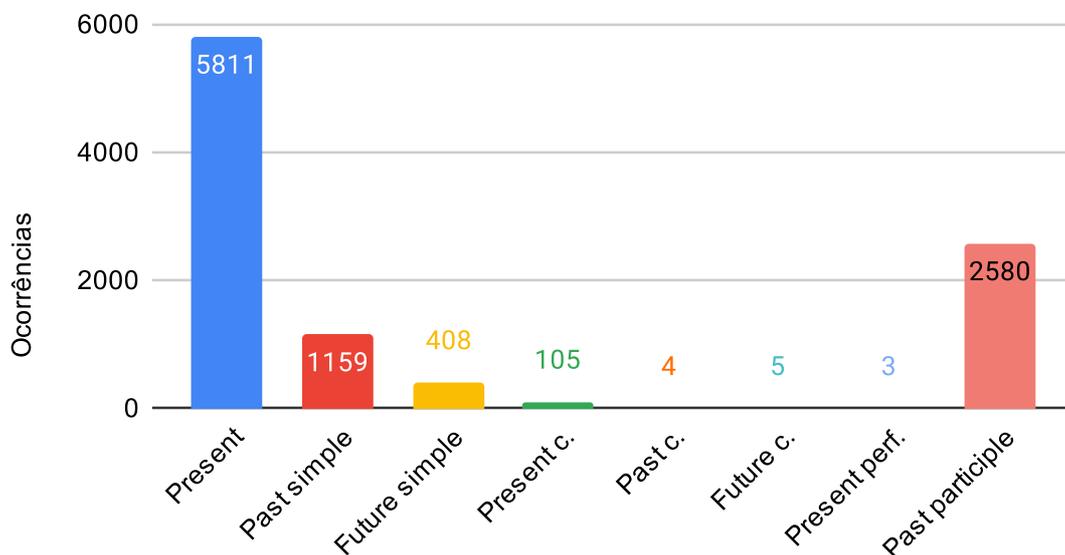
Como podemos observar pelo gráfico acima, os *nonfinite verbs* são os mais recorrentes no nosso corpus, contando com 23 149 ocorrências. Nesta categoria de formas não finitas, foram considerados as formas verbais no infinitivo, no gerúndio e no particípio. Dentro desta elevada quantia de formas verbais não finitas, 18 350 representam a forma no infinitivo. Considerando que estamos perante um corpus técnico, constituído por documentação como manuais de instruções, receitas, artigos de blogue, entre outros, é expectável que a linguagem imperativa seja rainha, visto ser mais direta e concisa para o recetor da informação. A linguagem imperativa parte do modo imperativo, contudo, os valores deônticos expressos por este modo podem ser traduzidos por outros processos linguísticos, como é o caso do infinitivo.

A segunda categoria mais frequente são os *finite verbs*. Esta categoria conta com 7 495 ocorrências e os verbos de forma finita são aqueles que têm flexão de tempo, pessoa e número, ou seja, foram contabilizados todos os tempos verbais, que se encontram expostos no gráfico abaixo.

**Gráfico 15**

*Tempos verbais presentes no corpus EN*

## Tempos Verbais

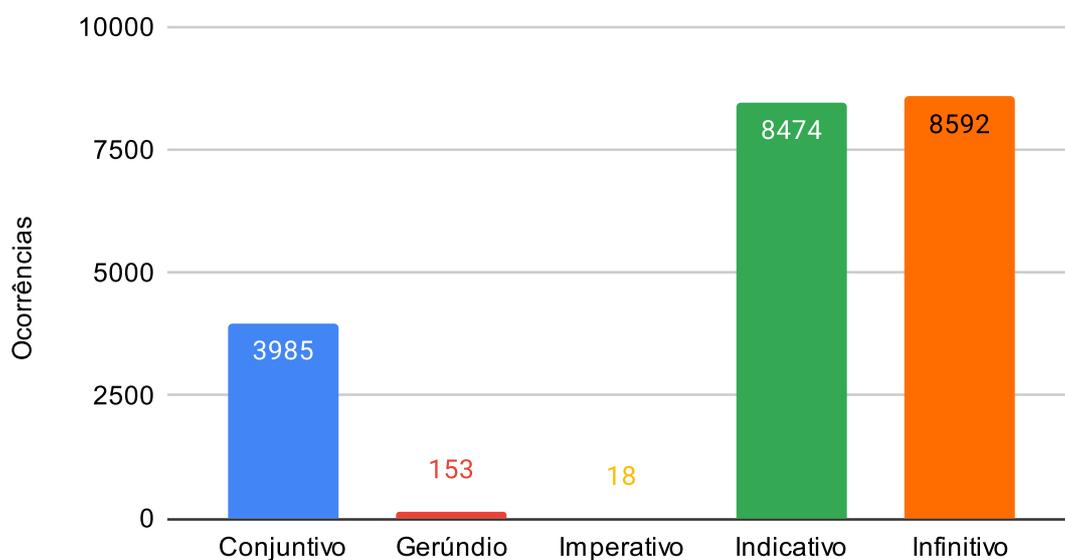


No presente gráfico, observamos uma enorme discrepância entre ocorrências de tempos verbais, sendo o presente o mais dominante, com 5 811 ocorrências, seguido do particípio passado (*past participle*) com 2 580.

**Gráfico 14**

*Ocorrências de modos verbais no corpus PT*

## Modos Verbais

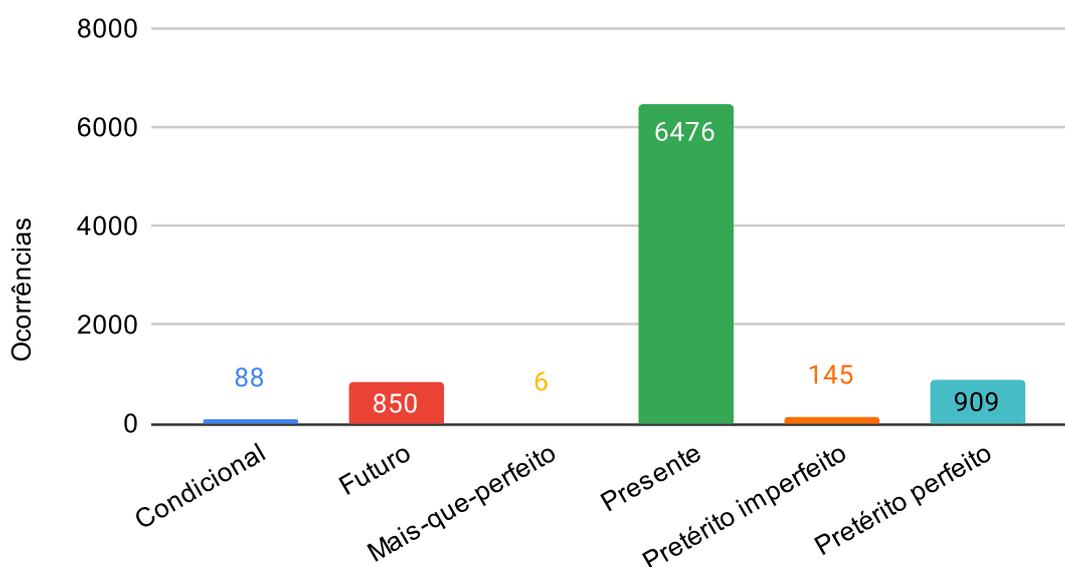


No corpus PT, temos dois modos verbais predominantes: os modos infinitivo e indicativo. O modo infinitivo conta com 8 592 ocorrências e, de forma semelhante ao corpus EN, era expectável esta predominância tendo em conta a nossa tipologia textual, o texto técnico. Já o modo indicativo, com 8 474 ocorrências, é utilizado de forma a refletir que o conteúdo descrito é atual, é corrente.

