



**UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL – UFMS  
CÂMPUS DE TRÊS LAGOAS – CPTL  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS – PPGL**

Eduardo Francisco Ferreira

**CERTIDÕES DE NASCIMENTO, CASAMENTO E ÓBITO: UM PERCURSO COM  
BASE EM CORPUS COMPARÁVEL PARA A EXTRAÇÃO DE FRASEOLOGIAS DE  
DOCUMENTOS OFICIAIS**

Tese submetida ao Programa de Pós-Graduação em Letras da Universidade Federal de Mato Grosso do Sul, Câmpus de Três Lagoas, como um dos requisitos para a obtenção do grau de Doutor em Letras.

Orientadora:  
Prof.<sup>a</sup> Dr.<sup>a</sup> Vanessa Hagemeyer Burgo

Três Lagoas

2024

EDUARDO FRANCISCO FERREIRA

**Certidões de nascimento, casamento e óbito:** um percurso com base em corpus comparável para a extração de fraseologias de documentos oficiais

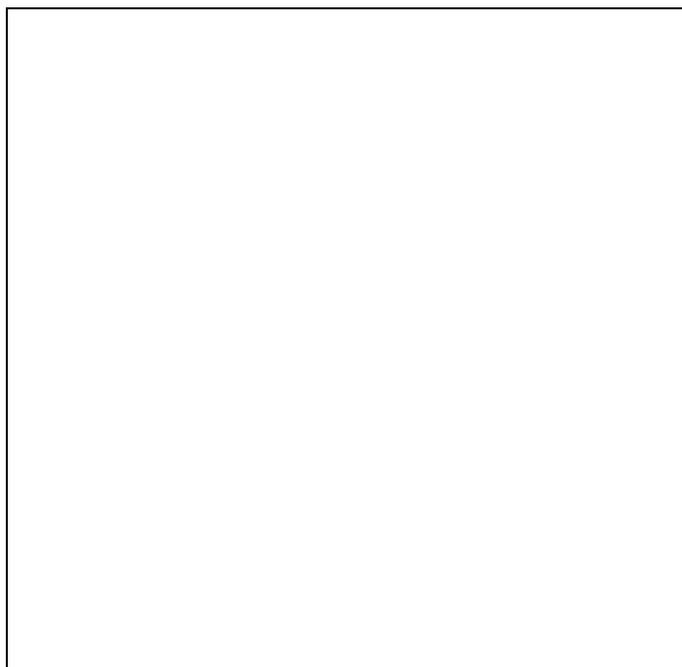
Tese submetida ao Programa de Pós-Graduação em Letras da Universidade Federal de Mato Grosso do Sul, Câmpus de Três Lagoas (UFMS-CPTL), como um dos requisitos para a obtenção do grau de Doutor em Letras.

Orientadora:

Prof<sup>a</sup>. Dr<sup>a</sup>. Vanessa Hagemeyer Burgo

TRÊS LAGOAS  
2024

AUTORIZO A DIVULGAÇÃO OU REPRODUÇÃO, TOTAL OU PARCIAL, DESTE  
TRABALHO, POR MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE  
ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.



## FOLHA DE APROVAÇÃO

EDUARDO FRANCISCO FERREIRA

**Certidões de nascimento, casamento e óbito:** um percurso com base em corpus comparável para a extração de fraseologias de documentos oficiais

Tese submetida ao Programa de Pós-Graduação em Letras da Universidade Federal de Mato Grosso do Sul, Câmpus de Três Lagoas (UFMS-CPTL), como um dos requisitos para a obtenção do grau de Doutor em Letras.

Resultado \_\_\_\_\_ (Preenchido pela banca: aprovado / reprovado)

Três Lagoas/MS, 30 de agosto de 2024

### BANCA EXAMINADORA

Prof<sup>a</sup>. Dr<sup>a</sup>. Vanessa Hagemeyer Burgo (Presidente)

Universidade Federal de Mato Grosso do Sul (UFMS-CPTL)

Prof<sup>a</sup>. Dr<sup>a</sup>. Cláudia Cristina Ferreira

Universidade Estadual de Londrina (UEL)

Prof<sup>a</sup>. Dr<sup>a</sup>. Letícia Jovelina Storto

Universidade Estadual do Norte do Paraná (UENP- CP)

Prof. Dr. Ulisses Tadeu Vaz de Oliveira

Universidade Federal de Mato Grosso do Sul (UFMS-CPTL)

Prof<sup>a</sup>. Dr<sup>a</sup>. Sheyla Cristina Araujo Matoso

Universidade Federal de Mato Grosso do Sul (UFMS-CPTL)

## AGRADECIMENTOS

Definitivamente, o desenvolvimento de uma tese não pode ocorrer de maneira satisfatória sem contar com algum suporte institucional e muito de pessoas ao longo do processo. Então, agradeço imensamente:

- A Deus, com quem eu conversei em alguns momentos, principalmente nos de aflição, pedindo por amparo, calma e saúde, e ser imediatamente acolhido, o que permitiu que eu pudesse terminar este estudo.
- À professora Vanessa Hagemeyer Burgo, minha orientadora, por ter desempenhado essa função com tanta dedicação, empenho e amizade, sempre disponível a compartilhar conhecimentos.
- À professora Mirian Ruffini, minha grande amiga, pelo apoio, conselhos, troca de ideias e tantas outras atitudes que muito contribuíram para a realização deste trabalho.
- À professora Letícia Jovelina Storto, também pelo apoio na hora em que as necessidades simplesmente aparecem. Sempre com alegria, rapidez e amizade.
- Aos professores Ulisses Tadeu Vaz de Oliveira e Sheyla Matoso por serem tão interessados em contribuir, de maneira tão pertinente, criativa e aprofundada com este estudo no momento da banca de qualificação do trabalho.
- Às professoras, e tradutoras juramentadas, Kilda Maria Prado Gimenez e Maud Vanessa Rugeroni pelos conselhos e apoio técnico no decorrer deste estudo.
- Ao Instituto Federal Catarinense pela oportunidade de afastamento integral para o desenvolvimento desta pesquisa.

## RESUMO

O trabalho tradutório não é tão simples como acessar o vocabulário do texto fonte e substituí-lo por aquele a ser apresentado no texto de chegada. Além dessa visão ingênua, existem outros elementos que devem ser levados em conta ao deslocarmos elementos comunicativos de uma língua para outra. Nesse viés, o tradutor, além de avançado conhecimento linguístico, precisa estar equipado com outros fatores envolvidos nesse processo para que o texto traduzido se encaixe adequadamente na língua de chegada, e lá consiga obter o mesmo benefício de que dispunha na língua de partida. Com o desenvolvimento tecnológico, o avanço das pesquisas, muitas interfaces são criadas, desvendando novas searas para estudos e, com isso, gerando a necessidade de produtos terminológicos que possam atender o tradutor ao se deparar com textos especializados. Diante do exposto, o objetivo deste trabalho é apresentar um percurso, com base em corpus comparável, para a compilação de um glossário a partir do texto contido em certidões e traduções juramentadas de certidões de nascimento, casamento e óbito brasileiras e americanas, realizadas por tradutores juramentados brasileiros. O arcabouço teórico está fundamentado nos postulados da Terminologia, associados aos da Linguística de Corpus e aos dos Estudos da Tradução. Nossos resultados mostram a existência de um caminho que pode ser desenhado na interface entre a Terminologia, Estudos da Tradução, mais especificamente a tradução técnica, e Linguística de Corpus, áreas de estudo ainda jovens, mas que, ao compartilhar elementos e objetivos, podem chegar a um produto que satisfaça o consulente na tarefa de superar desafios tradutórios. As análises foram realizadas em direção ao desenho do verbete considerado ideal para atender o tradutor na tarefa de localizar termos e fraseologias de certidões na cultura de chegada.

**Palavras-chave:** Certidões. Tradução. Corpus comparável. Fraseologia.

## ABSTRACT

Translation work is not as simple as accessing the source text's vocabulary and replacing it with equivalent terms in the target language. Beyond this naive view, there are other elements that must be considered when transferring communicative elements from one language to another. In this context, the translator, in addition to having advanced linguistic knowledge, needs to be equipped with other factors involved in this process to ensure that the translated text fits appropriately into the target language and achieves the same impact it had in the source language. With technological development and the advancement of research, many interfaces have been created, revealing new fields of study and, consequently, generating the need for terminological products that can assist translators when dealing with specialized texts. In light of this, the objective of this work is to present a path, based on a comparable corpus, for the compilation of a glossary derived from the content of certificates and certified translations of birth, marriage, and death certificates from Brazil and the United States, carried out by sworn Brazilian translators. The theoretical framework is grounded in the principles of Terminology, combined with those of Corpus Linguistics and Translation Studies. Our results indicate the existence of a path that can be drawn at the interface between Terminology and Translation Studies, specifically technical translation, two relatively young fields that, by sharing elements and objectives, can produce a product that meets the translator's needs in overcoming translational challenges. The analyses focused on designing the ideal dictionary entry to assist the translator in locating terms and phraseologies of certificates within the target culture.

**Keywords:** Certificates, Translation, Comparable Corpus, Phraseology.

## LISTA DE FIGURAS

Figura 1 - Modelo de Certidão de Nascimento no Brasil.....	28
Figura 2 - Modelo de Certidão de Casamento no Brasil. ....	28
Figura 3 - Modelo de Certidão de Óbito no Brasil. ....	29
Figura 4 - Modelo de Certidão de Casamento em Inteiro Teor, no Brasil. ....	29
Figura 5 - Modelo de Certidão de Nascimento Americana.....	34
Figura 6 - Modelo de Certidão de Casamento Americana. ....	37
Figura 7 - Modelo de Certidão de Óbito Americana. ....	39
Figura 8 - Modelo de tradução juramentada de certidão de nascimento .....	59
Figura 9 - WordSmith Tools 9.0.....	107
Figura 10 - Lista de frequência de certidões brasileiras.....	109
Figura 11 - Lista de frequência de certidões americanas. ....	109
Figura 12 - Lista de frequência de traduções em português.....	109
Figura 13 - Lista de frequência de traduções em inglês.....	109
Figura 14 - Concordanciador – o item “fé”.....	112

## LISTA DE QUADROS

Quadro 1 - Distinção entre com Base e Dirigido.....	87
Quadro 2 - Proposta de parâmetros para classificação de corpora. ....	88
Quadro 3 - Critérios e Tipologias Escolhidas. ....	98
Quadro 4 - O corpus comparável deste estudo. ....	105
Quadro 5 - Lista de frequência de traduções em inglês .....	110
Quadro 6 - Microestrutura do glossário em português. ....	116
Quadro 7 - Microestrutura do glossário em inglês. ....	117
Quadro 8 - Microestrutura do glossário em português. ....	117

## LISTA DE SIGLAS

ACETESP	Associação Cearense de Tradutores Públicos
ARPEN-Br	Associação dos Registradores das Pessoas Naturais No Brasil
ATIESP	Associação Profissional de Tradutores Públicos e Intérpretes do Estado de São Paulo
ATPMG	Associação dos Tradutores Públicos de Minas Gerais
ATP-Rio	Associação Profissional de Tradutores Públicos e Intérpretes Comerciais, Juramentados, do Estado do Rio de Janeiro
BNC	British National Corpus
CITRAT	Centro Interdepartamental de Tradução e Terminologia
CNJ	Conselho Nacional de Justiça
COCA	Corpus of Contemporary American English
DNRC	Departamento Nacional de Registro do Comércio
DRC	Departamento de Registros Cíveis
JUCEB	Junta Comercial do Estado da Bahia
JUCEMG	Junta Comercial do Estado de Minas Gerais
JUCEPE	Junta Comercial do Estado de Pernambuco
JUCESP	Junta Comercial do Estado de São Paulo
LC	Linguística de Corpus
PC	Computador Pessoal
TCT	Teoria Comunicativa da Terminologia
TGT	Teoria Geral da Terminologia
TJ	Tradução Juramentada
UF	Unidade Fraseológica
UFE	Unidade Fraseológica Especializada
UNICEF	Fundo das Nações Unidas para a Infância
UT	Unidade Terminológica

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	14
<b>1. O FOCO DO ESTUDO</b> .....	20
1.1. <b>CERTIDÕES NO BRASIL</b> .....	20
1.1.1. <b>Certidões de Nascimento, de Casamento e de Óbito</b> .....	23
1.1.2. <b>Serventia Notarial e Registral no Brasil, o Cartório</b> .....	24
1.1.3. <b>Registro Civil das Pessoas Naturais</b> .....	25
1.1.4. <b>Tipologia das Certidões Brasileiras</b> .....	26
1.2. <b>CERTIDÕES NOS ESTADOS UNIDOS</b> .....	30
1.2.1. <b>Notários Americanos</b> .....	30
1.2.2. <b>Departamento de Registros Cíveis</b> .....	31
1.3. <b>OS DESAFIOS NA TRADUÇÃO DE CERTIDÕES</b> .....	40
1.3.1. <b>Solucionando Os Desafios Tradutórios</b> .....	44
1.4. <b>CERTIDÕES TRADUZIDAS: O MODO JURAMENTADO</b> .....	45
1.4.1. <b>O Tradutor Juramentado</b> .....	45
1.5. <b>A TIPOLOGIA DOS DOCUMENTOS</b> .....	50
1.5.1. <b>O Texto Traduzido</b> .....	52
1.5.2. <b>Divisão Textual da Tradução Juramentada</b> .....	53
<b>2. FUNDAMENTAÇÃO TEÓRICA</b> .....	60
2.1. <b>TERMINOLOGIA</b> .....	60
2.1.1. <b>Teoria Geral da Terminologia (TGT)</b> .....	60
2.1.2. <b>Teoria Comunicativa da Terminologia (TCT)</b> .....	63
2.1.3. <b>Unidade Terminológica (UT)</b> .....	65
2.1.4. <b>Unidade Fraseológica (UF)</b> .....	66
2.2. <b>LINGÜÍSTICA DE CORPUS (LC)</b> .....	77
2.2.1. <b>Percurso da LC</b> .....	78
2.2.2. <b>Estatuto da LC</b> .....	81
2.2.3. <b>O Conceito de Corpus</b> .....	83
2.2.4. <b>Baseado em Corpus ou Dirigido pelo Corpus</b> .....	86
2.2.5. <b>Tipos de Corpus</b> .....	87
2.2.6. <b>Colocação</b> .....	90
<b>3. METODOLOGIA</b> .....	94

3.1. O PERCURSO NESTE ESTUDO .....	97
3.1.1. A Seleção .....	98
3.1.2. A Busca .....	98
3.1.3. O Tratamento .....	99
3.1.4. A Descrição .....	99
3.1.5. A Ferramenta .....	100
3.1.6. As Listas de Frequência e de Concordâncias .....	101
3.1.7. A Leitura da Concordância.....	101
3.1.8. A Fraseologia, a Colocação e a Polissemia .....	102
3.1.9. A Composição do Corpus.....	102
3.2. PROCEDIMENTOS, DADOS E ANÁLISES .....	104
3.2.1. A Utilização do <i>Software</i> .....	106
3.3. A CONFECÇÃO DO GLOSSÁRIO.....	115
3.3.1. Macroestrutura .....	115
3.3.2. Microestrutura .....	115
CONCLUSÃO .....	120
REFERÊNCIAS .....	125

## INTRODUÇÃO

Todo texto, em maior ou menor escala, ao ser traduzido, representa um desafio ao tradutor, visto que as escolhas tradutórias se localizam em ambientes complexos e desafiadores o que, por conseguinte, influenciam as tomadas de decisões (Toury, 2012, p. 76; Krieger e Maciel, 2001, p. 159; Cabré, 1999, p. 216; Baker, 1992, p. 9; Holmes, 1988, p. 67). Diante do exposto, é possível entender que, dependendo do ambiente cultural (Stupiello, 2008, p. 82) e do gênero a que pertence, cada texto representa nuances, particularidades e especialidades em variados níveis, fato que pode impor diferentes maneiras de se conduzir uma tradução.

Sempre que houver um desafio tradutório, o tradutor precisa recorrer a produtos terminológicos que o ajudem a transmitir o significado por meio da correta seleção de equivalentes da cultura de chegada. Sem a ajuda de um glossário ou de um dicionário terminológico, ele pode ficar inseguro durante a tentativa de transpor informações em ambiente interlinguístico, por exemplo, ao procurar o verbo que melhor se ajuste a determinado termo, carência que pode ocasionar atrasos na confecção de textos técnicos e científicos na língua de chegada (Bevilacqua, 2004/2005, p. 79).

Nesse momento, o profissional que pensava ter somente um desafio se depara com outro: os dicionários técnicos e/ou glossários são superficiais, falhos ou, simplesmente, não existem em determinadas áreas, fato que, de acordo com Cabré (1999a, p. 12), poderia ser dirimido por meio do envolvimento de tradutores, juramentados ou não, no momento da construção de produtos terminológicos, em razão de serem pessoas que nitidamente enxergam as reais necessidades para a realização de traduções competentes. A título de exemplificação, Teixeira (2008, p. 4) relata que muitos dos dicionários e glossários destinados ao tradutor são apenas listas de termos. Não apresentam exemplos de uso, explicações a respeito do grau de equivalência, nem da colocação de preposições ou adjetivos.

Diante do exposto, é possível inferir que a dificuldade para encontrar recursos terminológicos e terminográficos de consulta é diretamente proporcional ao nível de especificidade da área sobre a qual se realiza o trabalho tradutório. Com essa constatação, passo a tecer discussões a respeito dos textos de interesse neste estudo, mais especificamente os contidos em certidões de nascimento, casamento e óbito.

Em uma sociedade organizada, todo acontecimento civil deve ser registrado em algum órgão designado para guardá-lo indefinidamente. Dependendo do país, os registros são mantidos em formulários, como nos Estados Unidos da América, ou em livros, como no Brasil. Sempre que um indivíduo desejar exercer sua cidadania como, por exemplo, obter título de

eleitor, fazer matrícula em escola, ou se casar, é provável que surja a necessidade de comprovar os dados constantes no registro depositado. Para isso, deve procurar a repartição em que o assento foi realizado e solicitar uma certidão, que é o documento legal para transportar cópias das informações necessárias para além dos domínios da serventia em que estão.

Pode-se entender que a atividade registral que ocorre nos cartórios possui, no mínimo, duas funções: **1)** em direção ao cidadão, por fornecer certidões, instrumento que permite acesso a variados âmbitos institucionais e sociais; **2)** em direção a instâncias governamentais, por fornecer dados que, ao serem tabulados, tornam-se estatísticas que podem ser convertidas em melhorias nos procedimentos, seguidas de consequentes benefícios à população. Por exemplo: o número de nascimentos em determinada localidade pode desencadear políticas de planejamento e custeio de saúde, educação, habitação, etc., ou seja, o planejamento da infraestrutura necessária que possa atender adequadamente os cidadãos. Estatísticas a respeito dos óbitos em determinada região, por exemplo, são utilizadas para esclarecer em que medida o desenvolvimento de políticas, programas e iniciativas em saúde, educação, segurança, etc. para resolver ou mitigar os elementos causadores do desnível em questão devem ser empregados.

Existem vários tipos de registros arquivados a partir dos quais são emitidos variados tipos de certidões. Há certidões positivas, negativas, de propriedade, de tributos, de filiação e inúmeras outras. No entanto, neste trabalho, investigo somente as certidões de nascimento, casamento e óbito, tanto do Brasil quanto dos Estados Unidos, portanto, no par linguístico português brasileiro/inglês americano, com interesse em contribuir com o tradutor no momento em que se depara com esse tipo de texto.

Tendo em mente um trabalho que alinhe os elementos discutidos até aqui, acredito na possibilidade de existência de um caminho que possa ser desenhado na interface entre a Terminologia, Linguística de Corpus e os Estudos da Tradução, mais especificamente a tradução técnica, áreas de estudo ainda jovens, mas que, ao compartilhar elementos e objetivos, podem chegar a um produto que satisfaça o consulente na tarefa de superar desafios tradutórios.

Uma pesquisa a respeito de vocabulário contido em certidões, com vistas à compilação de um glossário, precisa levar em consideração os preceitos básicos da Terminologia e sua evolução, visto que é possível perceber uma crescente necessidade de padronização, em nível internacional, iniciada no século anterior a este (Williams e Chesterman, 2002, p. 20). Com nascimento marcado pela necessidade de nomear objetos e atividades, é correto afirmar que a Terminologia ainda mantém essa vertente em seu bojo e que, diante do aumento da complexidade nas comunicações, imposto pelos avanços científicos e tecnológicos, é a todo

momento convocada para exercer sua natureza essencial. Nesse novo cenário, esclarece Barros (2004, p. 57), os parâmetros utilizados pela Teoria Geral da Terminologia (TGT), idealizada por Wüster, começam a perder o vigor e não conseguem acompanhar a velocidade acelerada das mudanças, deixando espaço para críticas e novas propostas, lacuna preenchida com os postulados da Teoria Comunicativa da Terminologia (TCT), concebida por Cabré, na década de 1990.

Amplamente aceita pelos estudiosos da área, a TCT, mesmo com a possível existência de outras teorias, passa a funcionar como paradigma teórico para a geração de metodologias de trabalho rumo a proporcionar conhecimentos mais afinados com as complexidades comunicativas em acelerada evolução.

Redes de satélites e de computadores permitem a comunicação em tempo real. Diante disso, a Terminologia foi compelida a aumentar o escopo de atuação, isto é, além de termos e definições, é necessário prover fraseologias, também levando em conta os contextos em que são utilizadas. Pesquisas sobre fraseologia são determinantes na tentativa de resolver “problemas semânticos relacionados à transposição de significados de uma língua para a outra, posto que as estruturas fraseológicas costumam se caracterizar pela projeção de um sentido que não se depreende da soma de seus constituintes” (Krieger e Finatto, 2004, p. 86).

Diante da constante evolução social, científica e tecnológica aproximando povos e, por conseguinte, gerando necessidades de comunicação cada vez mais específicas e de ampla divulgação por meio de produtos terminológicos, é possível afirmar que o mercado de tradução técnica está em franca expansão. Tendo em mente o grande volume de textos a serem analisados, e trazendo em sua base o componente tecnológico que viabilize esse processamento, os profissionais da linguagem passam a contar com a rapidez de sistemas eletrônicos na apreciação do grande potencial oferecido pelo trabalho com corpus. De acordo com Hunston (2010, p. 3), o corpus pode ser reorganizado de maneira a possibilitar vários tipos de observação, a partir de uma nova perspectiva, elucidando elementos como frequência, fraseologia e colocação, trabalho extenuante e falho ao contar apenas com esforço e intuição humanos.

Existem várias denominações que se dirigem ao tipo de corpus a ser estudado. Pode ser classificado como especializado, geral, comparável, paralelo, pedagógico, de controle, entre outros, ou seja, o nome pode ser consequência do registro do tipo de texto, finalidade do estudo, ou maneira do possível alinhamento. Neste trabalho serão analisados alguns modelos de textos de certidões originais de nascimento, casamento e óbito tanto brasileiras quanto americanas, configurando, portanto, um corpus do tipo comparável bilíngue, no par linguístico Português

Brasileiro/Inglês Americano. O corpus comparável, afirma Hunston (2010, p. 15), é formado a partir de

dois (ou mais) corpora em diferentes línguas (ex. inglês e espanhol), ou diferentes variedades de língua (ex. inglês indiano e inglês canadense). São compostos ao longo das mesmas linhas e, por exemplo, vão conter a mesma proporção de textos de jornal, romances, conversas casuais e assim por diante [...]. Corpora comparáveis de diferentes línguas podem ser usados por tradutores e por aprendizes para identificar diferenças e equivalências em cada língua<sup>1</sup> (Tradução minha).

Frente à discussão realizada até este momento, é válido afirmar que, para se chegar a produtos terminológicos e terminográficos há possível caminho conduzido pela Terminologia, compartilhado ou não, guiado pela Linguística de Corpus. Com isso, podem-se rastrear percursos metodológicos para a extração de termos, defendido e utilizado pelos estudiosos e profissionais de cada uma delas, a fim de tentar revelar escopo e alcance.

Nesse ambiente, minha tese é a de que um produto terminográfico montado para atender as necessidades do tradutor no desempenho de seu ofício deve conter informações Terminológicas que possam garantir solidez de conhecimentos na área de interesse, associadas a conhecimentos estatísticos (frequência) proporcionados pela metodologia da Linguística de Corpus, ao mesmo tempo em que evidencia a maneira com a qual essas unidades terminológicas se comportam dentro dos corpora da mencionada área. Com isso, a metodologia para a compilação de glossários e dicionários especializados de uma determinada área ou tipo textual deve contemplar todos os itens terminológicos e arranjos fraseológicos que possam causar dificuldades ao tradutor, com e sem frequência relevante no corpus pesquisado.

Este trabalho se diferencia dos outros propostos até o momento por não haver metodologia atualizada e detalhada que combine os preceitos da Terminologia e da Linguística de Corpus para a extração de unidades terminológicas a partir de um corpus comparável bilíngue, Português/Inglês Americano, de certidões originais e de traduções juramentadas de certidões. Portanto, é necessário a implementação desta pesquisa, tendo em mente a superação desta lacuna.

Com isso, o **Objetivo Geral** será o de evidenciar um percurso detalhado para a extração de fraseologias a partir de um corpus comparável de certidões de nascimento, casamento e óbito

---

<sup>1</sup> Comparable Corpora: two (or more) corpora in different languages (e.g. English and Spanish) or different varieties of a language (e.g. Indian English and Canadian English). They are designed along the same lines, for example they will contain the same proportions of newspapers texts, novels, casual conversation, and so on. [...] Comparable corpora of different languages can be used by translators and by learners to identify differences and equivalences in each language.

brasileiras e americanas e traduções juramentadas desse mesmo tipo de certidões, que possa funcionar como modelo para tradutores e pesquisadores.

Para nortear este estudo em direção a resultados satisfatórios, serão levados em consideração os seguintes **Objetivos Específicos**:

- 1) Identificar a maneira com a qual um corpus comparável pode auxiliar na extração de itens lexicais candidatos a termos de um glossário de certidões;
- 2) Descrever os procedimentos básicos (passo-a-passo) para a extração de unidades terminológicas e fraseológicas de um corpus comparável de certidões;
- 3) Indicar quais seriam a macro e a micro estruturas mais adequadas para a compilação de um glossário desse tipo, com base nos candidatos extraídos do corpus;
- 4) Reportar o que pode ser mencionado a respeito do processo tradutório empreendido pelos tradutores juramentados que contribuíram com este estudo.

O empenho no desenvolvimento desta pesquisa se justifica por alguns motivos. O primeiro deles é justamente por não haver um glossário ou dicionário especializado bilíngue (português/inglês), desenvolvido a partir de corpus comparável de certidões originais e de traduções juramentadas de certidões de nascimento, casamento e óbito, nas duas direções, destinado a suprir as necessidades do tradutor em encontrar equivalentes tradutórios que funcionem de maneira eficaz nas línguas de chegada. Tal restrição pode ocorrer porque também não há trabalhos que indiquem os melhores encaminhamentos metodológicos para extração de unidades terminológicas a partir desse tipo de corpus contendo especificamente as mencionadas certidões. Buscarei converter elementos desta pesquisa em insumos que possam contribuir para sanar a lacuna existente no que se refere a evidenciar o passo-a-passo, por meio de corpus comparável, para se chegar a um produto terminológico capaz de satisfazer o consulente. Ao mesmo tempo, o trabalho aqui empreendido pode contribuir na condução de pesquisas que desvendem fenômenos tradutórios, fator de grande valia na formação e informação de tradutores, pesquisadores, alunos dos cursos de tradução e de todas as outras pessoas que se interessem por este tema.

Esta tese está dividida em três seções, de acordo com o descrito a seguir:

A **Seção 1** situa o leitor a respeito da tipologia dos textos utilizados nesta pesquisa. Consequentemente, aborda assuntos relacionados às certidões de nascimento, casamento e óbito utilizadas no Brasil, mencionando as bases de sua constituição, instâncias notariais onde solicitá-las a partir do assento realizado em livros e esclarecendo a maneira de como podem servir para o cidadão e para órgãos do governo. Ao mesmo tempo, também aborda assuntos relacionados às certidões em pauta, mais especificamente as americanas, traçando um paralelo

dentro dos mesmos parâmetros de finalidades e utilizações com as certidões brasileiras em questão, evidenciando semelhanças e diferenças. Também discute os possíveis desafios ao tradutor, em especial o juramentado brasileiro, na tarefa de ajustar uma certidão de nascimento, casamento, ou óbito aos parâmetros da cultura receptora, atividade que possibilita ao contratante acessar, no país de chegada, os mesmos direitos adquiridos no país de partida. Neste momento, a tradução juramentada de certidões de nascimento, casamento e óbito entram em tela. Nesta parte são explicados como ocorre o procedimento para que um tradutor obtenha a permissão para trabalhar como juramentado; os desafios que deve enfrentar e as possíveis soluções ao lidar com essa modalidade tradutória; e a tipologia dos documentos sob responsabilidade desse profissional, principalmente a respeito das certidões estudadas neste trabalho.

A **Seção 2** aborda os preceitos teóricos que norteiam o caminho, retratando os postulados da Terminologia, seu percurso e o objeto de estudo focalizado neste trabalho: a fraseologia. Mencionaremos aspectos terminográficos, incluindo elementos de como devem ser a macro e microestrutura de um glossário desenvolvido para as certidões em questão. Para finalizar o capítulo, abordarei os preceitos da Linguística de Corpus e seu comportamento na solução de desafios tradutórios, o que faz por meio de dados concretos de uso.

Na **Seção 3** destacarei a metodologia de trabalho utilizada para lidar com o corpus selecionado. Esclarecerei como o corpus foi compilado, como os textos foram selecionados, o objetivo de sua criação e as ferramentas utilizadas na extração terminológica. Também explicitarei as análises realizadas em direção ao desenho do verbete considerado ideal para atender o tradutor na tarefa de localizar termos e fraseologias de certidões, na língua de chegada.

## 1. O FOCO DO ESTUDO

Sempre que um indivíduo deseja exercitar sua cidadania como, por exemplo, adquirir produto bancário, participar de projetos ou benefícios governamentais ou, simplesmente, ingressar numa instituição escolar, algum tipo de comprovação da existência de um registro em cartório pode ser solicitado. Nesse contexto, as certidões existem justamente para cumprir esse papel de corroboração, transportando as informações contidas em registros arquivados, para que o usuário viabilize o acesso às finalidades pretendidas.

Nesta parte do estudo, pretendo discorrer a respeito dos documentos emitidos por órgãos competentes que visam a comprovar a existência de um registro: as certidões. Para tanto, apresentarei as definições para esse gênero textual, alguns tipos que circulam no Brasil e, em seguida, as partes comumente esperadas em sua configuração. De acordo com Costa (2014), os documentos originais em estudo neste trabalho pertencem ao gênero Certidão, ou seja, “documento, atestado ou certificado legal em que o serventuário oficial certifica alguma coisa de que tem provas, como, por exemplo, certidão de idade, certidão de casamento, etc.” (p. 68).

Mais adiante, haverá afunilamento neste assunto, momento em que abandono todos os outros tipos de certidão para tratar, especificamente, do subgênero das de nascimento, casamento e óbito, foco do interesse deste trabalho, conforme apontamentos nos estudos de Oliveira (2019). Apresento, também, um relato a respeito de cartórios em geral para, logo depois, focalizar o Registro Civil das Pessoas Naturais, elencando as garantias que oferta, a natureza do trabalho que desenvolve e como contribui no desenvolvimento de estatísticas. Para fechar esta parte, abordo a tipologia das certidões sob estudo, as informações que comportam e o significado das numerações/códigos utilizados, seguidos pelos modelos de cada tipo da nova certidão nacionalmente padronizada.

### 1.1. CERTIDÕES NO BRASIL

Na tentativa de encontrar uma definição que esclareça ao público geral, não especializado na área, o que é uma certidão, inicio a busca pelos dicionários a) Houaiss<sup>2</sup> e b) Aulete<sup>3</sup>, que retornaram as explicações mencionadas a seguir, respectivamente:

- a) Certidão: substantivo feminino

---

<sup>2</sup> <http://houaiss.uol.com.br/>

<sup>3</sup> <http://www.aulete.com.br/certid%C3%A3o>

- 1- Qualidade do que é certo; certeza.
- 2- Termo jurídico. Documento com fé pública emitido por tabelião ou escrivão, comprovando determinado evento. Ex.: c. de nascimento / c. de óbito
- 3- Rubrica: termo jurídico. Cópia autêntica ou transunto; traslado.

b) (cer.ti.dão) sf.

- 1- Documento de comprovação firmado por autoridade competente e de fé pública;  
ATESTADO: certidão de batismo: certidão de nascimento: certidão de casamento  
[Pl.: -dões.]  
[F.: Do lat. *certitudo*.]

Diferentemente de um glossário, os dicionários não são específicos de uma determinada área de estudo e, portanto, trazem apenas apontamentos, resultados genéricos que, muitas vezes, não satisfazem a necessidade do usuário técnico ou especializado. Nesse caso, por se tratar de uma pesquisa especializada, procurou-se aprofundamento nessa busca por meio de consulta ao que dizem os órgãos que lidam com as certidões em sua rotina, conforme a seguir. O Tribunal de Justiça da Bahia<sup>4</sup>, por exemplo, atende a condição mencionada e especifica um pouco mais a resposta ao mencionar que

[c]ertidão é um documento no qual o Oficial do cartório certifica que o registro se encontra devidamente lavrado nos livros sob sua responsabilidade. Independente de despacho judicial e dando fé pública, o oficial reproduz, de forma autêntica e absolutamente confiável, textos de um assento ou documento arquivado em sua serventia, fazendo inserir na certidão, obrigatoriamente, o número do livro, da folha e do termo sob o qual foi lavrado, ou ainda o número do registro ou pasta ou caixa em que o documento se encontra arquivado. Além do que, as certidões dos atos relativos ao estado civil devem conter sempre a indicação da serventia e respectivo oficial.

Uma certidão existe para comprovar a existência de um fato não transitório, mas sim, estável. Com isso, é possível entender que pode haver uma certidão para cada fato permanente e que, portanto, há um grande espectro de exemplares no que tange à quantidade de certidões existentes. Uma certidão pode ser requerida a partir de muitos setores, visto que existem várias instâncias de repartições públicas e privadas. Para evidenciar essa imensidão de possibilidades, serão citados alguns exemplos de repartições e do tipo de certidões que podem emitir, respectivamente:

---

<sup>4</sup>[http://www5.tjba.jus.br/index.php?option=com\\_content&view=article&id=77646&Itemid=193](http://www5.tjba.jus.br/index.php?option=com_content&view=article&id=77646&Itemid=193)

**Repartições:** Cartórios (de Registro, ou de Protesto); Distribuidores (Civil, ou criminal); Institutos; Juntas Comerciais; Justiças (do Trabalho, Eleitoral ou Federal); Prefeituras; Procuradorias; Registros de Imóveis; Secretarias; Tabelionatos; etc...

**Certidões:** Nascimento, Casamento, Óbito; Negativas, Positivas; Falência; Distribuição, Execução; Ações Cíveis e de Família; Executivos; Propriedade; Filiação; Tributos; Procução; etc...

Assim como qualquer outro país, com maior ou menor complexidade, o Brasil também adota procedimentos burocráticos para viabilizar o acesso a direitos e o cumprimento de deveres o que, por conseguinte, pode fomentar a criação ou a manutenção de uma gama de instituições de variados âmbitos, autorizadas a emitir certidões, como espelhado anteriormente. Ainda assim, é possível rapidamente fazer a distinção quando em meio a outros tipos de documentos, visto que as certidões são facilmente reconhecidas pelo aspecto que têm. A esse respeito, Asensio (2003, p. 65) explica que

[o] formato (ou o layout) é um dos elementos que ajudam a transmitir o significado global de um texto. Esse formato é comumente ligado ao tipo de texto, especialmente com a popular categorização de textos legais. Um texto que está de acordo com o formato esperado é percebido como natural e plausível; se não, é percebido como não natural e manipulado<sup>5</sup> (Tradução minha).

Em alguns casos, o que determina a manutenção de um formato é a força da lei, mas, na maioria das vezes, é mantido pela tradição, o hábito de ter sido sempre assim e, por esse motivo, existem variações na maneira como se apresentam (Asensio, 2003, p. 65). As certidões brasileiras emitidas por órgãos responsáveis parecem respeitar o formato que tem um documento oficial. Com isso, é comum, natural e esperado que tenham as seguintes partes:

1. **Identificação** – nome da entidade emissora.
2. **Título** - a palavra CERTIDÃO, em letras maiúsculas, sobre o texto.
3. **Texto** - um parágrafo com o teor da Certidão.
4. **Local e data** - por extenso.
5. **Assinatura(s)**.

---

<sup>5</sup> Format (or layout) is one of the elements that help transmit the global meaning of a text. It is commonly linked to the text type, especially with the popular categorization of legal texts. A text that conforms to an expected format is perceived as natural and plausible; if not, it is perceived as unnatural and manipulated.

Após essa breve discussão geral sobre certidões, a próxima seção focalizará os tipos de certidões a serem exploradas nesta pesquisa.

### 1.1.1. Certidões de Nascimento, de Casamento e de Óbito

Até o final do ano de 2010, as certidões de nascimento, casamento e óbito poderiam ter papel, tamanho, formato e informações que variavam de um lugar para outro dentro do território nacional. Por não se ter documentos com diretrizes nacionalmente unificadas e por negligência a questões de segurança, havia espaço para falsificações e fraudes. Para preencher essa lacuna, desde janeiro de 2011, a Casa da Moeda, em parceria com a Corregedoria do Conselho Nacional de Justiça (CNJ) e o Ministério da Justiça, emite certidões de casamento, nascimento e óbito padronizadas e com tecnologia antifraude. Os novos documentos são impressos em papel moeda e trazem itens de segurança como marca d'água e o brasão da República, entre outros, visando a inibir falsificações, ao mesmo tempo em que permite verificar a quantidade correta de brasileiros que nascem, se casam e/ou morrem, por terem emissões totalmente informatizadas. Ainda de acordo com a CNJ, por meio do Provimento nº 3, de 17 de novembro de 2009<sup>6</sup>, e em nota emitida nesse mesmo mês de novembro<sup>7</sup>, e reafirmada em janeiro de 2010<sup>8</sup>, por sua agência de notícias, as certidões antigas não perdem o valor e podem continuar a ser utilizadas por tempo indeterminado, não havendo, com isso, a necessidade de que uma via atualizada seja feita.

Neste trabalho, serão consideradas as traduções de certidões no par linguístico inglês/português, nas duas direções, não somente de certidões padronizadas, mas também as que foram traduzidas em época anterior a janeiro de 2011.

No intuito de entender o que são e a utilização das certidões do foco deste trabalho, a seguir, serão apresentadas as suas definições<sup>9</sup>:

- **Certidão de Nascimento:** comprova a existência de uma pessoa, seu local e data de nascimento, o nome dos seus pais e avós. É vista como o primeiro movimento para o pleno exercício da cidadania, já que viabiliza o acesso a instâncias sociais, políticas, econômicas, entre outras.

<sup>6</sup> [http://www.cnj.jus.br/images/stories/docs\\_corregedoria/provimentos/provimento\\_n\\_03.pdf](http://www.cnj.jus.br/images/stories/docs_corregedoria/provimentos/provimento_n_03.pdf)

<sup>7</sup> <http://www.cnj.jus.br/noticias-gerais/8021-orientacoes-para-a-implantacao-das-novas-certidoes-ja-estao-disponiveis->

<sup>8</sup> <http://www.cnj.jus.br/noticias-gerais/8342-corregedoria-divulga-nota-tecnica-sobre-obrigatoriedade-das-novas-certidoes>

<sup>9</sup> <http://www.brasil.gov.br/>

- **Certidão de Casamento:** documento passado em cartório para registrar a união de duas pessoas de sexo oposto. Registra informações como a data do casamento, o nome dos noivos, suas profissões, estado civil, endereços, naturalidade, nacionalidade, entre outros.
- **Certidão de Óbito:** é o registro do falecimento do cidadão, feito em Cartório Civil de onde ocorreu o óbito. É feita a partir de uma Declaração de Óbito - documento em que o médico atesta a causa mortis da pessoa. Essa certidão contém informações de identificação da pessoa falecida, com datas, horários e locais; causa do falecimento; herdeiros, lugar de sepultamento, entre outras.