### Gráfico 16

*Ocorrências do modo Indicativo no corpus PT*

## Indicativo



Como o gráfico 16 constata, o presente do indicativo é o tempo verbal mais frequente, contando com 6 476 ocorrências. O presente do indicativo é normalmente associado à retratação de uma ação habitual que ocorre no momento da produção da frase. No entanto, como mencionado anteriormente, o presente do indicativo pode ser uma tradução do modo imperativo. O que observamos no corpus PT, é uma manifestação do modo imperativo através do infinitivo e do presente do indicativo, que transmitem a informação ao seu consumidor através de linguagem concisa e direta, características chave da linguagem técnica.

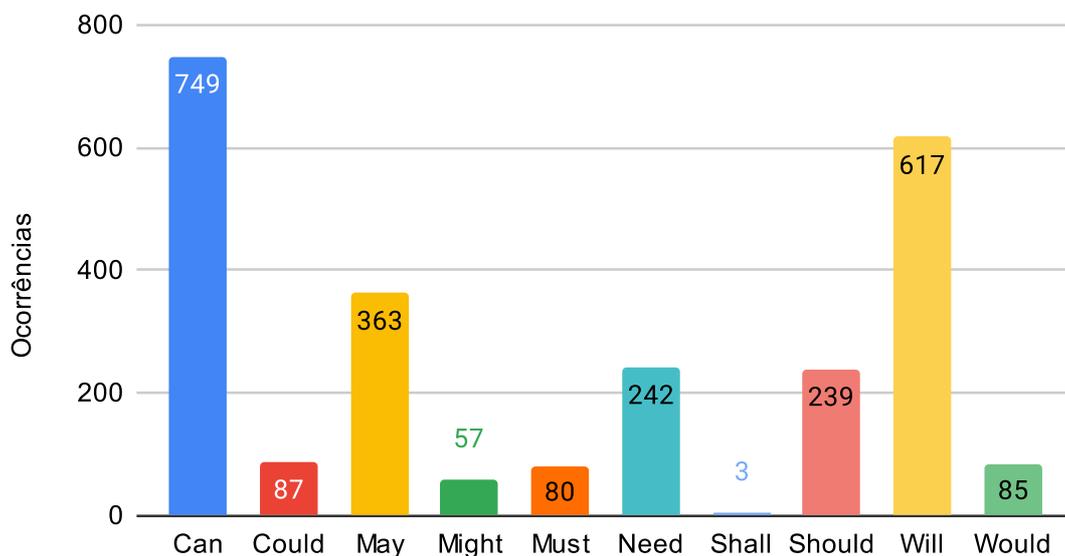
Passemos agora a analisar a presença de verbos modais nos nossos corpora. Um verbo modal é um verbo auxiliar que funciona como modificador e exprimem diferentes noções e valores, como desejo, possibilidade, probabilidade, dever, entre outros. Em português, a modalidade pode ser expressa através dos verbos modais dever, poder e ter de; de advérbios de frase como provavelmente, possivelmente e necessariamente; e ainda adjetivos como possível, provável, capaz. Já em inglês, a modalidade é expressa pelos verbos modais can, could, may, might, must, need, shall, should, will e would. O que encontramos em comum às duas línguas é o facto de a

modalidade ser expressa na forma verbo modal + infinitivo do verbo principal. Abaixo, observamos as ocorrências da modalidade em ambos os corpora.

### Gráfico 17

*Verbos modais presentes no corpus EN*

## Verbos Modais

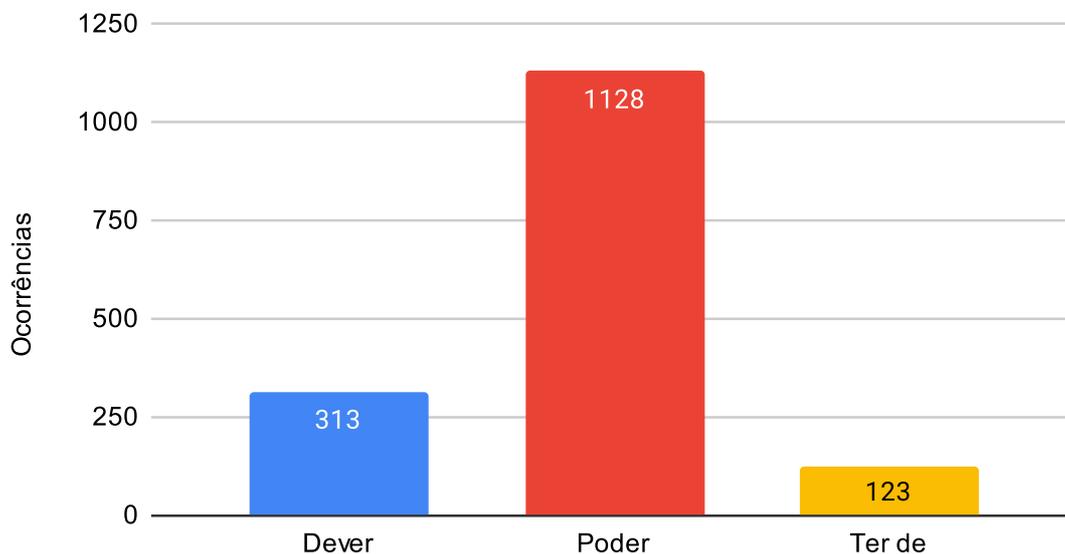


Como se observa, o verbo *can* é o mais recorrente no corpus, com 749 ocorrências, seguido dos verbos *will* e *may*, com 617 e 363 ocorrências, respetivamente. O verbo *can* demonstra capacidade, permissão e pedidos, ou seja, podemos concluir que estes são os valores mais presentes no nosso corpus, visto que são aqueles representados pelo verbo *can*. No que toca ao verbo *will*, este já representa as aceções correspondentes a hábitos e pedidos e o verbo *may* a possibilidade e permissão.

## Gráfico 18

Verbos modais presentes no corpus PT

### Verbos Modais

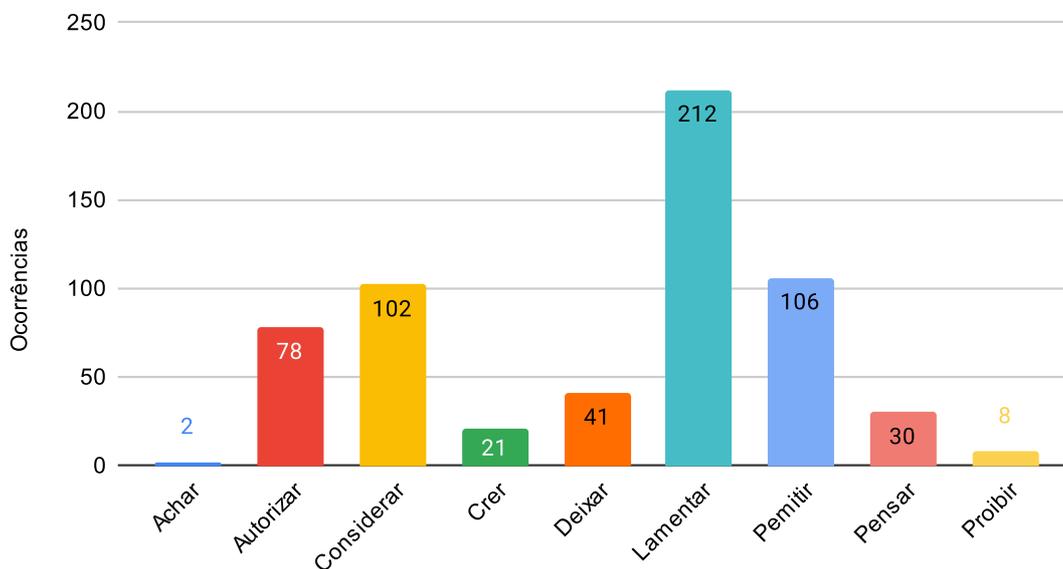


Em português, a modalidade verifica-se pela presença de verbos auxiliares modais, verbos principais com valor modal, advérbios de frase e adjetivos. Os verbos auxiliares modais em português são: *dever*, *poder*, *ter de* e *haver (de)*. No nosso corpus, só se verificou a presença de três destes verbos. O verbo com maior número de ocorrências é o verbo *poder*, que conta com 1 128 ocorrências, seguido de os verbos *dever* e *ter de*, com 313 e 123 ocorrências, respetivamente. De verbos principais com valor modal, foram seleccionados onze verbos de acordo com as modalidades epistémica (pensar, achar, crer, considerar), deôntica (autorizar, proibir, obrigar, permitir, deixar) e apreciativa (lamentar e alegrar-se).

## Gráfico 19

Verbos principais com valor modal no corpus PT

### Verbos principais modais



De acordo com o gráfico 19, os três verbos mais relevantes no nosso corpus são os verbos lamentar, permitir e considerar que contam com 212, 106 e 102 ocorrências, respectivamente. O verbo *lamentar* é um modal com valor apreciativo, ou seja, exprime avaliações positivas ou negativas sobre situações. Já o verbo *permitir* toma o valor deôntico, sendo utilizado para descrever situações que envolvam imposições ou autorizações, tomando o valor de obrigação ou permissão. Por fim, o verbo *considerar* representa a modalidade epistêmica, modalidade que exprime um juízo de valor acerca da verdade de uma proposição e toma valor de certeza ou valor de probabilidade.

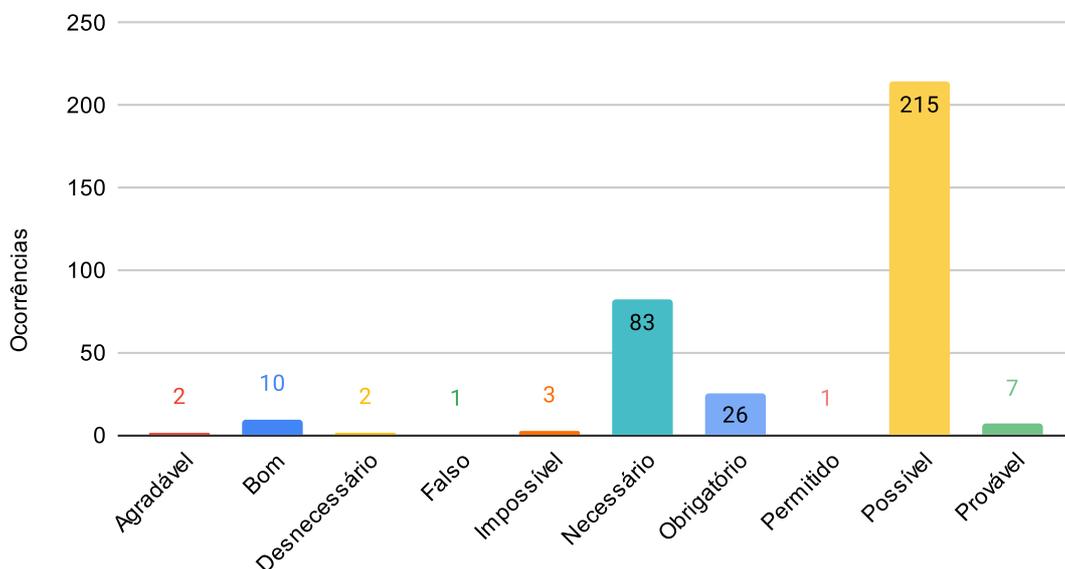
Outra forma de verificarmos a modalidade num texto, é através dos advérbios de frase, vulgo, advérbios com terminação em *-mente*. Fizemos uma pesquisa no nosso corpus por advérbios com esta terminação (através do uso da tag [lemma="\*. \*mente"]), contando assim com 637 ocorrências.

De adjetivos associados à modalidade, selecionamos quinze advérbios de acordo com as seguintes modalidades: epistêmica (possível, provável, impossível, improvável, desnecessário, verdadeiro, falso), deôntica (obrigatório, proibido, necessário) e apreciativa (bom, mau, agradável, desagradável, horrível).

## Gráfico 20

Adjetivos com valor modal presentes no corpus PT

### Adjetivos com valor modal



Como podemos observar, três dos adjetivos destacam-se, os adjetivos, *possível*, *necessário* e *obrigatório*, contando com 215, 83 e 26 ocorrências, respetivamente. O adjetivo modal mais presente no corpus (*possível*), encaixa-se na modalidade epistémica, representando valores de probabilidade ou certeza. De seguida, contamos com os adjetivos *necessário* e *obrigatório*, os representantes da modalidade deontica, podendo ser utilizados com valor de obrigação ou valor de permissão.

#### 4.6 Conclusões relativas aos corpora

Fazendo um apanhado sobre as características dos nossos corpora e retomando a nossa pergunta de investigação, podemos concluir que:

- A nível estrutural, verificamos que a macroestrutura textual é semelhante no TP e no TC, como observado pelo exemplo presente na secção 4.2.;
- Em relação ao tamanho dos corpora, verifica-se que o corpus PT é maior do que o EN. Tal fenómeno acontece porque o português pode ser considerado como uma língua mais complexa e com mais classes gramaticais. Isto é também verificável através da DeL e da DiL, que se manifestam como mais frequentes no corpus português;
- Ao nível terminológico, observamos que a multipalavra mais frequente nos corpora é a mesma; ainda a nível terminológico, desta vez olhando para as classes de palavras, verifica-se

um ligeiro aumento em todas as classes de palavras, com exceção da classe das interjeições e dos pronomes, onde o corpus inglês apresenta mais ocorrências;

- A nível morfosintático, novamente, o corpus português apresenta-se como sendo predominante, com mais ocorrências nas estruturas frásicas e na hipotaxe. A divergência a este nível encontra-se na parataxe, onde há mais ocorrências no corpus inglês;

- Por fim, a nível semântico-enunciativo. Em ambos os corpora, o tempo verbal mais utilizado foi o presente do indicativo. Partindo para a modalidade, verificamos que é no corpus inglês que esta se acentua mais. Contudo, comparando os corpora, verifica-se que a modalidade deôntica apresenta mais ocorrências a nível de verbos e a modalidade epistémica apresenta mais adjetivos. Isto acontece, pois, existe maior diversidade a nível modal na língua inglesa do que na portuguesa.

Como o papel da comunicação e linguagem técnica é o de informar, através de uma linguagem clara e simples, eram de esperar os resultados que obtivemos na nossa análise multinível. Ambos os corpora fazem mais uso do presente, sendo a forma mais simples e que todo o utilizador comum é capaz de interpretar, dando pouca margem para má interpretação. Verificado isto, é seguro afirmar que, embora os resultados apresentem discrepâncias algo significativas em certas ocorrências, há um padrão que se transporta do texto fonte para o texto alvo, as linguagens técnicas inglesa e portuguesa seguem a mesma lógica de raciocínio e de ocorrências.