No Brasil, as certidões de nascimento, casamento e óbito podem ser emitidas em três formatos: em breve relato; em inteiro teor; ou em relatório. Cada modalidade será discutida mais adiante ao relatar a tipologia de cada uma das certidões mencionadas nesta parte do estudo.

### 1.1.2. Serventia Notarial e Registral no Brasil, o Cartório

Disciplinado pela Lei nº 8.935/1994<sup>10</sup>, o cartório, ou serventia notarial e registral, é a entidade encarregada de eternizar os acontecimentos civis da vida dos cidadãos. Imprescindível na burocracia brasileira, é responsabilidade do cartorário “[...] garantir a publicidade, autenticidade, segurança e eficácia dos atos jurídicos preventivamente, desobstruindo o Judiciário do acúmulo de processos instaurados, no intuito de reestabelecer a Ordem Jurídica no país, e atuando como instrumento de pacificação social” (Rocha, 2007, p. 10).

O cartório é facilmente interpretado como pertencente ao âmbito governamental de serviços à comunidade, mas, de fato, a atividade notarial é exercida por particulares que colaboram com Poder Público, garantindo respeitabilidade jurídica aos eventos civis (Flach, 2004, p. 23).

Fiscalizada pelo Poder Judiciário, a rotina de uma serventia notarial e registral é repleta de tarefas, segundo Ribeiro (2008, p. 2), como: investigar elementos da realização de um ato; dar pareceres jurídicos; instrumentalizar adequadamente a vontade das partes; guardar de documentos; garantir maior segurança jurídica; entre outros.

De acordo com a Agência Senado<sup>11</sup>, existem vários tipos de cartórios como, por exemplo, o Tabelionato de notas; o Tabelionato de Protesto de Títulos; o Registro Civil das

<sup>10</sup> [http://www.planalto.gov.br/ccivil\\_03/leis/18935.htm](http://www.planalto.gov.br/ccivil_03/leis/18935.htm)

<sup>11</sup> <http://www.senado.gov.br/noticias/jornal/cidadania/Cart%C3%B3riosconcess%C3%B5esp%C3%BAblicas/not02.htm>

Pessoas Jurídicas; entre outros. Neste estudo, o cartório de interesse é chamado de Registro Civil das Pessoas Naturais, e tem a incumbência de registrar, e fornecer as respectivas certidões, os atos de nascimento, casamento, óbito, ausência ou morte presumida, emancipação ou interdição e tutela, opção de nacionalidade e sentença de deferimento de legitimação adotiva.

### 1.1.3. Registro Civil das Pessoas Naturais

A respeito dos cartórios, é possível entender a importância dos atos notariais ao garantir autenticidade, segurança, eficácia e publicidade (art. 1º, da Lei nº 8.935/1994), ou seja, fé pública, presunção de veracidade e produção dos efeitos jurídicos desejados. Capazes de gerar direitos e obrigações, os atos praticados no Registro Civil das Pessoas Naturais, de acordo com Ribeiro (2008, p. 23) e Rocha (2007, p. 10), possuem três vertentes distintas:

- a) **Natureza Declaratória ou Comprobatória:** declaram fato preexistente e têm efeitos comprobatórios como, por exemplo, as certidões de nascimento e de óbito.
- b) **Natureza Constitutiva:** Constituem direito a partir de ato registral. Nesse caso, pode-se citar a emancipação, ato que só tem validade a partir do registro competente.
- c) **Natureza Publicitária:** dá ciência do ato a todos os terceiros. O Registro de Interdição visa a informar os terceiros de que determinada pessoa não poderá praticar todos ou alguns atos da vida civil.

Assim como nas outras serventias de registros civis, o Registro Civil das Pessoas Naturais organiza os atos em livros. Com isso, os assentos de várias naturezas como registro de nascimento, adoção, casamento civil, óbito e natimorto, averbações e traslados de registros de brasileiros feitos no exterior, por exemplo, são anotados separadamente de acordo com o tipo. Ribeiro (2008, p. 28) esclarece que esses livros apresentam no máximo trezentas (300) folhas e são divididos por atos, indicados por letras, conforme a seguir:

- a) **Livro A** – Registro de Nascimento;
- b) **Livro B** – Registro de Casamento;
- c) **Livro B Auxiliar** – Registro de casamento religioso para efeitos civis;
- d) **Livro C** – Registro de óbito;
- e) **Livro C Auxiliar** – Registro de Natimorto;
- f) **Livro D** – Registro de Proclamas;
- g) **Livro E** – Demais atos da vida civil (máximo de 150 folhas).

Além de preservar a história dos cidadãos, os dados colhidos por ocasião dos registros são encaminhados para análise e montagem de quadros estatísticos, utilizados para direcionar

providências, recursos financeiros e políticas. Especificamente do Registro Civil das Pessoas Naturais, esclarece Ribeiro (2008, p. 33) “decorrem importantíssimas relações concernentes à família, à sucessão, à organização política do estado e a sua própria segurança interna e externa, além de ser uma fonte de estatística de sua população.”

#### 1.1.4. Tipologia das Certidões Brasileiras

Diante da natureza de divulgação atribuída aos cartórios, obter uma certidão é algo muito fácil, visto que qualquer pessoa pode fazer a solicitação, sem prestar esclarecimento ao oficial ou ao funcionário sobre os motivos ou interesses envolvidos no pedido, bastando, para isso, recolher os emolumentos referentes ao serviço a ser desempenhado (Rocha, 2007, p. 23). Para atender a necessidade ou desejo de uma reprodução de ato jurídico por escrito, o cartório disponibiliza três (03) modalidades de certidão, à livre escolha do solicitante: a) **em inteiro teor**: como o próprio nome diz, compreende o registro por inteiro, ou seja, todas as informações contidas no livro são transpostas para a certidão de maneira textual; b) **em resumo**, ou **em breve relato**: composta somente pelas informações principais, expressa por meio de formulário; c) **em relatório**: contempla os quesitos elaborados pelo próprio interessado, devidamente respondidos.

Segundo a Corregedoria Nacional de Justiça (CNJ)<sup>12</sup>, o dia 27 abril de 2009 marca o lançamento dos modelos padronizados de certidões de nascimento, casamento e óbito, seguido pelo Decreto 7.231, de 14 de junho de 2010<sup>13</sup>, que oficializa a data a partir da qual entra em vigor a obrigatoriedade de utilização de papel padronizado, ou seja, 02 de julho de 2012. Ainda de acordo com a CNJ, as mudanças no sistema e nas certidões visam a proporcionar mais segurança aos documentos, na tentativa de interromper erros e falsificações, ao mesmo tempo em que facilitam a conferência da autenticidade dos registros. O novo papel de segurança é emitido pela Casa da Moeda, que passa a atender pedidos dos cartórios de registro e unidades integradas implantadas em maternidades do país.

Além da Corregedoria Nacional de Justiça, a implantação das novas certidões contou com a participação do Ministério da Justiça, da Secretaria de Direitos Humanos da Presidência da República, da Casa da Moeda, dos Tribunais Estaduais e da Associação dos Registradores das Pessoas Naturais do Brasil (ARPEN-BR)<sup>14</sup>, visando ao aprimoramento do Provimento. A

---

<sup>12</sup> <http://www.cnj.jus.br/>

<sup>13</sup> [http://www.planalto.gov.br/ccivil\\_03/ato2007-2010/2010/Decreto/D7231.htm](http://www.planalto.gov.br/ccivil_03/ato2007-2010/2010/Decreto/D7231.htm)

<sup>14</sup> <http://www.arpenbrasil.org.br/home.php>

reunião dessas entidades gerou algumas decisões que, em seguida, foram publicadas pela Corregedoria do Conselho Nacional de Justiça. Por intermédio do Provimento 3, 2009<sup>15</sup>, fica determinado:

**Excluir:**

- a) Da certidão de nascimento: *Declarante da certidão*;
- b) Da certidão de casamento: os itens *nome do presidente da celebração, data da celebração, documentos apresentados, profissão e domicílio*;
- c) Da certidão de óbito: os itens *profissão, data do nascimento, nome do cônjuge e nome dos filhos*.

**Substituir:**

- a) Na certidão de casamento: as expressões *nomes e prenomes dos cônjuges* por *nomes completos de solteiro dos cônjuges*.

**Incluir:**

- a) Na certidão de óbito: campo para preenchimento do *nome e o número de registro de classe do médico que atestou o óbito*, quando houver a informação.

A seguir, para ilustrar os tipos de certidões considerados neste estudo, apresento os modelos de certidões adotados em todo o território brasileiro, comumente utilizados pelos solicitantes: **Figura 1** de nascimento, **Figura 2** de casamento e **Figura 3** óbito em breve relato (resumido, em formulário) e **Figura 4** de casamento em inteiro teor (textual, cópia exata do livro de registros). Na última ilustração, a de casamento em inteiro teor, informo que, por motivo de sigilo e também por não serem relevantes em auxiliar nas descobertas deste estudo, os nomes das pessoas envolvidas no registro foram deletados.

---

<sup>15</sup> <http://www.cnj.jus.br/provimentos-atos-corregedoria/12756-provimento-no-3-de-17-denovembro-de-2009>

**Figura 1 - Modelo de Certidão de Nascimento no Brasil**

REPUBLICA FEDERATIVA DO BRASIL  
 Registro Civil das Pessoas Naturais

### CERTIDÃO DE NASCIMENTO

NOME: \_\_\_\_\_

MATRÍCULA: 9999999999 9999 9 9999 999 99999999 99

DATA DE NASCIMENTO (POR EXTENSO): \_\_\_\_\_ DIA \_\_\_\_\_ MÊS \_\_\_\_\_ ANO \_\_\_\_\_

HORA DE NASCIMENTO \_\_\_\_\_ MUNICÍPIO DE NASCIMENTO E UNIDADE DA FEDERAÇÃO \_\_\_\_\_

MUNICÍPIO DE REGISTRO E UNIDADE DA FEDERAÇÃO \_\_\_\_\_ LOCAL DE NASCIMENTO \_\_\_\_\_ SEXO \_\_\_\_\_

FILIAÇÃO

PAI \_\_\_\_\_ MÃE \_\_\_\_\_

AVÓS

AVÔ PATERNO \_\_\_\_\_ AVÔ MATERNO \_\_\_\_\_

AVÓ PATERNA \_\_\_\_\_ AVÓ MATERNA \_\_\_\_\_

GÊMEOS \_\_\_\_\_ NOME E MATRÍCULA DOS GÊMEOS \_\_\_\_\_

DECLARANTE \_\_\_\_\_

DATA DO REGISTRO (POR EXTENSO) \_\_\_\_\_ NÚMERO DA ENVILHAGEM DE NASCIMENTO \_\_\_\_\_

OBSERVAÇÕES / AVERBAÇÕES \_\_\_\_\_

NOME DO OFÍCIO \_\_\_\_\_ O conteúdo da certidão é verdadeiro. Dou fé.  
 OFICIAL REGISTRADOR \_\_\_\_\_ Data e Local: \_\_\_\_\_  
 MUNICÍPIO /UF \_\_\_\_\_  
 ENDEREÇO \_\_\_\_\_ Assinatura do Oficial \_\_\_\_\_

**Figura 2 - Modelo de Certidão de Casamento no Brasil.**

REPUBLICA FEDERATIVA DO BRASIL  
 Registro Civil das Pessoas Naturais

### CERTIDÃO DE CASAMENTO

NOMES: \_\_\_\_\_

MATRÍCULA: 9999999999 9999 9 9999 999 99999999 99

\_\_\_\_\_  
 (Nomes, prenomes, data e locais de nascimento, nacionalidade, profissão, domicílio dos cônjuges e suas filiações)

DATA DA CELEBRAÇÃO DO CASAMENTO (POR EXTENSO): \_\_\_\_\_ DIA \_\_\_\_\_ MÊS \_\_\_\_\_ ANO \_\_\_\_\_

NOME DO PRESIDENTE DA CELEBRAÇÃO \_\_\_\_\_ REGIME DE BENS DO CASAMENTO \_\_\_\_\_

RELAÇÃO DOS DOCUMENTOS APRESENTADOS AO OFICIAL DO REGISTRO \_\_\_\_\_

NOME QUE CADA UM DOS CÔNJUGES PASSOU A UTILIZAR (QUANDO HOUVER ALTERAÇÃO) \_\_\_\_\_

OBSERVAÇÕES / AVERBAÇÕES \_\_\_\_\_

NOME DO OFÍCIO \_\_\_\_\_ O conteúdo da certidão é verdadeiro. Dou fé.  
 OFICIAL REGISTRADOR \_\_\_\_\_ Data e Local: \_\_\_\_\_  
 MUNICÍPIO /UF \_\_\_\_\_  
 ENDEREÇO \_\_\_\_\_ Assinatura do Oficial \_\_\_\_\_

Figura 3 - Modelo de Certidão de Óbito no Brasil.

The image shows a blank form for a Brazilian Death Certificate (Certidão de Óbito). At the top, it features the coat of arms of Brazil and the text 'REPÚBLICA FEDERATIVA DO BRASIL' and 'REGISTRO CIVIL DAS PESSOAS NATURAIS'. The title 'CERTIDÃO DE ÓBITO' is prominently displayed. Below the title, there are fields for 'NOME', 'MATRÍCULA', 'SEXO', 'COR', 'PROFISSÃO', and 'ESTADO CIVIL E IDADE'. Further down, there are sections for 'DATA DE NASCIMENTO (POR EXTERIORES)', 'DOMICÍLIO / RESIDÊNCIA', 'NACIONALIDADE', 'DOCUMENTO DE IDENTIFICAÇÃO', and 'ELEITOR'. The form also includes fields for 'FILIAÇÃO, PROFISSÃO E RESIDÊNCIA', 'DATA E HORA DE FALECIMENTO', 'LOCAL DE FALECIMENTO', 'CAUSA DA MORTE', 'SEPULTAMENTO', 'CARTÓRIO DE CASAMENTO', 'NOME DO CÓNJUGE', and 'DECLARANTE'. At the bottom, there are fields for 'FILHOS (NOME E IDADE)', 'OBSERVAÇÕES / AVERBAÇÕES', 'NOME DO OFÍCIO', 'OFICIAL REGISTRADOR', 'MUNICÍPIO/UF', and 'ENDEREÇO'. A signature line for the 'Assinatura do Oficial' is also present.

Figura 4 - Modelo de Certidão de Casamento em Inteiro Teor, no Brasil.

The image shows a filled-out form for a Brazilian Marriage Certificate (Certidão de Casamento em Inteiro Teor). The form is titled 'CERTIDÃO (Inteiro Teor)' and includes the matriculation number '085902 01 55 2006 2 00040 192 0002504 05'. The text of the certificate is as follows:

CERTIFICO a pedido verbal de pessoa interessada, para fins de direito, que revendo em Cartório a meu cargo, o Livro de Registros de Casamentos, encontrei o termo com o seguinte teor:

Aos cinco (05) dias do mês de dezembro (12) do ano de dois mil e seis (2006), nesta cidade e Comarca de Sertãozinho, Estado do Paraná, na rua Cecília Meireles, 90, Jd. Rebelo, residência da rubente, às 20:30 horas, perante o MM. Juiz de Casamentos, Sr. José Lucas Miglizzo, cartório Oficial do Registro Civil e no final assinado, depois de legalmente habilitados e na presença das testemunhas:

(RG nº \_\_\_\_\_) e (RG nº \_\_\_\_\_) brasileiros, aquela solteira, catariense, natural de Londrina-PR, com 19 anos de idade, residente em Londrina-PR, filha de \_\_\_\_\_ e \_\_\_\_\_, esta separada judicialmente, psicóloga, natural de Sertãozinho-PR, onde reside, com 41 anos de idade, filha de \_\_\_\_\_ e \_\_\_\_\_, nascida em Londrina-PR, cidade residente, sendo pelo MM. Juiz declarados casados;

e O contraente \_\_\_\_\_ de nacionalidade brasileira, de profissão administrador, estado civil solteiro, com 22 (vinte e dois) anos de idade, nascido em Londrina-PR, aos 25 (vinte e cinco) dias do mês de outubro (10) do ano de mil novecentos e oitenta e quatro (1984), (Certidão de Nascimento do CRC de Londrina-PR, 1º Ofício, Livro nº A-191, fs. 289, nº 47.174), residente e domiciliado em Londrina-PR, na rua Santiago, 603, Bairro Bela Súpia, filho de \_\_\_\_\_ e \_\_\_\_\_, natural, ele do Distrito Federal, com 45 anos de idade, empresário, ele de São Paulo-SP, com 45 anos de idade, do lar, residentes em Londrina-PR. A contraente \_\_\_\_\_ de nacionalidade brasileira, de profissão estudante, estado civil solteira, com 20 (vinte) anos de idade, nascida em Londrina-PR, aos 01 (primeiro) dias do mês de fevereiro (02) do ano de mil novecentos e oitenta e seis (1986), (Certidão de Nascimento do CRC de Sertãozinho-PR, Livro nº A-69, fs. 05v, nº 2.686), residente e domiciliada nesta cidade, na rua Cecília Meireles, 90, Jd. Rebelo, filha de \_\_\_\_\_ e \_\_\_\_\_, natural, ele de Sertãozinho-PR, com 48 anos de idade, médico veterinário, ele de Bela Vista do Paraíso-PR, com 44 anos de idade, professora, residentes nesta cidade. O contraente em virtude do casamento, passará a usar o nome de: \_\_\_\_\_ O mesmo nome.

A contraente em virtude do casamento, passará a usar o nome de: \_\_\_\_\_ O mesmo nome.

O regime adotado é o de Comunhão Parcial de Bens.

Exibiram os documentos exigidos pelo artigo 1.525 do Código Civil números I, III e IV a saber: Prova de idade, declaração de estado, profissão e residência de cada um, bem como de seus pais e atestado de duas testemunhas habilitadas de casamento consta com parecer favorável do Ministério Público e homologação judicial. Do que para constar serve esta Termo que vai devidamente assinado pelo: Juiz, rubricadas e testemunhas do ato no título qualificados. Eu (a) Mara Cristina Galles Calevares - Agente Delegada do Serviço de Registro Civil das Pessoas Naturais que digitei e subscrevi. (aa)

N.A.D.A. M.A.I.S.: Esta conforme o original, dou fé: Sertãozinho, Estado do Paraná, aos vinte dias do mês de setembro do ano de dois mil e treze (2006/2013). Eu, \_\_\_\_\_ (Thalita Tódero Galles) - Escrevente Juramentada subscrevi.