## 5 Considerações finais

O presente relatório debruçou-se sobre o trabalho realizado durante o estágio curricular de seis meses na empresa AP|Portugal e a forma como o mesmo permitiu experienciar, em primeira mão, as dinâmicas de trabalho que ocorrem na área da tradução e dos serviços linguísticos. Esta experiência possibilitou a utilização dos conhecimentos teóricos e metodológicos da tradução adquiridos no contexto das unidades curriculares do nosso percurso académico, assim como pôr em prática as bases já conseguidas no domínio das ferramentas CAT, de competências na ótica do tradutor e no desenvolvimento das competências linguísticas.

Os seis meses de estágio foram, sobretudo, dedicados à tradução técnica, permitindo o desenvolvimento e aprimoramento de novas competências e a aquisição de novos conhecimentos na ótica do tradutor, tais como competência cultural, competência em pesquisa, aquisição e processamento da informação e competência técnica, assim como descritas no referencial de competências do EMT. Ao adquirir e reforçar o conjunto de competências de tradutora proveniente do contexto da formação académica, era, por isso, mais fácil prever que tipo de dificuldades poderia encontrar em determinado tipo textual, nomeadamente os textos técnicos. Associada a esta previsão, foram criados corpora com os textos trabalhados e procedeu-se à análise multinível dos mesmos, tendo em consideração os níveis terminológico, morfossintático e semântico-enunciativo. Esta análise permitiu evidenciar que, apesar de se tratar de línguas diferentes, conseguimos estabelecer relações de padronização, ou seja, é possível observar que as ocorrências linguísticas e a estrutura do nosso texto fonte transparecem no nosso texto alvo. Era de esperar que estas fossem as nossas conclusões, uma vez que o papel da comunicação e da linguagem técnica é o de informar o leitor, fazendo uso de uma linguagem simples, como observamos nas conclusões da secção 4.6.

Este estágio concedeu ainda à estagiária uma perceção do trabalho de um tradutor no mercado de trabalho atual, um mercado que se encontra em constante evolução, devido à constante revolução tecnológica e à introdução de novas tecnologias que podem auxiliar no seio da prestação de serviços linguísticos, como é o caso da utilização da Inteligência Artificial e do Processamento de Linguagem Natural para a análise da linguagem e para o auxílio durante o processo de tradução. O constante convívio com os diferentes departamentos da empresa proporcionou um melhor entendimento das etapas que um processo de tradução leva, assim como o desenvolvimento da competência interpessoal da estagiária, onde foi notada mais evolução. Este

mergulho no mundo “empresarial” da tradução permitiu descobrir a profissão de tradutor através de uma nova ótica para o que o futuro poderá reservar para os prestadores de serviços linguísticos.

## 6 Bibliografia

- AP|Portugal. (9 de janeiro de 2020). *Os 9 princípios da metodologia Kaizen*. AP|Portugal. <https://blog.apportugal.com/pt/os-9-principios-da-metodologia-kaizen>
- Baerman, M., Brown, Dunstan, & Corbett, G. G. (2009). *Surrey Cross-linguistic Database on Defectiveness*. <http://dx.doi.org/10.15126/SMG.21/2>
- Berber Sardinha, T. (2004). *Linguística de Corpus*. Barueri, SP: Editora Manole.
- Bolrinha, M. (2017). *Clivada de SER e Estrutura Enfática com Mas SER: Uma mesma estrutura ou estruturas diferentes?* [Dissertação de Mestrado, Universidade Nova de Lisboa]. Repositório da UNL. <http://hdl.handle.net/10362/25221>
- Byrne, J. (2006). *Technical translation: Usability strategies for translating technical documentation* (1ª ed.). Springer Dordrecht. <https://doi.org/10.1007/1-4020-4653-7>
- Byrne, J. (2014). *Scientific and Technical Translation Explained*. Londres: Routledge. <https://doi.org/10.4324/9781315760391>
- Cambridge Dictionary. (s.d.). *Finite and non-finite verbs*. Obtido de British Grammar: <https://dictionary.cambridge.org/grammar/british-grammar/finite-and-non-finite-verbs>
- Cambridge Dictionary. (s.d.). *Modal verbs and modality*. Obtido de British Grammar: <https://dictionary.cambridge.org/pt/gramatica/gramatica-britanica/modal-verbs-and-modality?q=Modal+verbs>
- Cambridge Dictionary. (s.d.). *Questions: interrogative pronouns (what, who)*. Obtido de British Grammar: <https://dictionary.cambridge.org/pt/gramatica/gramatica-britanica/questions-interrogative-pronouns-what-who>
- Cantante, I. (2020). Sobre a esclaridade dos adjetivos adverbiais modais. *Revista de Estudos Linguísticos da Universidade do Porto*, 15, 41-69. <https://ojs.letras.up.pt/index.php/EL/article/view/9475>
- Coelho, S. R. (2023). *A tradução técnica e a pós-edição: estágio curricular na RWS Portugal*. [Relatório de Estágio, Universidade do Minho]. Repositório da Universidade do Minho. <https://hdl.handle.net/1822/82958>
- Collins Dictionary. (s.d.). *REFLEXIVE VERB definition and meaning*. Obtido de Collins English Dictionary: <https://www.collinsdictionary.com/dictionary/english/reflexive-verb>
- Demirgüneş, S. (2017). Microstructural (Cohesion and Coherence) Text Generation Problems of Syrian Refugee Students Learning Turkish. *Universal Journal of Educational Research*, 5(4), 581-590. <http://dx.doi.org/10.13189/ujer.2017.050407>

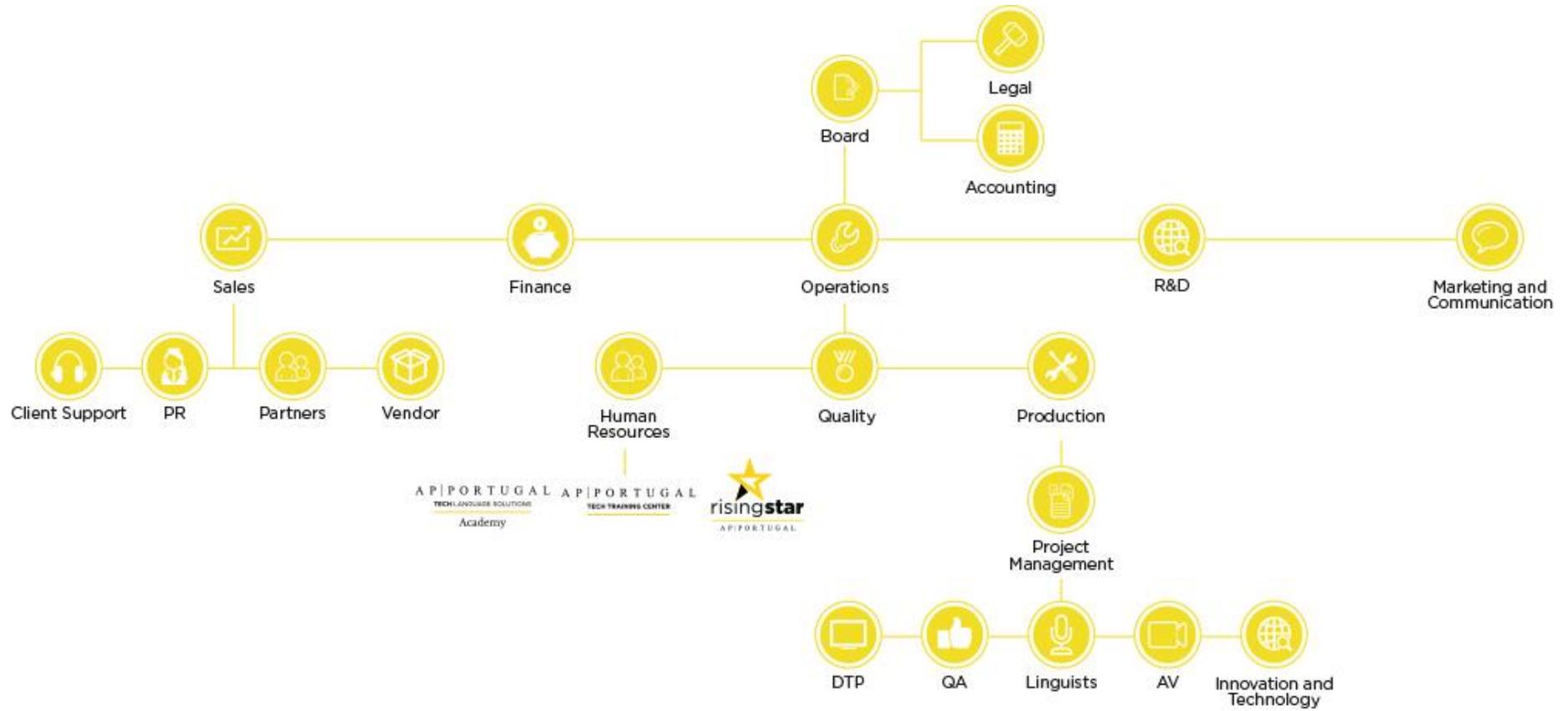
- Gouadec, D. (2007). *Translation as a Profession*. John Benjamins Publishing Company.  
<https://doi.org/10.1075/btl.73>
- House, J. (2014). *Translation Quality Assessment: Past and Present* (1ª ed.). Londres: Routledge.  
<https://doi.org/10.4324/9781315752839>
- Hurtado Albir, A. (2011). *Traducción y Traductología. Introducción a la Traductología*. Madrid: Cátedra.
- Instituto de Linguística Teórica e Computacional. (s.d.). *Dicionário de Termos Linguísticos*. Obtido de Portal da Língua Portuguesa: <http://www.portaldalinguaportuguesa.org/?action=terminology&act=view&id=3089>
- Iriarte Sanromán, Á. (2000). *A Unidade Lexicográfica. Palavras, Colocações, Frasemas, Pragmatemas*. [Tese de Doutoramento, Universidade do Minho]. Repositório da Universidade do Minho. <https://hdl.handle.net/1822/4573>
- Laviosa, S. (2010). Corpora. Em Y. Gambier, & L. van Doorslaer (Edits.), *Handbook of Translation Studies* (Vol. I, pp. 80-86). Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/hts.1>
- Leech, G. (2005). Adding Linguistic Annotation. Em M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice* (pp. 17-29). Oxford: Oxbow Books. <https://users.ox.ac.uk/~martinw/dlc/chapter2.htm>
- L'Homme, M.-C. (2020). *Lexical Semantics for Terminology: An introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.20>
- Milton, J., & Bandia, P. (Edits.). (2009). *Agents of Translation*. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/btl.81>
- Ministério da Educação e Ciência. (2023). *Morfologia - Dicionário Terminológico*. Obtido de Dicionário Terminológico: <https://dt.dge.mec.pt/>
- Ministério da Educação e Ciência. (2023). *Pragmática - Dicionário Terminológico*. Obtido de Dicionário Terminológico: <https://dt.dge.mec.pt/>
- Molina, L., & Hurtado Albir, A. (2002). Translation Techniques Revisited: A Dynamic and Functionalist Approach. *Meta*, 47(4), 498-512. <https://doi.org/10.7202/008033ar>
- Munday, J. (2016). *Introducing Translation Studies: Theories and Applications* (4ª ed.). Routledge. <https://doi.org/10.4324/9781315691862>

- Nord, C. (1997/2008). The Translation Brief as a Guideline for the Trainee Translator. *Ilha do Desterro*, 41-55.  
[https://www.researchgate.net/publication/49617959\\_Defining\\_translation\\_functions\\_The\\_translation\\_brief\\_as\\_a\\_guideline\\_for\\_the\\_trainee\\_translation](https://www.researchgate.net/publication/49617959_Defining_translation_functions_The_translation_brief_as_a_guideline_for_the_trainee_translation)
- Partington, A. (2011). Corpus Linguistics: what it is and what it can do. *Cultus: the Journal of Intercultural Mediation and Communication*, 4, 35-58.  
[https://www.cultusjournal.com/files/Archives/partington\\_p.pdf](https://www.cultusjournal.com/files/Archives/partington_p.pdf)
- Pereira, Y. d. (2023). *O conceito de agência aplicado à tradução especializada: estágio curricular na Wordattachment*. [Relatório de Estágio, Universidade do Minho]. Repositório da Universidade do Minho. <https://hdl.handle.net/1822/85369>
- Petraru, A.-M. (novembro de 2015). Translation and Agency - A Knowledge-Based Organisation. *International Conference Knowledge-Based Organization*, XX(2), 636-639. 10.1515/kbo-2015-0108
- Phoocharoensil, S. (2021). Semantic prosody and collocation: A corpus study of the near-synonyms persist and persevere. *Eurasian Journal of Applied Linguistics*, 7(1), 240-258.  
<https://doi.org/10.32601/ejal.911269>
- Reiss, K., & Vermeer, H. J. (2013). *Towards a General Theory of Translational Action. Skopos Theory Explained* (1ª ed.). (C. Nord, Trad.) Londres: Routledge.  
<https://doi.org/10.4324/9781315759715>
- Santos, E. d., Calil, E., Pereira, L. Á., & Coimbra, R. L. (2018). Diversidade e densidade lexical em textos escritos por alunos recém-alfabetizados: um estudo descritivo de produções individuais e em díades. *Calidoscópico*, 16(1), 25-32.  
<http://dx.doi.org/10.4013/cld.2018.161.03>
- Schäler, R. (2010). Localization and translation. Em Y. Gambier, & L. van Doorslaer (Edits.), *Handbook of Translation Studies* (Vol. I, pp. 209-214). Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/hts.1>
- Schubert, K. (2010). Technical Translation. Em Y. Gambier, & L. van Doorslaer (Edits.), *Handbook of Translation Studies* (Vol. I, pp. 350-355). Amsterdam/Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/hts.1>

- Shen, G.-r. (2010). Corpus-based Approaches to Translation Studies. *CROSS-CULTURAL COMMUNICATION*, *V*(4), 181-187.  
<http://dx.doi.org/10.3968/j.ccc.1923670020100604.010>
- SIL International. (2023). *What is a Verb (Linguistics)*. SIL Glossary of Linguistic Terms:  
<https://glossary.sil.org/term/verb-linguistics>
- Simeoni, D. (1995). Translating and Studying Translation: the View from the Agent. *Meta*, *40*(3), 445-460. <https://doi.org/10.7202/004146ar>
- Sousa, S. N. (janeiro de 2023). *As competências do tradutor: uma experiência de estágio na AP/Portugal*. [Relatório de Estágio, Universidade do Minho]. Repositório da Universidade do Minho. <https://hdl.handle.net/1822/82741>
- Xinyue, L. (2020). *Um estudo sobre os verbos auxiliares em português europeu*. [Dissertação de Mestrado, Universidade de Aveiro]. Repositório Institucional da Universidade de Aveiro. <http://hdl.handle.net/10773/30336>
- Zethsen, K. K. (1999). The Dogmas of the Technical Translation - Are They Still Valid? *HERMES - Journal of Language and Communication in Business*, *12*(23), 65-75.  
doi:<https://doi.org/10.7146/hjlc.v12i23.25549>