At: \_\_\_\_\_  
**THALITA TÓDERO GALLES**  
 Escrevente Juramentada

At: \_\_\_\_\_  
**REGISTRO CIVIL**  
 EUE12051

A complexidade da vida dos cidadãos faz com que haja a necessidade de rapidez nos serviços prestados à sociedade. Com isso, amenizar a burocracia mediante a eliminação de papéis e de presença física, e diminuir o tempo, tanto de espera quanto de deslocamento, são fatores que entraram na pauta de atendimento às pessoas que necessitam certidões. A internet é uma realidade presente em muitos âmbitos e não poderia ser diferente no que se refere ao modo de como acessar os produtos de um cartório, ou seja, não é preciso ir especificamente até lá. Desenvolvido pela Associação dos Notários e Registradores do Brasil (Anoreg-Brasil<sup>16</sup>), os cartórios possuem sítios na internet e, unidos aos Correios, proporcionam comodidade, rapidez e segurança, como os exemplos a seguir:

- Cartório 24 Horas: <https://www.cartorio24horas.not.br/>
- e-Cartório: <https://e-cartoriodobrasil.com/>
- Registro Civil Online: [www.registrocivil.org.br](http://www.registrocivil.org.br)
- Certidão Online Brasil: <https://cartorioonlinebrasil24h.com.br/>
- Documento no Brasil: <https://documentonobrasil.com.br/>

## 1.2. CERTIDÕES NOS ESTADOS UNIDOS

A partir deste ponto, pretendo elucidar como funciona a emissão de certidões nos Estados Unidos da América, para que seja possível o reconhecimento das diferenças neste âmbito, tendo em mente o que ocorre em território brasileiro. Discutirei a atuação dos notários, mostrando que sua função não tem impacto significativo no âmbito das certidões, também em comparação ao que ocorre no Brasil; explanarei a respeito do “*Department of Health – Division of Vital Records*”, o departamento de registros civis do governo americano, órgão que pode ter suas competências alinhadas aos Cartórios de Registros Civis brasileiros; por fim, e em paralelo com o que foi discutido anteriormente, transcorrerei sobre cada uma das certidões americanas em pauta neste estudo.

### 1.2.1. Notários Americanos

Necessário como testemunha ou terceira parte na assinatura de documentos, o notário nos Estados Unidos não precisa estar à frente de uma serventia notarial, já que se trata de uma

---

<sup>16</sup> <https://www.anoreg.org.br/site/category/noticias/cartorio-24horas/>

profissão que pode ser realizada e administrada sem ter local fixo para isso, mas, mesmo que tenha, está autorizado a atender clientes em residências, hospitais ou qualquer outro local para onde for chamado (Doliver e Figueredo, 2012, p. 5; Lorette e Spillane, 2011, p. 15).

Cada estado regulamenta a profissão ou estabelece pré-requisitos para seu exercício da maneira que lhe aprouver, mas, em geral, um notário tem por função lutar contra fraude e roubos mediante aplicação de instrumentos que evidenciem que estava presente no momento em que o documento foi assinado pelas partes, adicionando a ele o peso de uma atividade dentro do âmbito legal (Withers, 2005, p. 10). Starsky e Crossett (2010, p. 52) explicam que o mandato de um notário tem validade de quatro anos e que o candidato deve, por exemplo, ter mais de 18 anos; ser residente no estado em questão; não ter ficha criminal; e possuir uma espécie de seguro no valor de dez mil dólares para, no caso de cometer erro que resulte em prejuízo, ter dinheiro para ressarcir a pessoa atingida. Ao contrário do que ocorre no Brasil, um notário público não está autorizado a emitir nem/ou autenticar certidões (Starsky e Crossett, 2010, 15).

### **1.2.2. Departamento de Registros Cíveis**

O Departamento de Registros Cíveis<sup>17</sup> (DRC) é um órgão do governo americano que se responsabiliza por recolher dados, manter registros e fornecer certidões a respeito de alguns eventos da vida dos cidadãos como nascimento, casamento, união civil e óbito. Sempre que as pessoas se apresentam para comunicar algum desses atos, o registro é feito em uma repartição de âmbito estadual do DRC, que se responsabiliza por fazer as remessas dessas informações a esferas superiores.

Diferentemente do que ocorre no Brasil, as certidões de nascimento e de óbito não são emitidas por uma serventia notarial, mas sim pelo DRC, e não estão disponíveis ao público em geral. Segundo Porter e Crow (2003, p. 23), depois dos ataques de 11 de setembro de 2001, muitos estados seguiram recomendações federais e revisaram suas leis sobre privacidade e disponibilidade de registros cíveis para o público e enrijeceram o acesso ainda mais. Com relação ao tempo em que ficam mantidas sob sigilo, Staple (2007, p. 58) menciona que muitos estados americanos entendem que a certidão de nascimento, por exemplo, deve se tornar pública somente depois de 100 anos, outros não permitem que sejam disponibilizadas em momento algum, e outros, ao contrário, determinam períodos mais longos ou mais curtos com 50, 75 ou 125 anos.

---

<sup>17</sup> Department of Health – Division of Vital Records

De acordo com Jones-Vassey (2012, p. 1), os registros são utilizados para projeções populacionais, entender e propor intervenções em assuntos de saúde pública, examinar indicadores como fertilidade, mortalidade, mudanças culturais que refletem na formalização de casamentos e de separações, causas de morte, e os fatores associados a eles, e, também, para implementar vigilâncias sobre, por exemplo, a conduta de laboratórios, os métodos de registros de doenças, estudos que visam a identificar a ocorrência de mortes.

Vale ressaltar a dupla função dos registros civis. De acordo com as Publicações das Nações Unidas (2001, p. 5)<sup>18</sup>, o registro civil é contínuo, permanente, compulsório e universal e tem dois propósitos: de um lado administrativo e legal, ao manter registros e fornecer certidões, do outro lado estatístico, demográfico e epidemiológico, por meio da coleta de dados e realização de estatísticas.

#### 1.2.2.1. Certidão de Nascimento

Nos Estados Unidos, para uma pessoa iniciar o processo para obter documentos como carteira de identidade, seguro social, carteira de motorista, matrícula escolar, e vários outros, é necessário que apresente uma certidão de nascimento autenticada, reconhecida como o primeiro direito fundamental do ser humano (UNICEF, 2005, p. 1).

As barreiras para se chegar ao documento ficam mais difíceis de serem transpostas dependendo de quem é o solicitante. Tudo pode ser feito rapidamente por meio eletrônico, desde que o pedido seja encaminhado pela pessoa de quem o nome conste na certidão solicitada ou o nome de seu filho menor, do contrário, somente haverá prosseguimento no processo após o envio, pelo correio, de formulários autenticados que atendam a essa demanda ou, dependendo do caso, o solicitante deverá comparecer ao DRC, visto ser obrigatório informar o motivo do interesse no documento<sup>19</sup>.

Na função de servir a propósitos legais e estatísticos, a expressão 'certidão de nascimento' pode se referir a dois tipos de papeis. O **primeiro**, de acordo com Green e Bowie (2005, p. 125), diz respeito a um longo documento preenchido pelo médico, ou outro representante do hospital, descrevendo informação dos pais e do bebê como, por exemplo, profissão, grau de escolaridade, etnicidade e raça e informações sobre a gravidez como fatores de risco, complicações e/ou condições anormais do recém-nascido. Em seguida, esse

---

<sup>18</sup> <http://unstats.un.org/unsd/default.htm>

<sup>19</sup> <http://www.azdhs.gov/vital-records/birth-certificate/obtain.htm>

documento é encaminhado para registro e arquivamento<sup>20</sup>. A certidão de nascimento arquivada no Departamento de Registros Cíveis segue padrão de formulário, emitido pelo governo federal. Ainda conforme os mesmos autores (2005, p. 125), o **segundo** se refere à cópia autenticada, uma certidão mais curta que atesta a existência do nascimento a partir dos registros arquivados, entregue ao solicitante. Para que seja legitimada, essa cópia deve possuir chancela do registrador em alto relevo, selo impresso ou multicolorido, assinatura do registrador e a data em que a certidão foi preenchida. Staples (2007, p. 55) explica que a certidão de nascimento do primeiro tipo de documento mencionado acima possui duas seções: a **primeira** é a respeito do registro legal de nascimento, incluindo elementos como o nome da criança, sexo, hora e local do nascimento, nome atual e de solteira da mãe, o nome do pai, sua data e local de nascimento e o número do seu seguro social, por exemplo. Staples (2007, p. 56-7) informa que cópias autenticadas a partir desta parte da certidão arquivada são utilizadas para “[...] propósito de emprego; para a obtenção de outros documentos como carteira de motorista, passaportes e cartão de seguridade social; para matricular crianças na escola; para determinar elegibilidade em benefícios públicos [...]”<sup>21</sup> (Tradução minha). A segunda parte da certidão relata condições de saúde e informações médicas. O autor esclarece que a segunda parte é somente usada para coleta e análises de dados para fins estatísticos.

Como visto, as certidões americanas possuem cláusulas de confidencialidade recomendadas pelo governo federal dos Estados Unidos e acatadas, com ajustes ou não, pelos seus estados. Para que, neste estudo, seja possível a utilização de certidões americanas, sem esbarrar situações legalmente amparadas, seguirei a instrução de Staples (2007, p. 58) ao afirmar que “não há restrições, contudo, na liberação de informações que não identifiquem os indivíduos”.

Diante dessa afirmação, e por motivos éticos, informo que, para que não haja violação de cláusulas de confidencialidade neste trabalho, todos os nomes dos indivíduos interessados, ou seja, os constantes nos registros que deram origem às certidões, serão retirados. Isso acontecerá não apenas nas certidões de nascimento americanas, como a **Figura 5** a seguir, mas de todas as certidões desta pesquisa, para que sejam mantidos além do alcance de qualquer reconhecimento.

---

<sup>20</sup> [http://www.cdc.gov/nchs/nvss/vital\\_certificate\\_revisions.htm](http://www.cdc.gov/nchs/nvss/vital_certificate_revisions.htm)

<sup>21</sup> [...] for employment purposes; for obtaining other documents, such as driver’s license, passports, and Social Security cards; for enrolling children in school; for determining eligibility for public benefits [...].

**Figura 5 - Modelo de Certidão de Nascimento Americana.**

STATE OF OHIO  
OFFICE OF VITAL STATISTICS

CERTIFICATION OF BIRTH

STATE FILE NUMBER 1937056911 DATE RECORD FILED 09/07/1937

NAME [redacted]

DATE OF BIRTH 04/23/1937 SEX Male

BIRTHPLACE OHIO

MOTHER'S NAME [redacted] FATHER'S NAME [redacted]

MAIDEN NAME [redacted]

MOTHER'S BIRTHPLACE KENTUCKY FATHER'S BIRTHPLACE OHIO

Note: This is a true certification of the name and birth facts as recorded in the Office of Vital Statistics, Columbus, Ohio. Witness my signature and seal of the Department of Health this 07 day of October, 2014

[Signature]  
State Registrar of Vital Statistics

H 4 24 297 2

CENTRAL LOCATION

VOID WITHOUT WATERMARK OR IF ALTERED OR ERASED  
VERIFY PRESENCE OF ODH WATERMARK HOLD TO LIGHT TO VIEW

Fonte: NAZDAR – Digital Genealogy Archives<sup>22</sup>.

#### 1.2.2.2. Certidão de Casamento

Nos Estados Unidos, explica Funaro (1991, p. 3), o casamento pode ser realizado de duas maneiras: de acordo com o cerimonial ou sem o cerimonial. O **primeiro** tipo é dividido em duas partes: 1) o casal deve solicitar uma licença para realização do casamento e, 2) após a realização da cerimônia, deve retornar ao órgão expedidor da licença para que ela seja registrada e certificada. O **segundo** tipo é uma oportunidade para que o casal regularize diante da justiça a maneira com a qual já se apresenta para a comunidade, ou seja, transformar uma União Estável<sup>23</sup> em casamento.

Para obter uma licença para se casar, não basta apenas ir até o registrador e preencher uma solicitação. Na realidade, para obter o documento, existem alguns requisitos a serem cumpridos para que a licença seja autorizada. Sardone (2007, p. 12) informa que “[a]s leis e requisitos para licença de casamento variam de estado para estado e de localidade para

<sup>22</sup> <http://nazdar.org/>

<sup>23</sup> Common-Law Marriage.

localidade nos Estados Unidos<sup>24</sup>” (Tradução minha). No Estado do Texas<sup>25</sup>, por exemplo, existem as seguintes condições:

- Os noivos devem comparecer ao setor de registros juntos e provar idade e identidade.
- Além da carteira de identidade, de motorista, ou passaporte, cada um deve apresentar um documento que mostre seu nome e o número do seguro social.
- Pagar taxa de US\$ 82,00. Se o casal tiver participado de curso de aconselhamento pré-nupcial ofertado por órgão competente do Texas, a taxa passa a ser de US\$ 22,00. Obviamente, o valor da taxa não é o mesmo em todas as localidades. Na cidade de Nova Iorque<sup>26</sup>, por exemplo, a taxa é de US\$ 32,00.
- Ser maior de 18 anos. Com idade entre 16 e 18 é necessária a presença dos pais, ou pessoa legalmente autorizada. Se um dos dois nubentes for divorciado, deve provar que essa condição está legalmente resolvida e prover um certificado mostrando a mudança de nome, se diferente do que estiver na identificação.
- Licença válida por 90 dias, mas há um período de espera de 72 horas, a partir da concessão da licença, para a realização do casamento. Depois do prazo, mesmo que realizado, o casamento não vai ser autenticado. Para regularizar a situação, um novo processo de solicitação de licença deve ser iniciado. Na cidade de Nova Iorque vale por 60 dias e no estado de Nevada<sup>27</sup> não há período de espera entre o dia do pedido e a realização da cerimônia de casamento, por exemplo.

O tempo útil de uma licença também varia de acordo com a localidade em questão<sup>28</sup>, por exemplo: pode ser de 30 dias, como no Alabama e no Havaí, chegando a 60 dias em Illinois, na Flórida e em Nova Iorque, como mencionado anteriormente, subindo para 90 dias no Alaska e na Califórnia, podendo chegar a 1 ano, como no Arizona e em Nebraska.

No Estado da Califórnia há dois tipos de Licença para casamentos: pública ou confidencial<sup>29</sup>. Quando for **pública**, a licença se assemelha ao que já foi descrito para o Estado do Texas, mas pode não haver total simetria nos requisitos como, por exemplo:

- Menores de 18 anos dependem da autorização de pelo menos um dos pais, ou guardião legal, e da autorização do juiz da Corte Superior da Califórnia, mesmo que sejam emancipados.

<sup>24</sup> Marriage license laws and requirements differ from state to state and locality to locality in the United States.

<sup>25</sup> <http://tools.cira.state.tx.us/users/0034/docs/Clerk/Marriage%20License%20Application%20Requirements.pdf>

<sup>26</sup> <http://www.cityclerk.nyc.gov/html/marriage/license.shtml#intro>

<sup>27</sup> <http://www.clarkcountynv.gov/depts/clerk/services/pages/marriagelicense.aspx>

<sup>28</sup> <http://www.usmarriagelaws.com/>

<sup>29</sup> <http://www.cdph.ca.gov/certlic/birthdeathmar/Pages/TypesofMarriageLicenses.aspx>

- Pode-se adquirir a licença em uma comarca e se casar em qualquer outra. Entretanto, o casamento deve ser realizado dentro do próprio estado e o registro na comarca onde a licença foi solicitada.
- Neste caso a licença é pública, ou seja, qualquer pessoa pode, mediante o pagamento de taxa, requer uma cópia desse registro.
- É preciso haver pelo menos uma testemunha e não mais que duas. A assinatura de três, anula a licença.

No caso de ser uma licença **confidencial**, algumas diferenças e procedimentos podem ser observados:

- Não pode ser solicitada por menores;
- Os solicitantes devem estar morando juntos como esposos quando fizerem a solicitação e assinar um documento que ateste isso. É o caso do casamento sem cerimonial, como mencionado no início desta parte do capítulo;
- Também podem se casar em outra comarca, desde que dentro do Estado da Califórnia e que realizem o registro na comarca onde solicitou a licença;
- Não há testemunhas;
- O registro é confidencial, isto é, somente o casal pode solicitar cópias, mediante a apresentação de documento com foto e pagamento de taxa.

Após a realização da cerimônia, explica Wilson (2012, p. 1163), o sacerdote, ou autoridade competente que conduziu o evento, preenche os espaços em branco na parte inferior da licença a respeito da efetivação do casamento, deixados especificamente para esse fim. Em seguida, noivos, sacerdote e, em alguns lugares, as testemunhas afixam suas assinaturas. Esse documento é levado ao registrador para que uma certidão de casamento seja expedida, ou para que a licença seja endossada, registrada e passe a ter valor de certidão.

De acordo com Powell (2012, p. 102), nos Estados Unidos, é muito comum que a certidão de casamento seja registrada no mesmo documento da licença de casamento, já que, conforme mencionado anteriormente, muitos Estados estabelecem uma parte da licença de casamento para ser preenchida no dia da cerimônia, o que também pode incluir espaços para indicar a mudança nos nomes dos noivos, se houver. Depois disso, as certidões de casamento são mantidas nas Cortes Superiores dos Condados onde o evento ocorreu, ou em outro órgão local competente como o Departamento do Escrivão da Cidade, e não no DRC<sup>30</sup>.

---

<sup>30</sup> <http://www.azdhs.gov/vital-records/marriage.htm>

Bix (2013, p. 42) afirma que existem leis para casamento em todos os Estados americanos e, também, que em quase todos, a obtenção de licença é um pré-requisito para que a união seja validada. Segundo o autor, a licença espelha o que é estipulado pela lei do Estado em questão, revelando liberdades e restrições para que os interessados cheguem à certidão de casamento legalmente válida. Nesse âmbito, certas modalidades de casamento podem ser aceitas, ter regulamentação específica para serem aprovadas, ou podem ser proibidas. Aqui, podemos citar o casamento entre primos, entre pessoas do mesmo sexo, a partir de União Estável e casamento por procuração. A título de exemplificação, relataremos o comportamento de alguns estados ao se depararem com essas modalidades<sup>31</sup>:

A seguir, apresento a **Figura 6**, um modelo de certidão de casamento americana do Estado do Colorado que conta, na parte inicial, com a licença para casamento e, logo em seguida, com a certidão de casamento devidamente preenchida, retornando para registro.

**Figura 6 - Modelo de Certidão de Casamento Americana.**

FOR RECORDING USE

FAYE GRIFFIN JEFFERSON COUNTY, CO  
LICENSE #: 0002399  
DATE RETURNED: 10/13/2005 05:19 PM

20052403

**MARRIAGE LICENSE**  
THE STATE OF COLORADO.  
COUNTY OF JEFFERSON  
GOLDEN, COLORADO 80419

To any person or religious society, Authorized, by Law to Perform the Marriage Ceremony:

**GREETING:**  
You are hereby authorized to join in marriage \_\_\_\_\_ and \_\_\_\_\_  
and of this license you will make due return to my office.  
Witness, my hand and the seal of my office at Golden, Colorado, this 28<sup>th</sup> day of September, 2005

JEFFERSON COUNTY COLORADO SEAL

FAYE GRIFFIN County Clerk & Recorder  
By: [Signature] Deputy  
Issued at 09:44 o'clock A.M.

**MARRIAGE CERTIFICATE**

STATE OF COLORADO, ss.  
County of JEFFERSON

It is hereby certified that on the 9<sup>th</sup> day of October A.D. 2005 at Mount Vernon Country Club, Golden, COLORADO in said county, the undersigned, a Minister of the Universal Life Church did join in Matrimony in accordance with the laws of the state of Colorado and the authorization of the foregoing license

Groom: \_\_\_\_\_  
Bride: \_\_\_\_\_  
Witness, my hand and seal the day and year last above written

Fonte: Acervo pessoal.

<sup>31</sup> <http://www.usmarriagelaws.com/>

### 1.2.2.3. Certidão de Óbito

Nos Estados Unidos, a expressão ‘certidão de óbito’ também pode ser entendida a partir de dois significados: 1) documento expedido pelo médico certificando que a pessoa faleceu; ou 2) um documento expedido pelo registrador do DRC, declarando a data, local e causa da morte de uma pessoa<sup>32</sup>. A certidão de óbito não é disponível ao público em geral, mas pode ser imediatamente adquirida pelo cônjuge ou por parente de sangue do falecido, bastando, para eles, entrar em contato com o DRC do estado onde o ato está assentado, preencher a requisição, pagar as taxas e aguardar a chegada da cópia autenticada, pelo correio. Quando não se tratar das pessoas mencionadas, é preciso comparecer ao DRC e formalizar o pedido. Nesse momento, o solicitante deve esclarecer o motivo pelo qual deseja obter uma cópia dessa certidão<sup>33</sup>. Finkbeiner, Ursell e Davis (2009, p. 147) deixam claros os propósitos para a existência de uma certidão de óbito ao afirmar que

[a] certidão de óbito serve para dois propósitos: legal e estatístico. Legalmente, as certidões de óbito contribuem para o registro de óbito e são comumente utilizados em questões médico-legais, opções de sepultamento, seguro e assuntos de herança. Estatisticamente, a certidões de óbito são amplamente utilizadas em estudos epidemiológicos e de saúde pública<sup>34</sup> (Tradução minha).