## 7 Anexos

### 7.1 Anexo I: Organograma da AP | Portugal



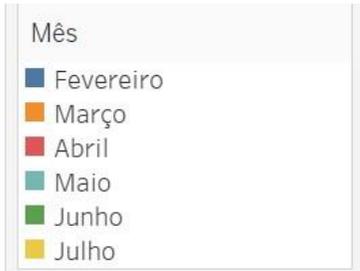
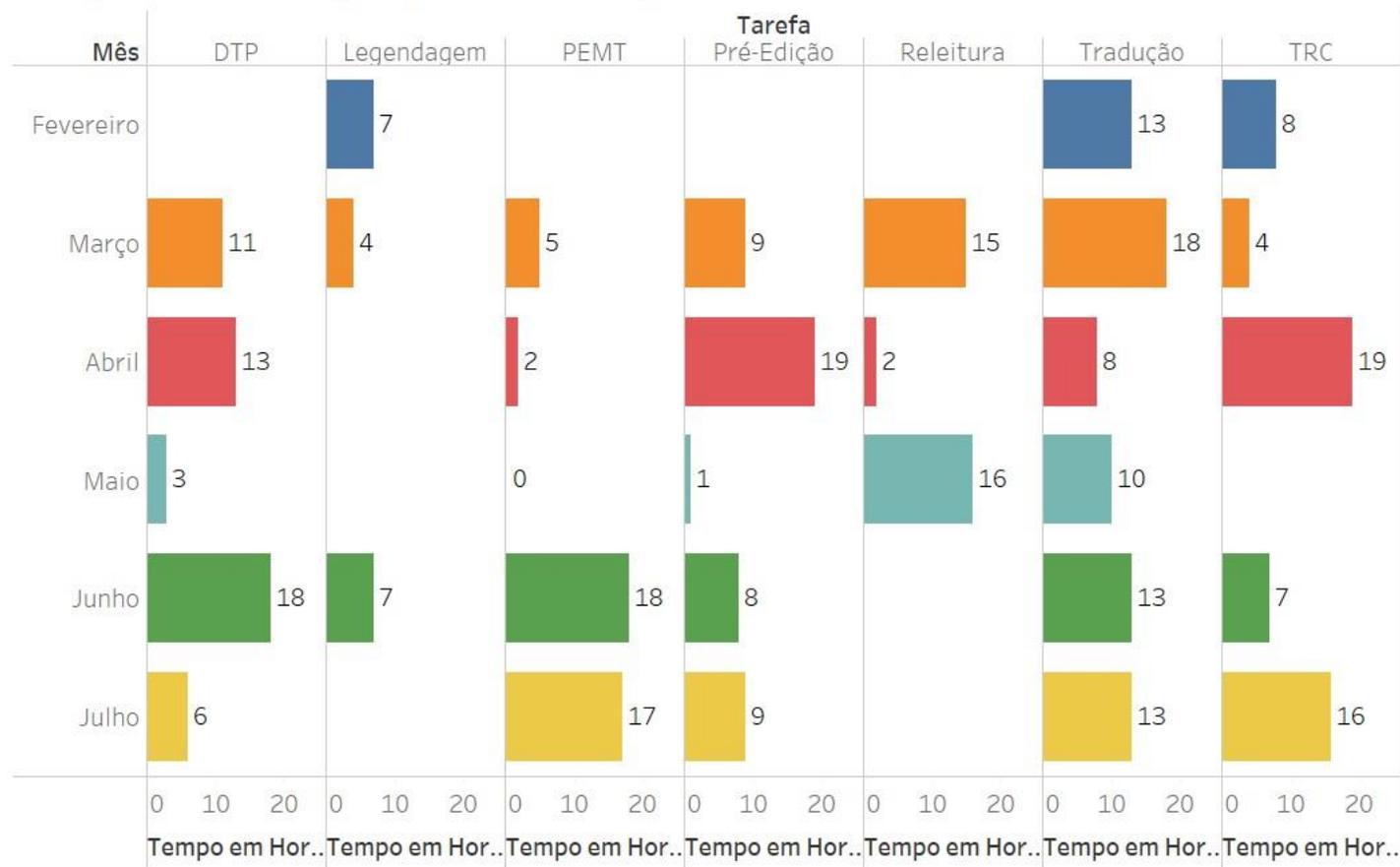
## 7.2 Anexo II: Exemplo do relatório mensal

Smartcat									
Total de Projetos	Ref. Projeto	Data	INTERNO / EXTERNO	Tradução	Revisão	Releitura	Outro	Duração da tarefa	Quantidade de palavras
1		02/05/2023	Externo	X				0:21:00	1625
2		02/05/2023	Externo				PEMT	1:08:00	2689
3		05/05/2023	Externo	X				0:44:00	1597
4		11/05/2023	Externo	X				0:34:41	1000
5		11/05/2023	Externo	X				0:07:00	7
6		12/05/2023	Externo	X				0:18:00	1100
7		12/05/2023	Externo	X				0:09:00	560
8		15/05/2023	Externo				(acompanh	0:14:00	763
9		17/05/2023	Externo	X				0:46:00	798
10		19/05/2023	Externo	X				0:47:00	1987
11		19/05/2023	Externo	X				2:14:05	19626
12		22/05/2023	Externo				PEMT	0:14:00	672
13		23/05/2023	Externo				PEMT	0:19:00	2457
14		24/05/2023	Externo	X				1:35:00	1553
15		29/05/2023	Externo				PEMT	0:05:00	114
16		29/05/2023	Externo				PEMT	0:02:00	1589
17		30/05/2023	Externo				PEMT	0:18:00	358
								9:55:46	38495