Igualmente ao que ocorre com as outras certidões mencionadas, o *National Center For Health And Statistics* (1991, p. 4-5) informa que “cada estado é independente para determinar a forma e o conteúdo dos seus registros civis”, instrução que pode, de acordo com as necessidades locais, propiciar que cada um tenha seu próprio formulário de certidão de óbito. Esse entendimento possibilita a existência de elementos diferentes do padrão recomendado pelo governo federal, mas, de maneira geral, tanto seu formato quanto os procedimentos para se chegar a ela são homogêneos como, por exemplo, requerer a certidão preenchida pelo médico para validar a causa da morte<sup>35</sup>. De acordo com Green e Bowie (2005, p. 128), os itens que fazem parte do conteúdo desse documento são:

1. Nome do falecido;

<sup>32</sup> <http://www.nyc.gov/html/doh/html/services/vr-death-general.shtml>

<sup>33</sup> <http://www.azdhs.gov/vital-records/death-certificate/obtain.htm>

<sup>34</sup> Death certificates serve two primary purposes: legal and statistical. Legally, death certificates contribute to the record of death and are commonly used in medicolegal, interment, insurance and inheritance matters. Statistically, death certificates are widely used in epidemiologic and public health studies.

<sup>35</sup> [https://www.cdc.gov/nchs/nvss/revisions-of-the-us-standard-certificates-and-reports.htm?CDC\\_AA\\_refVal=https%3A%2F%2Fwww.cdc.gov%2Fnchs%2Fvital\\_certificate\\_revisions.htm#print](https://www.cdc.gov/nchs/nvss/revisions-of-the-us-standard-certificates-and-reports.htm?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fnchs%2Fvital_certificate_revisions.htm#print)

2. Data do falecimento e local de nascimento;
3. Endereço residencial na hora da morte;
4. Causa da morte;
5. Local de enterro;
6. Nomes e locais de nascimento de ambos os pais;
7. Nome do informante (geralmente um parente);
8. Nome do médico;
9. Método e lugar para depositar o corpo;
10. Assinatura do diretor funeral;
11. Assinatura de um médico autorizado.

A seguir, apresento a **Figura 7**, uma certidão de óbito resumida do Estado do Maine.

**Figura 7 - Modelo de Certidão de Óbito Americana.**

STATE OF MAINE		
CERTIFIED ABSTRACT OF A CERTIFICATE OF DEATH		
DEPARTMENT OF HUMAN SERVICES		
DECEDENT PERSONAL DATA	FULL NAME OF DECEASED	DATE OF DEATH
	PLACE OF DEATH	Feb. 6, 1934
	Shawmut, Me	DATE OF BIRTH
	SOCIAL SECURITY NUMBER	SEX
		M
	FATHER'S NAME	MOTHER'S MAIDEN NAME
INFORMANT	NAME OF INFORMANT	ADDRESS
CAUSE(S) OF DEATH	Cerebral Hemorrhage	
CERTIFICATION	NAME OF PHYSICIAN OR MEDICAL EXAMINER CERTIFYING DEATH	DATE SIGNED
	W. L. Grousse M.D.	1934
DISPOSITION	NAME OF CEMETERY OR CREMATORY	CITY OR TOWN
		STATE
	NAME OF CLERK RECORDING THIS DEATH	CITY OR TOWN
	C.C. Porter	Castle Hill
		DATE OF FILING
I HEREBY CERTIFY THAT THE FOREGOING IS A TRUE ABSTRACT OF A CERTIFICATE OR RECORD WHICH IS IN MY OFFICIAL CUSTODY.		
ATTEST:	<i>Sue Dugan, Abst. Sec.</i>	MUNICIPAL CLERK
DATE ISSUED:	01/16/01	TOWN OF, Castle Hill
This copy not valid unless prepared on engraved border displaying seal and signature of Registrar.		

Fonte: NAZDAR – Digital Genealogy Archives<sup>36</sup>.

<sup>36</sup> <http://nazdar.org/>

Da mesma forma que ocorre no Brasil, também é possível obter certidões de nascimento, casamento e óbito americanas, sem mesmo sair de casa, por meio da internet. A seguir, menciono alguns exemplos de sítios disponíveis ao usuário:

- Vital Check: <https://www.vitalchek.com/v/>
- National Center for Health Statistics: <https://www.cdc.gov/nchs/w2w/index.htm>

Cada estado americano possui um *site* para a obtenção de certidões, por exemplo:

- Alabama Public Health: <https://www.alabamapublichealth.gov/>
- Florida Department of Health: <https://www.floridahealth.gov/>
- California Department of Public Health: <https://www.cdph.ca.gov/>

### 1.3. OS DESAFIOS NA TRADUÇÃO DE CERTIDÕES

Tradutores juramentados desempenham seu ofício em grande proximidade com órgãos públicos, tanto jurídicos quanto em vários outros âmbitos, governamentais ou não, em que a utilização de documentos “com fé pública” se faz necessária. Justamente por trabalhar sob essa condição, quando comparado a outros profissionais não juramentados, ainda mais responsabilidade é atribuída ao processo e ao produto de sua atividade. Acredita-se que ele detém grande poder, visto que lidar com documentos que podem trazer mudanças nas vidas das partes em questão, tanto de forma positiva quanto negativa, é procedimento corriqueiro. De acordo com Asensio (2003), a responsabilidade também é bastante alta em outros tipos de tradução, mas que, no modo juramentado, a proximidade com aspectos legais pode colocar o tradutor numa posição ainda mais delicada (p. 43). O profissional juramentado condiciona-se a obedecer às mesmas leis que todos os outros cidadãos, mas, quando comparado aos outros integrantes da sociedade, pode-se entender que o simples fato de ser tradutor na modalidade juramentada fomenta a expectativa de que um texto vindo de suas mãos não deve possuir erros, falhas, incompreensões e nem qualquer outra lacuna que possa deixá-lo exposto a objeções. A esse profissional, portanto, é vedado transgredir as regras, mesmo que não intencionalmente. O trabalho juramentado, por estar calcado na sensação de que abarca mais poder, pode tornar o tradutor ainda “mais culpado” quando as regras deixam de ser cumpridas. Nesse sentido, a sensação que traz mais poder aos procedimentos é a mesma que trará mais intensidade nos rigores legais sobre o tradutor se lacunas, contrariedades ou objeções, por exemplo, forem localizadas nesses mesmos procedimentos.

Outra restrição ao desempenhar uma tradução juramentada, com documentos de toda sorte, diz respeito a fatores éticos, que, nas palavras de Asensio (2003), discutidos ao longo deste parágrafo, podem ser **individuais** ou **coletivos**. O autor explica que os princípios éticos individuais dizem respeito às tomadas de “decisões do tradutor, tendo em mente o que considera certo e errado” (p. 44) e não extravasam os limites pessoais. Por outro lado, os coletivos, ou deontológicos, são códigos que regem um grupo e se localizam nas “associações profissionais e suas decisões a respeito do que consideram certo e errado” (p. 44). Nesse caso, o tradutor se submete aos códigos determinados pelo grupo, ou associação a que pertence, sob pena de ser expulso dessa posição em caso de não cumprimento das regras estipuladas. De acordo com a Associação Profissional dos Tradutores Públicos e Intérpretes Comerciais do Estado de São Paulo (ATIESP<sup>37</sup>), esses códigos tentam esclarecer, por exemplo, as responsabilidades do tradutor; a conduta que deve ter em mente ao realizar um trabalho; a postura esperada diante do cliente e/ou do colega; o tipo de texto que pode aceitar ou recusar; o tipo de conhecimento que precisa ter ao traduzir; entre outros. Em resumo, os preceitos éticos coletivos tentam exercer papel de norteadores de aspectos referentes ao dia-a-dia de um tradutor no desempenho de seu trabalho, explicitando o que é permitido e esperado, e o que não é.

Ainda permeado pelas regulamentações profissionais, Asensio (2003, p. 49) menciona que “a maioria das normas consideradas compulsórias são, na verdade, nada mais que usos e costumes seguidos por um significativo número de tradutores<sup>38</sup>” (Tradução minha). Nesse âmbito, Aubert (2012, p. 5) menciona que existe

[...]a constatação de que - ao contrário do que poderia parecer aos olhos do leigo, nem tudo - na realidade, quase nada - encontra-se regulamentado na prática da tradução juramentada. Para além de decretos e instruções normativas de teor relativamente genérico e formal, há um conjunto de praxes, de hábitos, de tendências e de costumes, alguns mais consensuais, outros menos, ditados quer pelo bom senso, quer pela tradição.

Asensio (2003, p. 49) vai além, e esclarece que traduzir de maneira mais conservadora é um traço comum a muitos tradutores por acreditarem que, dessa forma, podem reduzir os riscos jurídicos para si mesmos e para o trabalho, ao mesmo tempo em que atendem a expectativa do cliente e do órgão a que a tradução se destina. O autor defende que essas regulamentações

<sup>37</sup> <http://www.atiesp.org.br/associacao/tabelas-e-regulamentos/codigo-de-etica/>

<sup>38</sup> Most of the norms considered to be compulsory are, in fact, no more than uses and customs followed by a significant number of translators.

deveriam ser encaradas como provenientes de um hábito tradutório e não de uma vigente obrigatoriedade.

Ainda a respeito de fatores e influências externas à própria tradução, que podem funcionar como dificuldade, restringindo o seu desenvolvimento, é prudente não deixar de lado os que outros participantes no ato tradutório inculcem, ou tentam inculcir, no trabalho a ser realizado. Ao contratar um tradutor, o cliente pode se posicionar de maneira a exercer certa influência, almejando que a tradução seja desempenhada de maneira a lhe viabilizar mais facilmente as benesses referentes ao objetivo que tenta atingir. Claramente, a intenção do contratante de sutilmente embutir uma seleção lexical capaz de alterar a realidade expressa no texto esbarra em princípios éticos pessoais e deontológicos, e esse tipo de comportamento não deve ser aceito pelo tradutor. Do outro lado dessa relação está o “recipiente”, ou seja, o órgão ou instituição a quem o resultado tradutório se destina. Justamente por saber que documentos tendenciosos podem ser apresentados, é comum que o recipiente prefira que a tradução seja a mais literal possível, muitas vezes, ignorando que tal procedimento pode resultar em um trabalho inadequado (Asensio, 2000, p. 49).

Como visto, existem várias regras que incidem sobre a tradução juramentada, causando dificuldades e restrições ao tradutor e que, de acordo com Asensio (2003, p. 50), não devem ser rigidamente observadas. Ainda segundo o mesmo autor (2003),

[e]xiste tensão na tradução oficial entre ser adequado aos padrões de qualidade profissionais e ser aceitável para as autoridades públicas, da mesma forma que existe tensão entre credibilidade e fácil identificação por um lado, estilo e plausibilidade por outro. Cada tradutor tem seu próprio critério para resolver esse dilema e não se devem esperar conclusões uniformes<sup>39</sup> (p. 50) (Tradução minha).

De acordo com o discutido até este momento, pode-se entender que muitos são os fatores externos que incidem no desempenho fluente de um processo tradutório e, nesse sentido, antever os obstáculos pode guiar o tradutor por caminhos mais seguros na tarefa de ajustar seu trabalho ao ambiente adverso em que está inserido.

Muito já foi abordado a respeito da tentativa de alinhar a tradução ao seu texto fonte. Por entender que essa atitude é infrutífera, o tradutor se depara com uma vasta gama de lacunas a serem preenchidas no texto de chegada, o que demanda conhecimento dos aspectos envolvidos, adequação aos parâmetros de aceitação e criatividade. Nesse contexto, segundo

---

<sup>39</sup> There is tension in official translation between being adequate to professional quality standards and being acceptable to the public authorities, just as there is a tension between the faithfulness and easy identification on the one hand, and style and plausibility on the other. Each translator has their own criteria for solving this dilemma and you should not expect uniform conclusions.

Asensio (2003, p. 50), o ato de traduzir não é simplesmente comparar línguas, mas sim um ato comunicativo e criativo. Por outro lado, de acordo com o mesmo autor, o que ocorre não é uma total liberdade de ação, visto que a simples postura de observar os parâmetros macrotextuais ou microtextuais envolvidos pode funcionar como fator que reduz a quantidade de possíveis escolhas tradutórias. Os fatores que podem causar a mencionada redução de escolhas, de acordo com Asensio (2003, p. 51), são as:

- restrições de significado: no texto fonte e no texto de chegada;
- restrições textuais: tipo, função, conteúdo (no texto fonte e no de chegada);
- restrições tradutórias: dependendo da situação social do ato de traduzir, lei e regulamentações, etc.;
- restrições éticas em vários níveis;
- restrições comunicativas: comunicadores, objetivo, veículo, meios e sinais;
- restrições do trabalho: prazos, preço, outras considerações mercantis;
- restrições conceituais: dependendo do conceito de tradução seguido pelo tradutor;
- restrições de habilidade: habilidades do tradutor;
- restrições de risco: os riscos assumidos pelo tradutor<sup>40</sup> (Tradução minha).

Com vistas a padronizar o texto e acomodá-lo em literalidade, as restrições têm início nas pressões sociais, nas exigências das partes envolvidas no processo de tradução, nas considerações éticas, nos aspectos legais, e nas circunstâncias inerentes ao próprio texto.

Uma certidão pode conter elementos textuais e/ou semióticos passíveis de acarretar dificuldades no momento de serem trazidos para a cultura alvo por meio de uma tradução, e é sobre esses elementos específicos que se discute a partir de agora.

Ao lidar com documentos oficiais, é comum ao tradutor se deparar com formulários, tabelas e gráficos. Percebe-se no texto fonte um cuidado em ajustar o que é necessário ser dito ao pouco espaço ofertado. Com isso, o autor recorre às abreviaturas, contrações ou, simplesmente, reduções, para que o documento fique completo.

Criados para informar que o documento em que está afixado é autêntico, os itens semióticos têm sua presença estabelecida de maneira variada na comparação entre países,

---

<sup>40</sup> Factors that reduce the number of choices in translation (translation constraints) are:

- meaning constraints: both in source and target texts;
- textual constraints: type, function, content (both in source and target texts);
- translation constraints: depending on the social situation of the act of translation, law and colligate regulations, etc.;
- ethical constraints at various levels;
- communicative constraints: communicants, goal vehicle, means and signals;
- assignment constraints: depending on the concept of translation held by the translator;
- ability constraints: the translator's skills;
- risk constraints: the risks assumed by the translator.

portanto, nem todos utilizam os mesmos tipos de recursos em um mesmo tipo de documento (Nascimento, 2006, p. 46-47).

Ao tratar de certidões de nascimento, de casamento e de óbito, Asensio (2003, p. 88) descreve outros problemas específicos quando se trabalha com esses documentos. De acordo com o autor, esses textos podem servir a outros propósitos além de somente informar que houve um nascimento, casamento ou morte. Existem países em que o registro desses eventos é feito de maneiras diferentes, o que pode resultar em: 1) inclusão de informações relevantes somente para a cultura fonte; 2) existência de nomenclaturas adversas para denominar os órgãos competentes; 3) interesse em coleta de dados para o desenvolvimento de estudos estatísticos; 4) presença de dados médicos; 5) inclusão de detalhes sobre linhagem, casta; 6) registro da afiliação religiosa; e 7) utilização de sequência e significados diferentes para nomes e sobrenomes. A esse respeito é justo dizer que para nomear uma pessoa existem vários sistemas a serem seguidos, e cada país tem o seu. No Brasil, por exemplo, o sistema permite que haja nomes simples ou compostos e a utilização de um ou vários sobrenomes, havendo a expectativa de que o último sobrenome seja sempre o do pai. Em vários outros países a sequência pode não significar o mesmo que aqui e, muitas vezes, nem existindo distinção entre o que é nome e sobrenome. O problema persiste num eventual casamento, quando se adquire um nome de casado(a), visto que a legislação de cada país permite variadas combinações entre os sobrenomes dos nubentes, possibilidades legais que podem ser, muitas vezes, nulas perante a tradição. A respeito disso, e voltando a citar o Brasil como exemplo, a legislação permite que somente a mulher acrescente o sobrenome do marido, que somente o marido acrescente o sobrenome da mulher, que ambos acrescentem os sobrenomes um do outro aos seus, ou que ambos mantenham os nomes de solteiros. Pode-se perceber, mesmo que remotamente, a ocorrência desta última categoria, mas, tradicionalmente, a esposa é quem acrescenta o sobrenome do marido ao seu, procedimento aceito com mais naturalidade que os outros.

### **1.3.1. Solucionando Os Desafios Tradutórios**

Em uma tradução, quando há interesse em se elencar elementos valorizados pelos integrantes da cultura de chegada, para que a considerem aceitável, o interessado pode ficar surpreso ao se deparar com reduzidos requisitos em vez de extensa lista. Nesse âmbito, para que traduções sejam consideradas ‘bem feitas’, um critério relevante é o de justamente não parecerem ser traduções, isto é, que o texto proporcione leitura fluente (Venutti, 1995; Reiss, 2000), desprovido de interferências ou lacunas. Ao mesmo tempo, ainda que supostamente

antagônico à ideia de fluência, os usuários também parecem alimentar a expectativa de que o trabalho do tradutor esteja firmemente calcado no texto de partida, critério sem o qual a credibilidade do texto traduzido pode ser posta em discussão.

#### 1.4. CERTIDÕES TRADUZIDAS: O MODO JURAMENTADO

Em Tradução Juramentada (doravante TJ), ter sólida proficiência nas línguas envolvidas no projeto tradutório pode não ser o bastante para realizar um trabalho competente. Para o tradutor juramentado, além de intimidade com aspectos sociais, históricos e culturais do par linguístico em questão, o conhecimento técnico do material a ser traduzido e seu funcionamento na língua de chegada são fatores determinantes que um profissional eficiente deve levar em consideração.

Utilizada para validar o original e calcada em ambiente que denota cuidados e praxes específicos, a TJ tenta suprir o espaço existente na comunicação entre países, visando a estabelecer claras relações comerciais, sociais e jurídicas (Barros, Babini e Aubert, 2010, p. 234). Em face dessa amplitude, torna-se necessário aumentar a investigação sobre essa modalidade tradutória, visando a descobrir como tradutores públicos expressam suas escolhas ao se depararem com os limites de uma área que não parece abrir espaço para que intercedam.

##### 1.4.1. O Tradutor Juramentado

De acordo com Asensio (2003, p. 5), pelo mundo, o nome para o tradutor que realiza a tradução de documentos, para que sejam oficialmente reconhecidos, pode ser, por exemplo, Tradutor Autorizado pelo Estado (Dinamarca), Tradutor Especialista Judicial (França), Tradutor Certificado (Marrocos), ou Tradutor Público Juramentado (Peru). No Brasil, depois de passar por um concurso público, o profissional que se encarrega desse trabalho pode ser chamado simplesmente de Tradutor, de Tradutor e Intérprete Comercial, de Tradutor Público e Intérprete Comercial, ou de Tradutor Juramentado. Essa variação é encontrada entre os estados brasileiros, mas é apenas terminológica, visto que na prática, esses profissionais desempenham o mesmo trabalho. Com isso, é plausível aceitar que se trata da mesma operação, área e finalidade e que, portanto, a questão terminológica envolvendo o nome dessa atividade não se aplica ao objetivo aqui proposto. Por esse motivo, neste trabalho, o profissional que desenvolve traduções, para que sejam oficialmente reconhecidas, será denominado Tradutor Juramentado,

e o documento produzido com o objetivo de ser reconhecido e aceito em alguma instância, Tradução Juramentada.