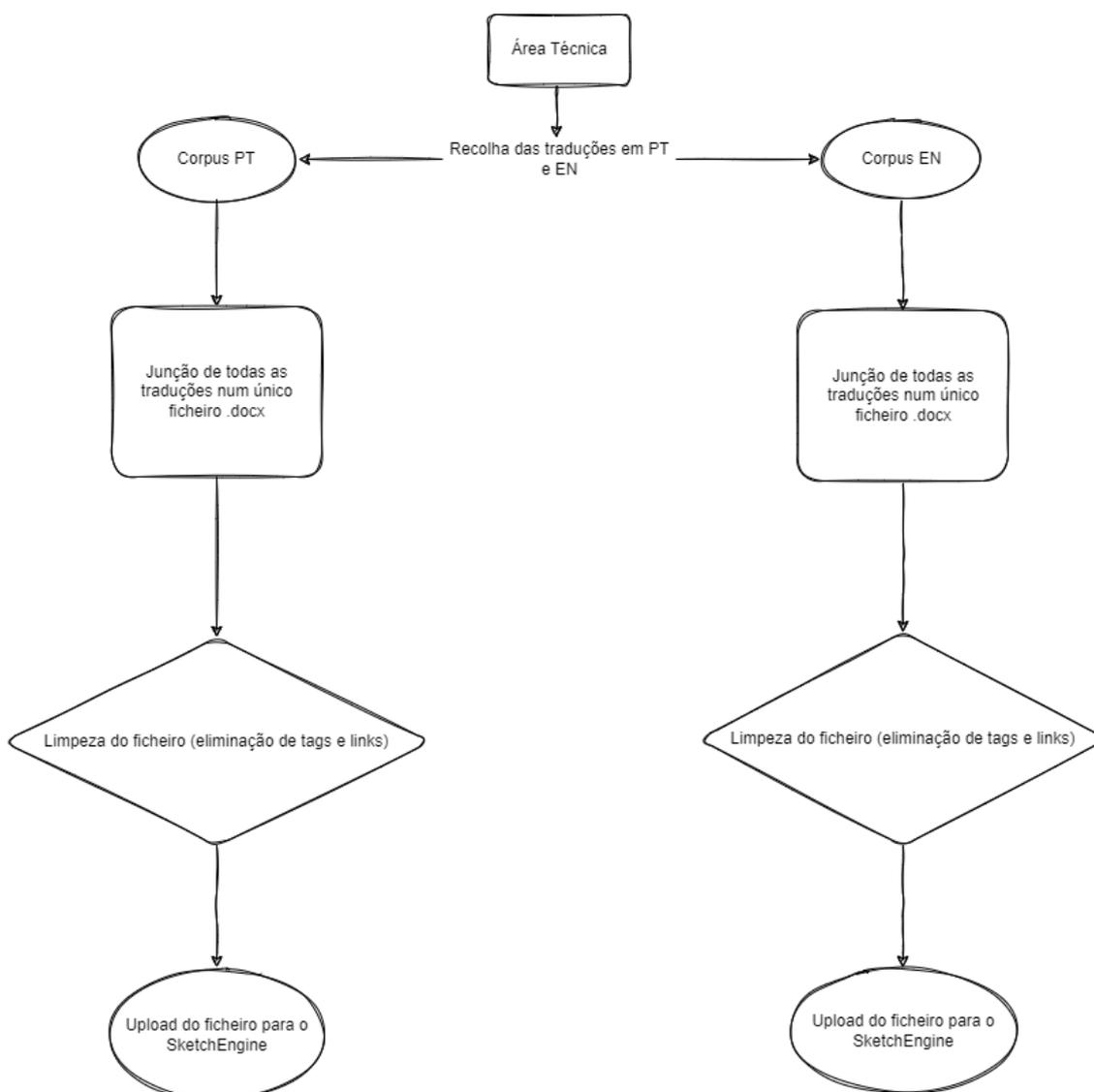
Smartling									
Total de Projetos	Ref. Projeto	Data	INTERNO / EXTERNO	Tradução	Revisão	Releitura	Outro	Duração da tarefa	Quantidade de palavras
1		02/05/2023	Externo				PEMT	0:02:00	8
2		02/05/2023	Externo				PEMT	0:02:00	42
3		08/05/2023	Externo				PEMT	0:19:00	265
4		08/05/2023	Externo				PEMT	0:12:00	131
5		09/05/2023	Externo				PEMT	0:20:00	292
6		09/05/2023	Externo				PEMT	0:13:00	140
7		09/05/2023	Externo				PEMT	0:41:00	519
8		10/05/2023	Externo				PEMT	0:01:00	20
9		11/05/2023	Externo				PEMT	0:22:00	118
10		15/05/2023	Externo				PEMT	0:09:00	229
11		15/05/2023	Externo				PEMT	0:06:00	80
12		16/05/2023	Externo				PEMT	0:08:00	73
13		17/05/2023	Externo				PEMT	0:02:00	9
14		17/05/2023	Externo				PEMT	0:42:00	697
15		18/05/2023	Externo				PEMT	0:04:00	107
16		23/05/2023	Externo				PEMT	0:17:00	366
17		23/05/2023	Externo				PEMT	0:07:00	115
18		31/05/2023	Externo				PEMT	0:09:00	522
								3:56:00	3733

**7.3 Anexo III:** Tempo gasto, em horas, por tarefa e por mês

**Tempo de realização por tarefa e por mês**



## 7.4 Anexo IV: Representação visual do workflow de criação dos corpora



## 7.5 Anexo V: Lista de POS tags para pesquisa nos corpora

Português	
Adjetivos	[tag="A.*"]
Adjetivos com valor modal	
<ul style="list-style-type: none"> <li>• Agradável</li> <li>• Bom</li> <li>• Desnecessário</li> <li>• Falso</li> <li>• Impossível</li> <li>• Necessário</li> <li>• Obrigatório</li> <li>• Permitido</li> <li>• Possível</li> <li>• Provável</li> </ul>	<ul style="list-style-type: none"> <li>• [word="agradável"]</li> <li>• [word="bom"]</li> <li>• [word="desnecessário"]</li> <li>• [word="falso"]</li> <li>• [word="impossível"]</li> <li>• [word="necessário"]</li> <li>• [word="obrigatório"]</li> <li>• [word="permitido"]</li> <li>• [word="possível"]</li> <li>• [word="provável"]</li> </ul>
Advérbios	[tag="R.*"]
<ul style="list-style-type: none"> <li>• Advérbios modais</li> </ul>	<ul style="list-style-type: none"> <li>• [lemma="*.mente"]</li> </ul>

Colocações	<ul style="list-style-type: none"> <li>• Adjetivo + advérbio</li> <li>• Adjetivo + preposição + nome</li> <li>• Nome + adjetivo</li> <li>• Nome + preposição + nome</li> <li>• Nome + verbo</li> <li>• Verbo + advérbio</li> <li>• Verbo + nome</li> <li>• Verbo + preposição + nome</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="A.*"][tag="R.*"]</li> <li>• [tag="A.*"][tag="SP.*"][tag="N.*"]</li> <li>• [tag="N.*"][tag="A.*"]</li> <li>• [tag="N.*"][tag="SP.*"][tag="N.*"]</li> <li>• [tag="N.*"][tag="V.*"]</li> <li>• [tag="V.*"][tag="R.*"]</li> <li>• [tag="V.*"][tag="N.*"]</li> <li>• [tag="V.*"][tag="SP.*"][tag="N.*"]</li> </ul>
Conjunções	<ul style="list-style-type: none"> <li>• Coordenadas</li> <li>• Subordinadas</li> </ul>	[tag="C.*"] <ul style="list-style-type: none"> <li>• [tag="CC.*"]</li> <li>• [tag="CS.*"]</li> </ul>
Determinantes	<ul style="list-style-type: none"> <li>• Determinante artigo</li> <li>• Determinante demonstrativo</li> <li>• Determinante indefinido</li> <li>• Determinante possessivo</li> </ul>	[tag="D.*"] <ul style="list-style-type: none"> <li>• [tag="DA.*"]</li> <li>• [tag="DD.*"]</li> <li>• [tag="DI.*"]</li> <li>• [tag="DP.*"]</li> </ul>
Estrutura clivada		[lemma="ser"][] [word="que"]
Frases negativas		[word="não"]
Interjeições		[tag="I.*"]
Modos verbais	<ul style="list-style-type: none"> <li>• Conjuntivo</li> <li>• Gerúndio</li> <li>• Imperativo</li> <li>• Indicativo</li> <li>• Infinitivo</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="VMS.*"]</li> <li>• [tag="VMG.*"]</li> <li>• [tag="VMM.*"]</li> <li>• [tag="VMI.*"]</li> <li>• [tag="VMN.*"]</li> </ul>
Modo Indicativo	<ul style="list-style-type: none"> <li>• Condicional</li> <li>• Futuro</li> <li>• Mais-que-perfeito</li> <li>• Presente</li> <li>• Pretérito imperfeito</li> <li>• Pretérito perfeito</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="VMIC.*"]</li> <li>• [tag="VMIF.*"]</li> <li>• [tag="VMIM.*"]</li> <li>• [tag="VMIP.*"]</li> <li>• [tag="VMII.*"]</li> <li>• [tag="VMIS.*"]</li> </ul>
Nomes	<ul style="list-style-type: none"> <li>• Nomes comuns</li> <li>• Nomes próprios</li> </ul>	[tag="N.*"] <ul style="list-style-type: none"> <li>• [tag="NC.*"]</li> <li>• [tag="NP.*"]</li> </ul>
Preposições		[tag="SP.* "]
Pronomes	<ul style="list-style-type: none"> <li>• Pronome demonstrativo</li> <li>• Pronome exclamativo</li> <li>• Pronome indefinido</li> <li>• Pronome interrogativo</li> <li>• Pronome pessoal</li> <li>• Pronome relativo</li> </ul>	[tag="P.* "] <ul style="list-style-type: none"> <li>• [tag="PD.*"]</li> <li>• [tag="PE.*"]</li> <li>• [tag="PI.*"]</li> <li>• [tag="PT.*"]</li> <li>• [tag="PP.*"]</li> <li>• [tag="PR.*"]</li> </ul>
Verbos		[tag="V.*"]
Verbos modais	<ul style="list-style-type: none"> <li>• Achar</li> <li>• Autorizar</li> </ul>	<ul style="list-style-type: none"> <li>• [lemma="achar"]</li> <li>• [lemma="autorizar"]</li> </ul>

<ul style="list-style-type: none"> <li>• Considerar</li> <li>• Crer</li> <li>• Deixar</li> <li>• Dever</li> <li>• Lamentar</li> <li>• Pensar</li> <li>• Permitir</li> <li>• Poder</li> <li>• Proibir</li> <li>• Ter de</li> </ul>	<ul style="list-style-type: none"> <li>• [lemma="considerar"]</li> <li>• [lemma="crer"]</li> <li>• [lemma="deixar"]</li> <li>• [lemma="dever"]</li> <li>• [lemma="lamentar"]</li> <li>• [lemma="pensar"]</li> <li>• [lemma="permitir"]</li> <li>• [lemma="poder"]</li> <li>• [lemma="proibir"]</li> <li>• [lemma="ter"][word="de"]</li> </ul>
Voz passiva	[lemma="ser"][tag="VMP.*"]

<b>Inglês</b>	
Adjetivos	[tag="JJ.*"]
<ul style="list-style-type: none"> <li>• Comparativo</li> <li>• Superlativo</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="JJR.*"]</li> <li>• [tag="JJS.*"]</li> </ul>
Advérbios	[tag="RB.*"]
<ul style="list-style-type: none"> <li>• Comparativos</li> <li>• Superlativos</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="RBR.*"]</li> <li>• [tag="RBS.*"]</li> </ul>
Colocações	
<ul style="list-style-type: none"> <li>• Adjetivo + nome</li> <li>• Advérbio + adjetivo</li> <li>• Nome + nome</li> <li>• Nome + verbo</li> <li>• Verbo + advérbio</li> <li>• Verbo + nome</li> <li>• Verbo + preposição + nome</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="JJ.*"][tag="NN.*"]</li> <li>• [tag="RB.*"][tag="JJ.*"]</li> <li>• [tag="NN.*"][tag="NN.*"]</li> <li>• [tag="NN.*"][tag="VV.*"]</li> <li>• [tag="VV.*"][tag="RB.*"]</li> <li>• [tag="VV.*"][tag="NN.*"]</li> <li>• [tag="VV.*"][tag="IN.*"][tag="NN.*"]</li> </ul>
Conjunções	
<ul style="list-style-type: none"> <li>• Coordenadas</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="CC.*"]</li> </ul>
Determinantes	[tag="DT.*"]
Estrutura clivada	[lemma="it"][lemma="be"][] [word="that"][]
Frases negativas	[lemma="no"]   [lemma="n't"]
Interjeições	[tag="UH.*"]
Tipos verbais	
<ul style="list-style-type: none"> <li>• Defective (modal verbs)</li> <li>• Finite</li> <li>• <b>Nonfinite (infinitive verbs, gerund verbs, participle verbs)</b></li> <li>• Phrasal</li> <li>• Reflexive</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="MD.*"]</li> <li>• [tag="VVP.*"]; [tag="VVZ.*"]; [tag="VHP.*"]; [tag="VHZ.*"]; [tag="VBP.*"]; [tag="VBZ.*"]; [tag="VVD.*"]; [tag="VBD.*"]; [word="will"][tag="VV.*"]; [tag="VBP.*"][lemma="*.ing"]; [tag="VBZ.*"][lemma="*.ing"]; [tag="VBD.*"][lemma="*.ing"]; [word="will   shall"][tag="VB.*"][lemma="*.ing"]; [tag="VHP.*"][tag="VBN.*"]; [lemma="*.ing"]; [tag="VBN.*"]; [tag="VVN.*"]</li> <li>• [tag="VV.*"]; [tag="VVG.*"]; [tag="VVN.*"]</li> </ul>

	<ul style="list-style-type: none"> <li>• [tag="V.*"][tag="RP.*"]</li> <li>• [tag="V.*"][word="myself   yourself   himself   herself   itself   ourselves   themselves   yourselves"]</li> </ul>
<p>Tempos verbais</p> <ul style="list-style-type: none"> <li>• Present</li> <li>• Past simple</li> <li>• Future simple</li> <li>• Present continuous</li> <li>• Past continuous</li> <li>• Future continuous</li> <li>• Present perfect</li> <li>• Past participle</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="VVP.*"]; [tag="VZ.*"]; [tag="VHP.*"]; [tag="VHZ.*"]; [tag="VBP.*"]; [tag="VBZ.*"]</li> <li>• [tag="VVD.*"]; [tag="VBD.*"]</li> <li>• [word="will"][tag="VV.*"]</li> <li>• [tag="VBP.*"][lemma=".ing"]; [tag="VBZ.*"][lemma=".ing"]</li> <li>• [tag="VBD.*"][lemma=".ing"]</li> <li>• [word="will   shall"][tag="VB.*"][lemma=".ing"]</li> <li>• [tag="VHP.*"][tag="VBN.*"][lemma=".ing"]</li> <li>• [tag="VBN.*"]; [tag="VFN.*"]</li> </ul>
<p>Nomes</p> <ul style="list-style-type: none"> <li>• Nomes possessivos</li> <li>• Nomes próprios</li> </ul>	<p>[tag="NN.*"]</p> <ul style="list-style-type: none"> <li>• [tag="NNZ.*"]</li> <li>• [tag="NP.*"]</li> </ul>
Preposições	[tag="IN.*"]
<p>Pronomes</p> <ul style="list-style-type: none"> <li>• Interrogativos</li> <li>• Pessoais</li> <li>• Possessivos</li> </ul>	<ul style="list-style-type: none"> <li>• [tag="WP.*"]</li> <li>• [tag="PP.*"]</li> <li>• [tag="PPZ.*"]</li> </ul>
<p>Verbos</p> <ul style="list-style-type: none"> <li>• Verbo <i>to be</i>.</li> <li>• Verbo <i>to have</i></li> </ul>	<p>[tag="VV.*"]</p> <ul style="list-style-type: none"> <li>• [tag="VB.*"]</li> <li>• [tag="VH.*"]</li> </ul>
<p>Verbos modais</p> <ul style="list-style-type: none"> <li>• Can</li> <li>• Could</li> <li>• May</li> <li>• Might</li> <li>• Must</li> <li>• Need</li> <li>• Shall</li> <li>• Should</li> <li>• Will</li> <li>• Would</li> </ul>	<p>[tag="MD.*"]</p> <ul style="list-style-type: none"> <li>• [lemma="can"]</li> <li>• [lemma="could"]</li> <li>• [lemma="may"]</li> <li>• [lemma="might"]</li> <li>• [lemma="must"]</li> <li>• [lemma="need"]</li> <li>• [lemma="shall"]</li> <li>• [lemma="should"]</li> <li>• [lemma="will"]</li> <li>• [lemma="would"]</li> </ul>
Voz passiva	[lemma="be"][tag="VFN.*"]