Quando o assunto é TJ, o direcionamento inicial pode estar na busca de definições que evidenciem a abrangência dessa modalidade tradutória, ao mesmo tempo em que viabilizem o entendimento sobre o que a diferencia das outras modalidades não juramentadas. Dentro dessa proposta, algumas fontes que corroboram escopo e atuação da TJ foram encontradas, como apresentadas a seguir. A primeira, de acordo com Asensio (2003, p. 1), que prefere chamar a TJ de tradução oficial, menciona que os trabalhos desenvolvidos nessa área podem ser definidos, grosso modo, como “traduções que atendem os requisitos para servirem como instrumentos legalmente válidos numa cultura alvo<sup>41</sup>” (Tradução minha). De um ponto de vista acadêmico, e não tão generalizada quanto a anterior, Aubert (1998a, p. 14) assevera que por TJ entende-se

a tradução de textos – de qualquer espécie – que resulte em um texto traduzido legalmente reconhecido como uma reprodução fiel do original (com *fé pública*). Esta característica de fidelidade, por sua vez, significa que, por meio de tal tradução, o texto original, expresso em um idioma estrangeiro, torna-se capaz de produzir efeitos legais no país da língua de chegada e, ainda, que tal tradução é correta, precisa, exaustiva e semanticamente invariante em relação ao original (obviamente, dentro dos limites dos meios de expressão disponíveis nas respectivas línguas/culturas que se confrontam no ato tradutório específico).

No Brasil, existem escritórios que divulgam a realização de trabalhos tradutórios tanto juramentados quanto não juramentados, o que costumam denominar de traduções “simples”. Aceitam textos em vários idiomas e explicam ao contratante a diferença entre uma tradução juramentada e outra que não é. A Andrart Traduções<sup>42</sup>, por exemplo, veicula que

[a] tradução juramentada consiste na tradução feita em formato apropriado para ter validade oficial e legal perante órgãos e instituições públicas; é impressa em papel timbrado personalizado do Tradutor Público Juramentado, que assina a tradução, cuja assinatura pode ser reconhecida em cartórios e repartições consulares no Brasil a fim de ser oficialmente reconhecida no exterior.

Diferente do que acontece em outros países, onde o tradutor juramentado se liga diretamente ao Ministério da Justiça, das Relações Exteriores, a cartórios, ou a alguma entidade ou pessoa que detém o poder ou alguma jurisdição (Asensio 2003, p. 22), no Brasil esse tradutor está subordinado unicamente à Junta Comercial do estado em que reside, instituição que o

<sup>41</sup> [...] as translations that meet the requirements to serve as legally valid instruments in a target country.

<sup>42</sup> <http://www.andrart.com/htm/serv.htm>

classifica como Agente Auxiliar do Comércio. Em alguns estados também é chamado de Tradutor Oficial e Intérprete Comercial, podendo ser concursado/empossado ou “ad hoc”, o que acontece quando um concurso para a contratação de um tradutor ainda não foi realizado, ou quando o escopo de concursados não se estende a determinados idiomas que, por ventura, precisem de reconhecimento na cultura alvo. Nesse caso, o consulado, embaixada, ou a Junta Comercial nomeia um tradutor com proficiência comprovada no idioma em questão, somente para um determinado trabalho que, depois de realizado, não vincula o profissional àquelas instituições.

O âmbito do trabalho de um tradutor juramentado em um país pode variar quando comparado ao que um colega de outro país está autorizado a fazer. Asensio (2003, p. 6) menciona que, em alguns lugares, o tradutor somente tem habilitação para trabalhar com documentos escritos e que, em outros, pode fazer tradução oral. Do mesmo modo acontece quanto à direção do par linguístico em questão: alguns podem traduzir nas duas direções, enquanto outros, somente em direção à língua materna.

Subordinadas administrativamente aos governos estaduais e tecnicamente ao Departamento Nacional de Registro do Comércio (DNRC), todas as unidades federativas do Brasil possuem uma Junta Comercial que, no intuito de se comunicar com o público a respeito dos serviços que oferecem, procuram esclarecer alguns aspectos a respeito da TJ ao divulgar informações em formato eletrônico, tais como: 1) Legislação – decretos, instruções e portarias; 2) Tarifário – tabela de emolumentos dos Tradutores Públicos, ou seja, a remuneração recebida por eles mediante o término de um trabalho tradutório. Os emolumentos não saem dos cofres públicos, mas sim do pagamento efetuado pelo contratante, e têm valores diferentes de um estado brasileiro para outro (Campbell, 1984, p. 117); 3) Relação dos tradutores autorizados, classificados pelo idioma com que trabalham; e 4) explicação sobre o tradutor juramentado, uma tentativa de definição no intuito de informar quem ele é e o tipo de trabalho que desenvolve para o contratante. A partir da explicação, é possível depreender o que essas entidades entendem por TJ. Nem todas as juntas comerciais ofertam essa informação, mas as que foram encontradas constam a seguir.

1 - Junta Comercial do Estado da Bahia (JUCEB)<sup>43</sup>: “Tradutores Juramentados ou Tradutores Públicos e Intérpretes Comerciais são profissionais autônomos habilitados por concurso público promovido pelas Juntas Comerciais de cada Estado, cujas traduções possuem fé pública em todo território nacional.”

---

<sup>43</sup> <http://www.juceb.ba.gov.br/>

2 – Junta Comercial do Estado de Minas Gerais (JUCEMG)<sup>44</sup>: “O ofício de tradutor público e intérprete comercial tem fé pública, reflete em português o conteúdo original do documento traduzido, dando existência legal no país a um documento redigido no estrangeiro.”

3 – Junta Comercial do Estado de Pernambuco (JUCEPE)<sup>45</sup> e Junta Comercial do Estado de São Paulo (JUCESP)<sup>46</sup>: “Tradutor Público e Intérprete Comercial são agentes auxiliares do comércio, que dão fé pública às traduções ou versões por eles feitas de documentações empresariais, certidões, contratos sociais e procurações.”

Mediante o que foi publicado pelas Juntas Comerciais mencionadas anteriormente, entende-se que Traduções Juramentadas são: 1) traduções realizadas por um tradutor juramentado e que possuem fé pública, com validade em todo o território nacional; 2) traduções realizadas por um tradutor juramentado que são o reflexo em português do conteúdo original, dando a ele existência legal; e 3) traduções ou versões de documentos oficiais, que tem fé pública e que são realizadas por um tradutor público e intérprete comercial.

Além das Juntas comerciais, os tradutores podem se ligar a sindicatos e/ou a associações. Essas entidades, de acordo com o perfil com que foram criadas, podem congrega diferentes profissionais em seus quadros desde que, de alguma maneira, estejam ligados à área de tradução. Procuram desempenhar várias atividades que apoiam o tradutor e, dentre elas, o de valorizar o profissional de tradução, promover o intercâmbio de informações, e realizar cursos e eventos. Nesse âmbito, existem as associações destinadas a integrar somente tradutores juramentados como é o caso, por exemplo, da Associação Profissional Dos Tradutores Públicos e Intérpretes do Estado de São Paulo (ATPIESP)<sup>47</sup>, Associação Profissional de Tradutores Públicos e Intérpretes Comerciais, Juramentados, do Estado do Rio de Janeiro (ATP-Rio)<sup>48</sup>, Associação dos Tradutores Públicos de Minas Gerais (ATPMG)<sup>49</sup>, e da Associação Cearense de Tradutores Públicos (ACETESP)<sup>50</sup>. No intuito de se comunicar com seus associados e fomentar o intercâmbio de informações, as instituições supracitadas, com exceção da ACETESP, expõem o que entendem por TJ. A ATPIESP, por exemplo, divulga que a tradução juramentada

[é] a tradução oficial, feita por tradutor público (comumente conhecido como tradutor juramentado), exigida legalmente em todo o território nacional para que documentos

<sup>44</sup> <http://www.jucemg.mg.gov.br/ibr/>

<sup>45</sup> <http://www.jucepe.pe.gov.br/>

<sup>46</sup> <http://www.jucesp.fazenda.sp.gov.br/>

<sup>47</sup> <http://www.atpiesp.org.br/>

<sup>48</sup> <http://www.atprio.com.br/pages/trad.asp>

<sup>49</sup> <http://www.atpmg.com.br/index.php>

<sup>50</sup> <http://www.acetesp.org.br/>

redigidos em língua estrangeira produzam efeito em repartições da União, dos Estados ou dos Municípios, em qualquer instância, juízo ou tribunal ou entidades mantidas, fiscalizadas ou orientadas pelos poderes públicos (artigo 157 do Código de Processo Civil e Decreto Federal nº 13.609 de 21.10.1943). A tradução juramentada tem fé pública em todo o território nacional, e as versões são reconhecidas na maior parte dos países estrangeiros.<sup>51</sup>

As explicações ou definições divulgadas pelas outras associações não vão além do que já foi observado neste capítulo e que, portanto, não atingem suficiente relevo para que sejam replicadas.

De acordo com o sistema legal brasileiro, qualquer documento vindo do exterior não tem validade no território nacional se permanecer na língua estrangeira. Portanto, para que sejam imbuídos do efeito pretendido, e sejam aceitos nas respectivas instâncias, é obrigatório que estejam em língua portuguesa. Para fazer esse ajuste, o usuário deve seguir o trâmite imposto pela CNJ, por meio da Resolução nº 155<sup>52</sup>, de 16 de julho de 2012, ao estabelecer que esses papéis têm que ser validados mediante a intervenção de um tradutor juramentado. Depois disso, não importa em que estado ocorreu, a tradução tem validade legal em todo Brasil. A respeito da legitimação da TJ, Camargo (2007, p. 48) e Orenha e Camargo (2009, p. 59) asseveram que a obrigatoriedade dessa modalidade tradutória no Brasil está prevista no art. 18 do decreto nº 13.609, de 21/10/1943, ao estipular que

[n]enhum livro, documento ou papel de qualquer natureza, que for exarado em idioma estrangeiro, produzirá efeitos em repartições da União, dos Estados ou Municípios, em qualquer instância, Juízo ou Tribunal ou entidades mantidas, fiscalizadas ou orientadas pelos poderes públicos, sem ser acompanhado da respectiva tradução feita na conformidade desse regulamento (p. 48).

Como já mencionado anteriormente, é sabido que cada unidade federativa brasileira possui uma Junta Comercial, um órgão com atuação no estado em que está presente. Ainda de acordo com o mesmo decreto de 1943, é na Junta Comercial que todos os trabalhos produzidos pelos tradutores juramentados ficam depositados, por meio de um “livro”, constituindo um grande arquivo. Segundo Aubert e Tagnin (2004, 2003), esses livros ficam sob a responsabilidade do tradutor durante o período em que estiver atuando na profissão. Depois disso, havendo algum motivo para que não mais exerça a profissão, os livros devem ser encaminhados à Junta Comercial do estado em questão. A esse respeito, o art. 33 estabelece que “[h]averá em cada ofício [de tradutor público] um livro “Registro de Traduções”,

<sup>51</sup> <http://www.atpiesp.org.br/associacao/traducao-juramentada/>

<sup>52</sup> <http://www.cnj.jus.br/atos-administrativos/atos-da-presidencia/resolucoespresidencia/20313-resolucao-n-155-de-16-de-julho-de-2012>

encadernado e numerado em todas as suas folhas que, com isenção de selos e emolumentos, serão rubricadas pela Junta Comercial ou órgãos encarregados do registro do comércio.”

Depois de ser aprovado em um concurso público e ser nomeado, é na Junta Comercial do estado de residência que o nome do “Tradutor Público e Intérprete Comercial” fica registrado. Para conquistar a essa posição, Campbell (1984, p. 113, 114) elenca os requisitos que o candidato a tradutor juramentado no Brasil deve obedecer: 1) ter idade a partir de 21 anos – a regulamentação menciona que o candidato deve ter 21 anos de idade completos, mas não estipula uma idade máxima, fato que pode ocorrer em outros órgãos governamentais e empresas; 2) ser cidadão brasileiro nato ou naturalizado; 3) não ser negociante falido irreabilitado – essa norma foi incluída para garantir que, tendo saúde financeira, o candidato apresente a independência moral presumida do Tradutor Público; 4) não estar respondendo a processo e nem ser condenado por crime que implique demissão ou incapacidade de exercer o ofício – de acordo com Campbell (1984, p. 114), trata-se de uma condição excessivamente rigorosa, já que é necessário ter em mente a presunção de inocência; 5) morar por mais de um ano na praça onde pretende atuar, para garantir que o candidato tenha ligação com o local onde está; 6) estar quite com o serviço militar; 7) ser aprovado em concurso público; e 8) ser nomeado pela Junta Comercial do seu estado domiciliar.

Tendo em vista os pontos levantados até o presente momento, pode-se mencionar que a TJ ocupa um lugar importante na comunicação entre países que desejam iniciar ou reforçar relações de vários âmbitos.

Na seção a seguir, discuto os tipos de documentos a serem investigados nesta pesquisa.

## 1.5. A TIPOLOGIA DOS DOCUMENTOS

Existem muitos documentos que podem ser submetidos à TJ, com variados objetivos que podem motivar os clientes a contratar esse serviço. Sobre essa amplitude e diversificação, Campbell (1984, p. 116) esclarece que

[qualquer texto pode ter sua tradução juramentada, e isto pode significar desde correspondência amorosa – para fins de instrução de processos de desquite ou divórcio – [traduções tão difíceis como as dos mais complexos textos literários, e tendo obrigatoriamente que trazer para o vernáculo clima emocional e insinuações tão próximos do original, quanto possível], como cheques bancários.

À primeira vista, a intenção é a de validar, em território brasileiro, um documento produzido em outro país, na tentativa de conquistar aqui algum direito obtido lá. Em outras

palavras, a tradução juramentada é requerida para que a situação legal originada em outro país possa aqui ser reconhecida ao ser apresentada a alguma instância do governo, ou de qualquer outra autoridade. Nesse caso, o tradutor juramentado, por estar nessa condição por meio de aprovação em concurso, funciona como um “autenticador público” (Asensio, 2003, p. 3), ou seja, torna um documento estrangeiro inteligível na língua de chegada para que obtenha o mesmo efeito que tem na cultura de partida. O mesmo autor (p. 27-29) seleciona exemplos dos tipos de textos que podem se tornar documentos e, por conseguinte, objetos de requerimento, petição, relatório, contrato ou qualquer outro ato legal. Trata-se de uma tentativa de evidenciar o quanto diversa a área pode ser e não de mostrar a total existência de tipologia textual: 1) nota de suicídio; 2) carta particular; 3) artigo científico; 4) histórico escolar; 5) documentação acadêmica; 6) documento administrativo; 7) relatório de acidente de trânsito; 8) estatutos de uma empresa; 9) sentença judicial; 10) situação financeira; 11) fiança; e 12) testamento. Tais documentos podem não ocorrer na mesma proporção e que, portanto, uns podem ser mais frequentemente do que outros, podendo ser a localização um fator preponderante. Nas proximidades de uma universidade internacionalizada, por exemplo, históricos escolares possivelmente estejam no topo da lista de procura por tradução juramentada, visto que, para ingressar nessa instituição, é necessário validar os documentos que comprovam os estudos realizados no exterior.

De acordo com Barros, Babini e Aubert (2010, p. 235), a maioria desses documentos pode ser dividida em seis grupos, conforme a seguir:

- a) **documentos pessoais:** carteira de identidade, certidões de nascimento, casamento, divórcio ou óbito; documentos escolares, carteiras de habilitação de motoristas, passaportes e outros;
- b) **documentos societários:** termos de incorporação, deliberações de conselhos de empresas, atas de reuniões, contratos em geral etc.
- c) **documentos financeiro-comerciais:** balanços de empresas, faturas, notas de débito, letras de câmbio, conhecimentos de embarque, notas promissórias, correspondência comercial etc.
- d) **documentos legais:** cartas rogatórias, atestados de antecedentes, procurações etc.
- e) **documentos de diferentes naturezas:** patentes, transferência de tecnologia, correspondência eletrônica etc.

f) **documentos executivos** (títulos que, na relação credor/devedor, são passíveis de serem levados a processo): notas promissórias, conhecimentos de embarque, cheques, letras de câmbio, faturas, notas de débito, etc.

Ainda a respeito de tradução de documentos oficiais, Asensio (2003, p. 3) esclarece que podem conter os seguintes elementos:

- a) **Elementos de registro:** nascimentos, casamentos, estudos acadêmicos, óbitos, testamentos, atividades ilegais, ou outros atos judiciais ou administrativos (como contratos de venda ou de prescrição médica).
- b) **Elementos documentais,** como cartas, relatórios, formulários de certificação em branco, formulários de certificação completados, certificações validadas, traduções, traduções autorizadas e traduções autorizadas e validadas.
- c) **Elementos de validação,** como a certificação das formalidades correspondentes a diferentes momentos de um processo (as certificações de um escrivão, um notário, um funcionário do tribunal, a apostila de Haia, por exemplo)<sup>53</sup> (Tradução minha).

O autor esclarece também que, mesmo havendo um grande escopo de documentos, com variada recorrência a depender do local, alguns são ainda mais recorrentes que outros e explicita os tipos mais comuns (p. 29): 1. Nascimento, Casamento e Óbito; 2. Divórcio; 3. Acadêmico; 4. Acordos e Garantias; 5. Procuração; 6. Testamentos; 7. Adoções; 8. Companhias; 9. Carta Rogatória; 10. Sentenças e Resoluções; 11. Identificação.

Diante do exposto, pode-se verificar que, mesmo sob fatores diversos que podem influenciar, as certidões de nascimento, de casamento e de óbito, são as mais frequentemente solicitadas na demanda por tradução juramentada. Pela recorrência nas citações, fica evidenciada sua importância e, com isso, o meu interesse em conduzir um trabalho que examine a tradução desses documentos de registro pessoal.

### 1.5.1. O Texto Traduzido

Em narrativas literárias, de acordo com O'Sullivan (2003), o texto é a mensagem transmitida por um *autor real* a um *leitor real*, ou seja, de alguém que de fato escreveu o texto

---

<sup>53</sup> A) Recorded elements: births, marriages, academic studies, deaths, wills, illegal activities, or other legal or administrative acts (such as sales agreements or medical prescriptions). B) Documentary elements, such as letters, reports, blank certification forms, completed certification forms, validated certifications, translations, authorized translations and authorized and validated translations. C) Validating elements, such as the certification of formalities corresponding to different moments of a proceeding (the certifications of a registrar, a notary public, a court clerk, the Hague apostille, for example).

para alguém que, em seguida, o tem nas mãos e realiza a leitura. Visto que essas pessoas não são encontradas no texto, nem o autor real se comunica diretamente com o leitor real, a comunicação acontece por meio de um “par construído na narrativa do texto, denominados *autor implícito e leitor implícito*” (p. 199). Em seguida à sua finalização, o texto segue seu trajeto e, nas palavras de Chatman (1990), o autor real se afasta, retira-se do texto tão logo seja, de alguma maneira, veiculado em direção ao público.

Para circular além dos limites da língua fonte, e informar um número maior de leitores, é necessária a presença de um leitor real que “intercepte a comunicação e a transmita – reprocessada – para o novo leitor que vai receber a mensagem<sup>54</sup>” (Schiavi, 1996, p. 15) (Tradução minha) e, para atender a essa demanda, entra em cena o tradutor. No entender de O’Sullivan (2003, p. 199),

[o] tradutor age, em primeira instância, como o leitor real do texto fonte. Como alguém familiarizado com a língua fonte e com as convenções e normas daquela cultura, o tradutor está em posição de desempenhar o papel de leitor implícito do texto fonte (Tradução minha<sup>55</sup>).

Nesse ambiente, ao tradutor é imputada a responsabilidade de criação de um novo texto em que exista a implantação de mecanismos que permitam a essa tradução funcionar adequadamente na cultura alvo. Ao tomar a frente de um projeto tradutório, o tradutor passa a reorientar o texto, levando em consideração um novo público (Hermans, 2002, p. 12), sempre diferente daquele do texto fonte, um novo *leitor implícito*, o leitor do texto traduzido. O leitor implícito de um texto fonte é gerado por um autor implícito e, por outro lado, o leitor implícito de uma tradução passa a ser gerado por um *tradutor implícito* (O’Sullivan, 2003, p. 200).

Diante do exposto, nos parágrafos seguintes, buscarei rastrear o modo com o qual o tradutor (implícito) realiza a seleção de componentes para a montagem do documento traduzido, elencando sua sequência e os respectivos conteúdos adotados.

### 1.5.2. Divisão Textual da Tradução Juramentada

<sup>54</sup> [...] intercepts the communication and transmits it – re-processed – to the new reader who will receive the message.

<sup>55</sup> The translator acts in the first instance as the real reader of the source text. As someone familiar with the source language as well as the conventions and norms of that culture, s/he is in a position to slip into the role of the implied reader of the source text.

Diferente do que acontece em outras modalidades de tradução, a presença do tradutor oficial nos textos juramentados é mais perceptível o que, de acordo com Asensio (2003, p. 64), é característica da tradução oficial e geralmente ocorre por dois motivos: o primeiro é pela “necessidade de ser preciso, fiel, para evitar fraude, etc.”; o segundo por “refletir seu entendimento imperfeito do original. O que é o caso, por exemplo, quando aparece o comentário ‘ilegível’<sup>56</sup>” (Tradução minha). Segundo o mesmo autor (p. 64), ao incluir registros no texto de chegada, o tradutor faz “o máximo para que seus comentários fiquem o mais perfeitamente distintos do que é incluído no texto fonte<sup>57</sup>” (Tradução minha).

De posse de traduções juramentadas, é possível observar que os documentos são geralmente compostos por três partes:

- 1) elementos pré-textuais, contendo a abertura feita pelo tradutor;
- 2) o texto traduzido, contendo paratextos;
- 3) elementos pós-textuais, contendo o fechamento do documento traduzido.

Início, agora, o detalhamento de cada uma dessas partes.

#### 1.5.2.1. Elementos iniciais ou pré-textuais: o cabeçalho

Mediante verificação do trabalho tradutório em certidões de nascimento, casamento e óbito, é possível encontrar elementos pré-textuais (ou primeira parte da certificação – Asensio, 2003, p. 65). Esses elementos ocorrem da mesma maneira nas duas direções do par linguístico em questão, ou seja, em traduções a partir do inglês para o português e do português para o inglês, e são comuns, com maior ou menor evidência, aos vários tradutores que colaboraram na composição do corpus deste estudo.

A respeito das palavras iniciais de um documento traduzido, Nascimento (2006, p. 54) afirma que

[c]omeçam com um cabeçalho (geralmente) meio longo que inclui diversos elementos que identificam tanto o tradutor quanto sua posição oficial [...] Apesar de variarem muito, os cabeçalhos apresentam informações similares: nome e endereço do tradutor, seu número de registro na Junta Comercial e a certificação de que recebeu um documento a ser traduzido [...] <sup>58</sup> (Tradução minha).

<sup>56</sup> [...] need to be accurate, faithful, and to avoid forgery, etc. [...] reflects only their imperfect understanding of the original. That is the case, for instance, with the appearance of the commentary ‘illegible’.

<sup>57</sup> [...] do their best to make their commentaries perfectly distinguishable from what is included in the original text.

<sup>58</sup> [s]tarting with a (usually) somewhat long letterhead that includes several elements that identify both the translator and his/her official position [...] Although letterheads vary, they all include similar information: the

Nesse sentido, a primeira parte da tradução juramentada, chamada de cabeçalho ou de elementos pré-textuais, pode conter as informações a seguir. Informo que, com base no que ocorre nas traduções juramentadas, os elementos de 1 a 12 foram nomeados por mim, e que a presença/ausência de cada um deles depende do que o tradutor considera suficiente para a finalidade em questão, não havendo orientações ou normativas para isso.

- 1) **Identificação:** no topo da página, o documento se inicia informando o nome completo do tradutor, acompanhado ou não do seu número no Cadastro de Pessoa Física (CPF);
- 2) **Autorização:** neste item, o tradutor explicita a função que exerce. Pode se autodenominar Tradutor Público e Intérprete Comercial, ou Tradutor e Intérprete Juramentado, Intérprete do Comércio e Tradutor Público Juramentado, ou simplesmente Tradutor Juramentado;
- 3) **Qualificação:** Em seguida, pode haver o acréscimo da titulação acadêmica. O tradutor menciona o curso em que se formou, seguido pelo nome da instituição escolar onde conquistou esse nível de escolaridade;
- 4) **Localização:** endereço completo do lugar onde o tradutor atende o público interessado em realizar traduções;
- 5) **Comunicação:** divulgação dos canais de acesso ao tradutor – endereço eletrônico e/ou telefone(s);
- 6) **Numeração:** as traduções obedecem a um sequenciamento organizacional estipulado por meio de número, mês e, muitas vezes, a pasta onde estão arquivadas;
- 7) **Nomeação:** o tradutor pode informar a localidade onde exerce o seu ofício, o idioma para o qual está autorizado a trabalhar, o número da portaria que o nomeou, a data em que esse fato ocorreu, e que foi assinada pelo presidente da Junta Comercial da localidade em questão;
- 8) **Apresentação:** o tradutor atesta o tipo de certidão que foi apresentada para ser traduzida (nascimento, casamento, óbito ou *birth, marriage, death*) e pode informar se está de posse de um documento original ou de uma cópia autenticada. Depois disso, pode mencionar a direção em que a tradução vai ser realizada.

Para finalizar a parte pré-textual, o tradutor informa que a próxima etapa será cumprir sua função, ou seja, traduzir o documento apresentado.

---

translator's name and address, his/her registration number with the Commercial Registry, and his/her certification to have received a document for translation [...] (p. 54)

### 1.5.2.2. Elementos textuais: o documento a ser traduzido

Após detalhar informações a respeito de si mesmo e do documento apresentado, o tradutor agora se debruça sobre o texto fonte. Tendo em vista a intenção de desenvolver um trabalho tradutório que espelhe o texto de partida, ao mesmo tempo em que almeja ir ao encontro das expectativas dos recipientes, o tradutor pode interferir no texto para torná-lo fluente, o que é esperado que faça. Isso pode ocorrer de maneira sutil, como a tentativa de emparelhar elementos interlinguísticos, ou mais explícita, com paratextos, para informar ou explicar elementos contidos no documento apresentado.

Ao falar sobre a presença dos paratextos na tradução, Nascimento (2006, p. 55) esclarece as razões pelas quais esse recurso é largamente utilizado, ao explicar que

[o]s comentários do tradutor aparecem nas traduções oficiais por diversos motivos, como: para destacar a existência de um item semiótico (e.g. [logo de uma escola]; [assinatura ilegível]); para dar informações de onde um material textual específico aparece no texto fonte (e.g. [atrás do documento]; [do outro lado da página 1]), ou para explicar algum item cultural (e.g. CNPJ [Corporate Taxpayer Registration]). São geralmente utilizados para comentários breves, e geralmente aparecem no texto traduzido entre colchetes ([ ]), chaves ({ }), ou parênteses<sup>59</sup> (Tradução minha).

Em visita ao corpus desta pesquisa, pode-se notar que os paratextos apresentam os elementos previstos por Nascimento (2006), e são utilizados com os mesmos propósitos descritos anteriormente, nas duas direções do par linguístico em questão. No caso das certidões, podem-se destacar os seguintes exemplos, utilizados em alguns dos documentos: [texto cortado na lateral esquerda]; [código de barras]; [selo]; [data]; [espaço em branco]; [chancela]; [logo]; [assinalado]; [campo não preenchido]; [iniciais]; [brasão]; [carimbo]; [assinatura]; [verso]; [cancelado]; etc...

### 1.5.2.3. Elementos pós-textuais: o fechamento da tradução

---

<sup>59</sup> Translator's comments appear in Official Translations for several purposes, such as: to highlight the existence of a semiotic item (e.g. [school logo]; [illegible signature]); to provide information as to where some specific textual material appears in the ST (e.g. [back of the document]; [Reverse side of page 01]), or to explain some cultural item (e.g. CNPJ [Corporate Taxpayer Registration]). They are usually used for brief comments, and usually appear in TTs in square brackets ([ ]), braces ({ }), or parentheses.

Logo em seguida à tradução do texto fonte, o tradutor inicia a trajetória que leva ao fechamento do trabalho, ou seja, a segunda parte da certificação (Asensio, 2003, p. 65-66).

- 9) **Notificação:** o tradutor pode inserir uma nota ou comentário a qualquer momento no desenvolver do trabalho. Nascimento (2006, p. 55-56) explica que as notas do tradutor podem ter diversos objetivos, mas geralmente são utilizadas para dar explicações como, por exemplo, a respeito de algum aspecto culturalmente embasado, ou para dar informações como, por exemplo, sobre a validade da tradução.

Observando o corpus deste trabalho, pode-se verificar que, em direção ao inglês, o tradutor escreve apenas:

- **Note:** *Note: The deceased was a naturalized Brazilian. The informant is fully liable for all the information above.*

No sentido inverso, ou seja, em direção ao português, ele escreve:

- **Nota:** *Nota: Fotocópia autenticada da tradução para o inglês do documento original emitido em chinês.;*
- **N. T.:** *N.T. Os itens deixados em branco no documento original não foram traduzidos por serem irrelevantes.;*
- **N. da T.:** *[N. da T.: somente traduzidos os campos pertinentes a este óbito. Documento legalizado pelo Consulado Geral do Brasil em Miami].*

Também é possível localizar uma nota inserida para avisar ao leitor de que não há notas: *Notes: None.*

Como visto, a presença de comentários e notas é bastante intensa na tradução juramentada de documentos. Como anteriormente mencionado, além da maneira explícita com a qual se posiciona no texto, o tradutor também interfere de maneira que pode não ser muito aparente, tornando mais difícil fazer o rastreamento. A esse respeito, Nascimento (2006, p. 199) se remete à sua pesquisa, revelando uma lacuna carente de investigação ao informar que

[a]o contrário do esperado, apesar de a maioria dos comentários aparecer entre colchetes ou chaves, um número muito grande de comentários apareceu no texto sem nenhuma marca gráfica [...]. Esses comentários invadem o texto e podem, ocasionalmente, ser erroneamente interpretados como algo pertencente ao texto fonte e não como intervenção do tradutor<sup>60</sup> (Tradução minha).

---

<sup>60</sup> [c]ontrary to what was expected, although most comments appeared in brackets or parentheses, a very large number of comments appeared in the text without any graphic marks [...]. These comments intrude on the text and may occasionally be mistaken for something belonging to the ST, not a translator's intervention.

10) **Conclusão:** é o momento em que o tradutor diz que a tradução está terminada, informa que o trabalho foi realizado de maneira completa e manifesta a credibilidade que o documento passa a ter mediante sua autenticação. Esse fechamento pode ser feito de várias maneiras que levam ao mesmo entendimento.

**Em direção ao português:**

- E nada mais continha o documento, que dou fé.
- Foi o que traduzi e dou fé.
- Dou fé.
- Este é o teor do documento original que fielmente traduzi.

**Em direção ao inglês pode-se encontrar:**

- In witness whereof I sent my hand.
- In witness hereof I sign this English translation of a \_\_\_\_\_ certificate.
- This is a complete and faithful translation witness my hand and official stamp.
- Such is the tenor of said document, which I hereby certify.
- In my capacity as a sworn translator, I hereby certify that the above text is a true and accurate translation from its original done to the best of my knowledge and ability on this [date].

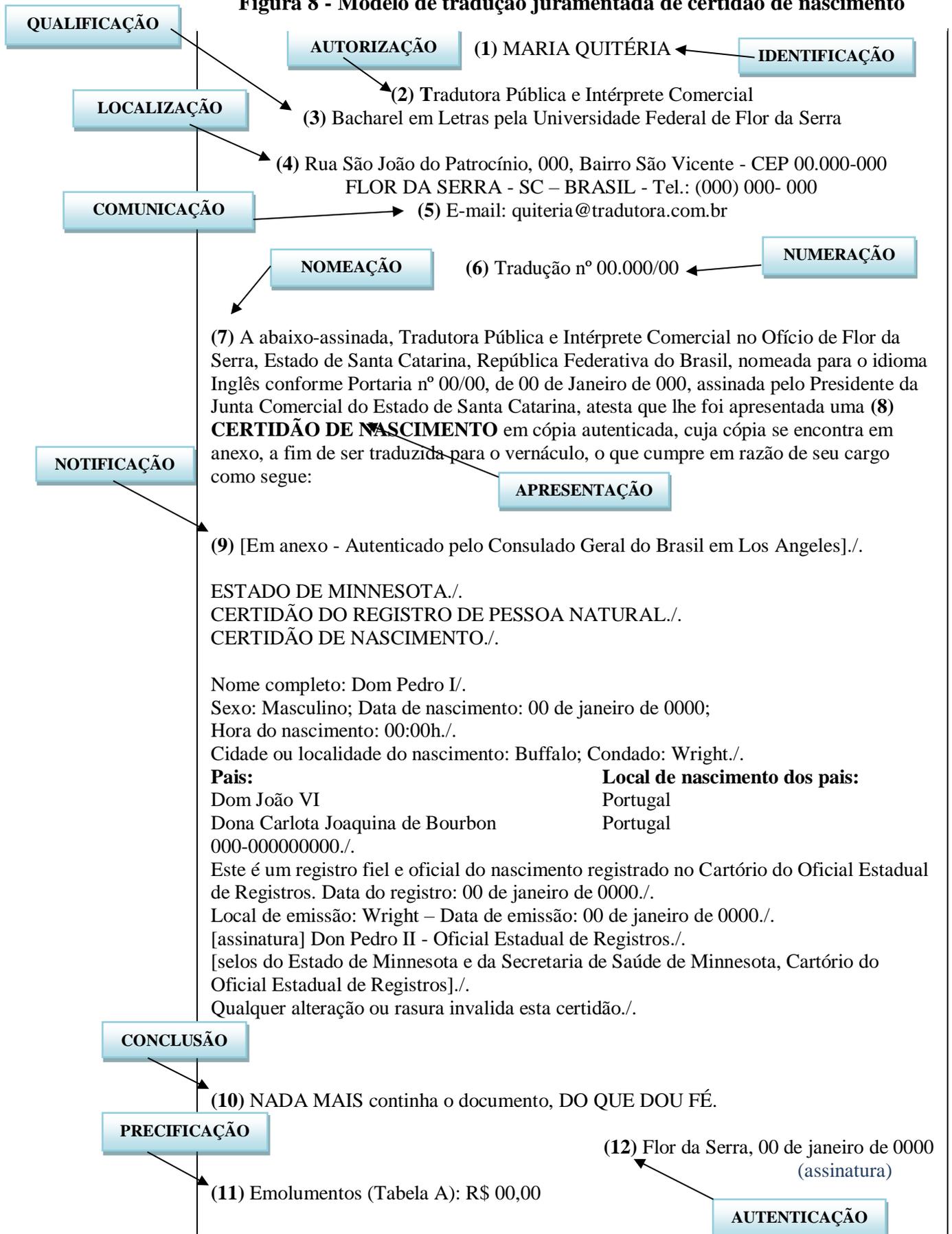
11) **Precificação:** é possível incluir os emolumentos da tradução, ou seja, o preço a ser pago pelo contratante, em moeda local, em virtude da entrega do trabalho. Para que não haja dúvidas quanto ao valor cobrado, o tradutor pode incluir a quantidade de caracteres utilizados ao fazer a tradução e estipular a quantidade de laudas atingidas pelo documento. Em seguida, basta aplicar o valor de acordo com enquadramento na tabela de emolumentos publicada pela Junta Comercial do estado onde o tradutor é atuante.

12) **Autenticação:** nesta parte do documento são incluídos alguns elementos como o local e a data em que a tradução foi realizada, seguidos pela assinatura e carimbo, que atestam a autoria da tradução.

Assim como mencionado para os elementos pré-textuais, o tradutor também pode decidir os elementos que julgar necessários e suficientes como pós-textuais.

A seguir, apresento o layout de uma tradução juramentada, por meio da Figura 8, realizada do Inglês em direção ao Português de uma certidão de nascimento. Neste exemplo, ilustro a disposição dos elementos pré e pós-textuais, constituintes desse documento, conforme os parâmetros discutidos anteriormente nesta parte do capítulo.

**Figura 8 - Modelo de tradução juramentada de certidão de nascimento**



## 2. FUNDAMENTAÇÃO TEÓRICA

Nesta parte, apresentarei a fundamentação que embasa teoricamente o estudo aqui previsto. Como se trata de uma interface, há uma organização de todos os preceitos asseverados pelos estudiosos de várias áreas. Por se tratar de uma pesquisa em âmbito especializado a partir de corpus comparável, inicio a discussão a partir de orientações da Terminologia, mapeando seu trajeto e tecendo considerações que se iniciam no termo e chegam até a unidade fraseológica especializada. Em seguida, abordo a Linguística de Corpus, estudos de corpora, e seus modos de tratamento.

### 2.1. TERMINOLOGIA

Uma pesquisa a respeito do vocabulário contido em traduções juramentadas de certidões, visando à compilação de um glossário, precisa levar em consideração os preceitos básicos da Terminologia e sua evolução, visto que é possível perceber uma crescente necessidade de padronização, em nível internacional, iniciada no século anterior a este (Williams e Chesterman, 2002, p. 20). Nesse contexto, tecerei considerações a respeito do termo e sua diferenciação do léxico em geral; do movimento precursor realizado por Wüster (1998), por meio da Teoria Geral da Terminologia; da revisão dessa teoria preconizada por Cabré (1999a), ao postular a Teoria Comunicativa da Terminologia; discutirei as mútuas contribuições entre Terminologia, Estudos da Tradução e Linguística de Corpus e, por fim, buscarei esclarecer a maneira com a qual um glossário será construído, mediante a utilização dos pressupostos da Terminografia.

Conforme percebido nos elementos mencionados anteriormente, abordarei questões teóricas da Terminologia como, por exemplo, conceitos, sinonímia e polissemia, assim como questões práticas como identificação, extração e compilação de fraseologias.

#### 2.1.1. Teoria Geral da Terminologia (TGT)

O desenvolvimento da linguagem nos seres humanos fomenta a necessidade de nomear atividades e objetos presentes no cotidiano, empreendimento que pode ser considerado precursor da terminologia. Ao entrarem em contato, as civilizações se depararam com um mundo não monolíngue, situação que gera “[...] a necessidade de compreender o universo nomeado por outros homens e começa, então, a compilar palavras, relacionar conteúdos,

identificar equivalentes. Nascem os dicionários bilíngües e obras símiles, nos quais os termos [...] ocupam lugar de destaque” (Barros, 2004, p. 28). Acompanhando a complexidade imposta pela evolução, o interesse a respeito do ‘funcionamento’ da linguagem é o novo componente adicionado à práxis estabelecida, fator que certamente contribuiu para o futuro desenvolvimento de uma disciplina que se dedica ao estudo terminológico e a proposições metodológicas para o manejo desse campo lexical (Idem, p. 30).

Diferentemente do que poderia ser esperado, as iniciativas para a implantação e/ou consolidação dessa área linguística não ocorreram através de postulados de cientistas da linguagem, mas sim por meio da necessidade de especialistas e técnicos de normalizar o repertório lexical de suas respectivas áreas de atuação, tudo isso “a fim de garantir a comunicação profissional e a transferência de conhecimento, num momento em que se buscava uma língua universal que permitisse superar os obstáculos que a linguagem comum causava” (Remenche, 2010, p. 35; Barros, 2004, p. 31).

Dentro dessa perspectiva, alguns trabalhos começam a ser desenvolvidos em vários países, mas é na Áustria que surgem os considerados marco histórico na Terminologia, impulsionando o desenvolvimento da Escola de Viena. Na década de 1930, o engenheiro Eugen Wüster, tendo em mente a padronização de termos de sua área, postula a Teoria Geral da Terminologia (TGT) em sua tese de doutorado, trabalho de base normativa, com enfoque cognitivo, em que aplica os fundamentos mencionados anteriormente (Holanda e Cialdine Arruda, 2013, p. 06; Remenche, 2010, p. 348, 351; Barros, 2004, p. 32; Cabré, 1998, p. 5).

Segundo Wüster (1998, p. 23), acreditava-se não ser possível normalizar a linguagem, visto que os parâmetros para que isso acontecesse passavam pelo âmbito da língua geral, mas, esclarece o autor, o respaldo a essa atividade foi calcado em novos métodos descobertos em um número crescente de trabalhos linguísticos, tornando possível a normalização.

Ao normalizar a terminologia, a intenção pragmática de Wüster era a de buscar fluência e eficácia nas comunicações científicas e técnicas, diretrizes que reduziram a TGT a uma terminologia representativa e prescritiva, visto que os conceitos eram identificados por meio de rótulos, isto é, etiquetas criadas para fugir das ambiguidades do léxico comum. Sob tal enfoque, o termo é orientado de maneira onomasiológica, ou seja, expressa ‘conceitos’ e não ‘significados’ (Cabré, 2011, p. 2; Lara, 2010, p. 11; Remenche, 2010, p. 351; Krieger e Finatto, 2004, p. 33).

É indiscutível que os pressupostos de Wüster são relevantes para o desenvolvimento da área de Terminologia, estudos pelos quais é considerado o “fundador da investigação terminológica moderna” (Galinski e Budin, 1998, p. 15). No entanto, existe uma lacuna:

princípios inconsistentes diante da comunicação real, visto que “não ampliou as possibilidades explicativas, tornando-se apenas orientações metodológicas para a produção terminográfica, alicerçada numa visão idealizada e normalizadora do componente lexical temático das línguas” (Remenche, 2010, p. 351; Almeida, 2003, p. 214).

A partir da década de 1990, informação e comunicação se intensificam e passam a ocorrer com mais fluência e conseqüente rapidez. Diante disso, torna-se necessário haver uma revisão dos fundamentos da TGT, uma atitude visando à implementação de fatores que possam anular suas restrições. O novo cenário de necessidades terminológicas motiva fortes ataques aos pilares da TGT, derrubando o ideal de univocidade, que “significa que, em princípio, um conceito está adscrito a uma só denominação e vice-versa<sup>61</sup>” (Wüster, 1998, p. 137) (Tradução minha). Nesse sentido, Krieger, Maciel e Finatto (2000, p. 2) esclarecem os fatores restritivos a respeito da teoria da Terminologia, evidenciando a carência de reavaliação, ao mencionar que

[...] seus parâmetros teórico-metodológicos fundados sobre uma visão normalizadora e estática dos termos não mais cobrem a produtiva e diversificada realidade dos léxicos temáticos. Nesse sentido, as “insuficiências” maiores da teoria estão relacionadas à sua incapacidade de dar conta da abrangência e diversidade denominativa desses léxicos que, inclusive, não se distinguem do funcionamento normal de toda e qualquer unidade lexical dos sistemas lingüísticos, sofrendo igualmente processos de sinonímia e variação.

O desajuste entre teoria e prática terminológicas começa a dar sinais de inadequação frente ao real funcionamento da linguagem, pois, “mesmo no interior de uma única área de conhecimento, nem sempre há um só conceito ou uma única denominação correspondente. Tal constatação desfaz a ilusão do controle sobre o emprego dos vocabulários especializados” (Krieger, 2001, p. 102), “propondo que o artificialismo do ideal normalizador seja suplantado pelo exame do contexto de produção dos léxicos especializados. A primeira conseqüência é o reconhecimento da variação terminológica nas comunicações especializadas” (Krieger e Finatto, 2004, p. 35).

Agora considerada falha, insuficiente, ou de baixo alcance, a TGT se percebe forçada a ceder lugar a modelos mais amplos e flexíveis que funcionem a partir de um ponto de vista descritivo que possa “explicar melhor os fenômenos que envolvem a comunicação especializada e melhor descrever suas unidades mais representativas, os termos, de forma a abranger toda sua complexidade” (Krieger e Finatto, 2004, p. 34). Essa nova orientação para os estudos e atividades terminológicos é levada em conta, respeitando-se o fato de que

---

<sup>61</sup> Esto significa que, en principio, un concepto está adscrito a una sola denominación, y viceversa.

[a]tualmente, o trabalho de identificação terminológica exige o reconhecimento de grande quantidade de unidades lexicais especializadas que integram a linguagem de campos de conhecimento técnico e/ou científico em contínua e acelerada expansão. Mais ainda, assiste-se também à expansão das linguagens especializadas como decorrência das novas áreas tecnológicas que se agregam ao cenário já estabelecido (Krieger, Maciel e Finatto, 2000, p. 1).

Neste novo cenário de mudanças e necessidades, Maria Teresa Cabré propõe a Teoria Comunicativa da Terminologia (TCT), modelo que pretende abarcar os elementos não contemplados pela TGT, discussão que inicio a seguir.

### **2.1.2. Teoria Comunicativa da Terminologia (TCT)**

Como discutido anteriormente, o início do século 20 não conta com o interesse dos linguistas nem dos cientistas sociais nos estudos e análises terminológicas. Diante disso, movidos pela necessidade, engenheiros e técnicos enveredam por esse caminho na tentativa de compensar essa negligência e, conseqüentemente, atender a crescente demanda por produtos terminológicos de seus respectivos campos de especialização. Na segunda metade do mesmo século, alguns estudiosos da linguagem voltam seus olhares para os fenômenos terminológicos, mas a maioria continua debruçada sobre a tentativa de mapear os princípios que governem todas as línguas humanas. Somente no final do século é que os estudos começam a incluir aspectos comunicativos e, a partir desse interesse, a terminologia passa a configurar nas análises linguísticas (Cabré, 1998, p. 2).

Evidentemente, essa inclusão não ocorre mediante a prosperidade das pesquisas que desenvolviam, mas sim pela pressão das rápidas e contínuas mudanças sociais, que passam a motivar a construção de novos conceitos e de novos paradigmas culturais. Fatores como urbanização, industrialização, acesso à educação e democratização passam a fomentar e a valorizar a tecnologia e a informação e contribuem com a necessidade de avanço da terminologia. De acordo com Cabré (1998, p. 3), “algumas dessas mudanças surtiram especial efeito na linguagem e comunicação interpessoal e criou a necessidade por novos produtos linguísticos, novas profissões relativas à linguagem e novas maneiras de organizar a comunicação<sup>62</sup>” (Tradução minha). Diante do exposto, Cabré (1998, p. 4-7) elenca as mudanças que passam a afetar a terminologia:

---

<sup>62</sup> Some of these changes have had major effects on language and interpersonal communication and have created the need for new linguistic products, new professions related to language, and new ways of organizing communication.

- Acelerado desenvolvimento da ciência e tecnologia, trazendo novos conceitos e novos campos conceituais que requerem novos nomes;
- Criação de novas maneiras de comunicação, que exigem constante atualização de vocabulário, motivando o surgimento das chamadas indústrias da linguagem;
- Produção em massa;
- Comunicação em massa, permitindo a disseminação de terminologia e criando interação entre léxicos geral e especializado;
- Transferência de conhecimento e de produtos, forçando o aparecimento de novos mercados para a ciência, tecnologia, cultura e comércio, o que evidencia a necessidade de lidar com o multilinguagem.
- Informação se torna ferramenta emancipadora e se multiplica em escala geométrica, gerando a criação e a frequente necessidade de ampliação, atualização e sofisticação dos bancos de dados;
- Fundação de entidades governamentais que tratam de políticas para assuntos voltados para a área terminológica, regulamentando a transferência de termos e focalizando a padronização.

Diante dos fatores expostos, não é difícil verificar que necessidades comunicativas passam a governar as tomadas de decisões em âmbito terminológico, necessitando de teorias que abarquem essa vertente, tendo em vista a complexidade do novo parâmetro instalado. Apresentada por Maria Teresa Cabré, a Teoria Comunicativa da Terminologia (TCT) destaca-se como alternativa em substituição à TGT de Wüster, redimensionando os estudos (Barros, 2004, p. 57) e passando a valorizar o aspecto comunicativo das linguagens especializadas e a ampliar o escopo das pesquisas até então realizadas.

Diante do exposto, apresento os encaminhamentos propostos pela TCT, fatores que evidenciam o distanciamento em relação à TGT:

- **TERMO:** Uma unidade lexical pode se tornar um termo ao ser usada em determinados contextos, ou seja, nas comunicações especializadas e, “[c]onsequentemente, o conteúdo de um termo não é fixo, mas relativo, variando conforme o cenário comunicativo em que se inscreve” (Krieger e Finatto, 2004, p. 35). Passando a ser visto como elemento central da terminologia, o termo não se distingue de palavra e não é algo alheio ao sistema linguístico, mas ao contrário, está submetido às mesmas regras que governam o funcionamento do léxico (Remenche, 2010, p. 356; Almeida, 2006, p. 86; Barros, 2004, p. 57; Cabré et al., 1998, p. 4).

- **VARIAÇÃO:** A língua de especialidade é uma realização normal da língua natural, contando com todas as ações que atuam nas línguas vivas (Cabré, 2003, 1998, p. 04; Silva, 2010, p. 45), portanto, o termo é “unidade lexical que sofre todas as implicações sistêmicas e contextuais como qualquer palavra da língua, é o acolhimento do princípio da variação em toda sua dimensionalidade” (Krieger e Finatto, 2004, p. 36). Incluir esse aspecto, segundo Krieger (2001, p. 102), também passa a ser “procedimento teórico-metodológico que qualifica os instrumentos terminográficos elaborados segundo essa ótica, uma vez que eles permitem situar o usuário no panorama da diversidade de realizações do termo.”
- **POLISSEMIA:** Ao aceitar a linguagem como fenômeno complexo, dinâmico e, por conseguinte, imerso em variação, reconhecer a polissemia das comunicações especializadas é mais um passo em direção ao alargamento da abrangência de atuação da TCT (Krieger e Finatto, 2004, p. 36). Além da variação e da polissemia, a sinonímia e a homonímia também passam a ser previstas, aceitas e tratadas (Barros, 2004, p. 59).
- **CONCEITO:** Critérios temáticos e pragmáticos determinam o grau de especialização de um termo, regidos pelas mesmas normas de funcionamento da língua em geral, e os aspectos “sociocultural e linguístico de uma comunidade como fatores que influem nos conceitos, assim como nos canais comunicativos por onde circulam e se difundem<sup>63</sup>” (Cabré et al., 1998, p. 4) (Tradução minha).

Como se pode perceber, a TCT contribui para o redirecionamento das práticas terminológicas ao propor que novos elementos comunicativos, antes não previstos, sejam observados, motivando encaminhamentos para o avanço da área, proporcionando respostas adequadas às necessidades comunicativas dos usuários de terminologia e iluminando o caminho para o lançamento de novos produtos terminográficos, inclusive por meio de modernas ferramentas tecnológicas (Remenche, 2010, p. 360; Cabré et al., 1998, p. 5).

### 2.1.3. Unidade Terminológica (UT)

Entendido como o objeto sobre o qual se baseia uma área de investigação, o termo, de acordo com Krieger e Finatto (2004, p. 76-80) e Krieger (2001, p. 68), continua impondo desafios para o seu reconhecimento. Esse fato pode ser verificado pela impossibilidade de

---

<sup>63</sup> El trasfondo sociocultural y lingüístico de una comunidad como factor que influye en los conceptos, así como los canales comunicativos por donde circulan y se difunden.

delimitação de um campo essencialmente terminológico desprovido de aproximações e interfaces com o léxico comum da língua. Alguns fatores contribuem para a existência dessa opacidade nas fronteiras, e todos eles podem ser atribuídos à ampliação do escopo dos estudos terminológicos, que passou a abarcar novos integrantes como: 1. Ampla abrangência do que pode representar: conteúdos, objetos, abstrações e processos; 2. Mobilidade: o termo, auxiliado pela rápida difusão comunicativa, pode sofrer *vulgarização*, ou seja, caminhar em direção ao léxico comum. Nesse viés, o contrário também pode ocorrer, isto é, uma palavra pode caminhar em direção ao contexto especializado e assumir a função de termo, atividade chamada *terminologização*; 3. Polivalência funcional: o termo se comporta de maneira natural e eficaz em mais de uma terminologia.

Mesmo diante da complexidade para seu reconhecimento, algumas afirmações, de acordo com Krieger e Finatto (2004, p. 81) e Krieger (2001, p. 63), podem ser feitas a respeito dessa unidade complexa e multifacetada (Cabré, 1999):

1. Termos são predominantemente substantivos;
2. Podem ser unidades simples ou complexas, conhecidas como sintagmas terminológicos;
3. Orientam a produção de produtos terminográficos;
4. Siglas, acrônimos, abreviaturas e fórmulas também podem ser termos;
5. Necessitam de um contexto pragmático de comunicação;
6. Integram as línguas naturais; e
7. Estão presentes nas fraseologias, inclusive nas especializadas, característica que vai ao encontro do interesse deste trabalho.

Como visto, pode-se dizer que há ausência de consenso sobre aspectos relativos ao termo. Diante disso, Santiago (2013, p. 59) afirma que é possível inferir que situação idêntica ocorre com as unidades fraseológicas especializadas, já que é aceitável dizer que apresentam o termo como um de seus elementos constituintes.

#### **2.1.4. Unidade Fraseológica (UF)**

Presentes em todas as línguas e consideradas desafio ao tradutor no desempenho de seu trabalho (Orenha e Camargo, 2009, p. 61), as unidades fraseológicas, além de múltiplas, diferentes e complexas, podem não ser facilmente encontradas em produtos terminográficos (Paxeco, 2010, p. 1), visto que os estudos terminológicos focalizam o termo, considerado elemento essencial da Terminologia e, muitas vezes, priorizado diante da opção por

combinatórias de palavras (Orenha-Ottaiano; Hasmann; Camargo, 2013, p. 173). Tal preferência leva Hanks (2010, p. 1299) a afirmar que a produção de dicionários “[...] tem focalizado principalmente a tendência terminológica, isto é, o significado das palavras em relação ao mundo, e tem seriamente negligenciado a fraseologia<sup>64</sup>” (Tradução minha) e não levando em consideração aspectos de uso ao mesmo tempo em que trata de aspectos do significado.

Os perceptíveis avanços nos estudos do léxico e a necessidade de acomodar, de maneira “natural”, determinadas sequências de palavras na língua de chegada, fazem com que a lacuna a ser preenchida com material fraseológico fique ainda mais notória, assinala Tagnin (2002b, 735). Ainda segundo Tagnin (2002a, p. 194), existem poucos dicionários de colocações em inglês, alguns em português que apenas ofertam uma lista restrita de ocorrências, e o terreno fica ainda mais árido quando se trata de dicionários bilíngues, ou quando as línguas envolvidas não forem as mencionadas neste parágrafo (p. 200).

Isoladamente, os componentes da fraseologia não conduzem ao correto entendimento da sua totalidade e, portanto, insistir em desconstruí-la, ou separá-la em partes, certamente trará resultados infrutíferos. Sobre isso, Orenha e Camargo (2009, p. 62) afirmam que “[...] uma unidade fraseológica (UF) é uma expressão formada por muitas palavras, caracterizada por estar já estabelecida, institucionalizada, por ser estável, por apresentar certa particularidade sintática ou semântica e pela possibilidade de variação de seus elementos integrantes.”

No corpus de certidões brasileiras pode-se, a título de exemplificação do que foi mencionado no parágrafo anterior, citar a expressão “O referido é verdade e dou fé”, e nas certidões americanas sempre há a expressão “*This is to certify that this is true and correct representation [...]*.” Não é proibido que se faça a troca de quaisquer dos componentes dos exemplos anteriormente mencionados, visto que várias outras possibilidades também gerariam expressões gramaticalmente possíveis. Nesse âmbito, “O mencionado é verídico, aplico fé” e “*This is to guarantee that this is correct and authentic representation [...]*” são expressões que atingem o mesmo efeito semântico em comparação com as que foram supracitadas. Entretanto, a escolha, o posicionamento e a quantidade de itens lexicais são características sedimentadas pelo uso e, diante de algum tipo de alteração, a sequência resultante estaria suscetível a perder a naturalidade, causando, dessa forma, estranhamento ao ser lida pelo usuário (Monteiro, 2012, p. 5). Nesse viés, Tagnin (2002a, p. 194) previne o tradutor para que não recorra a escolhas

---

<sup>64</sup> [...] have focused mainly on the terminological tendency - i.e. the meaning of words in relation to the world - and have seriously neglected phraseology.

alheias às convenções de determinado gênero textual, no momento em que toma decisões tradutórias, ao argumentar que

[p]or estranho que pareça, mesmo como falante nativo da língua alvo, o tradutor pode ter problemas no nível da produção para conseguir soluções naturais, caso se atenha tanto ao texto de partida a ponto de não perceber que, entre formas igualmente gramaticais, uma delas é de uso mais corrente. Em outras palavras, pode não se dar conta de que, dentro de uma gama de formas gramaticalmente *possíveis*, há certas formas que têm uma *probabilidade* maior de ocorrerem. Caso o tradutor selecione uma dessas formas *possíveis*, em detrimento da mais *provável*, produzirá uma tradução não natural, não fluente. Esse problema certamente se agravará quando o tradutor não estiver traduzindo para sua língua materna.

Após o entendimento de que não são meras expressões idiomáticas cristalizadas, o interesse nos estudos sobre a fraseologia começa a decolar ao longo do século 20 (Gries, 2008, p. 3), “[d]esde então foram desenvolvidas importantes reflexões sobre o assunto, intimamente relacionado à problemática da formação de palavras, de sintagmas e de locuções, bem como de estruturas chamadas de colocações” (Krieger e Finatto, 2004, p. 84). Ainda que figurando em todos os campos linguísticos, pode-se afirmar que os estudos sobre a fraseologia ainda necessitam de contribuição, realidade que passa a assumir novos contornos diante da constatação de Granger e Meunier (2008, p. XIX) ao esclarecer que

[p]or muito tempo considerada assunto periférico, a fraseologia está agora entrando em posição central dentro de um vasto espectro de campos de estudos, de processamento de linguagem natural a ensino de línguas estrangeiras, e agora, 25 anos depois da publicação do artigo seminal de Pawley e Syder, está gradualmente adquirindo o lugar que merece na teoria linguística<sup>65</sup> (Tradução minha).

Estudos de corpora são capazes de evidenciar que as unidades lexicais não atuam de maneira estanque em âmbito textual, mas estabelecem relações umas com as outras por meio de padronização fraseológica em que o significado respeita o contexto em que estão inseridos (Orenha-Ottaiano; Hasmann; Camargo, 2013, p. 176). Diante do exposto, é possível afirmar que existe afinidade entre palavras para coocorrerem e formarem grupos preferenciais (Desmet, 2002, p. 37) como, por exemplo, compostos, colocações, expressões idiomáticas, locuções, expressões fixas, coocorrentes, etc.

De acordo com Coimbra e Bevilacqua (2009, p. 237) “[m]uitos teóricos têm a fraseologia como o foco de seus estudos, não havendo consenso sobre sua denominação,

---

<sup>65</sup> Long regarded as a peripheral issue, phraseology is now taking centre stage in a wide range of fields, from natural language processing to foreign language teaching and now, 25 years after the publication of Pawley’s e Syder’s seminal article, it is gradually acquiring the place it deserves in linguistic theory.

delimitação e critérios para sua identificação, uma vez que é um objeto altamente complexo”, situação que, a partir do ponto de vista de Moreno (2014, p. 43), deve-se à “[...] enorme quantidade de pesquisa dedicada à fraseologia nas últimas décadas, [o que] tem impulsionado o desenvolvimento desta disciplina, mas não resolveu a questão da diversidade na terminologia usada para descrever o objeto de estudo<sup>66</sup>” (Tradução minha).

A falta de critérios para simetria propiciou que a “unidade fraseológica” fosse rebatizada a partir do ponto de vista do pesquisador, na tentativa de que se alinhasse aos contornos do estudo que estivesse em tela. Sem os devidos parâmetros teóricos a serem seguidos, esse expediente autorizou a multiplicidade denominativa dos itens lexicais produzidos em sequência, fato ilustrado nos exemplos a seguir:

- a) **Desmet** (2002, p. 34) se refere à unidade fraseológica como “combinação mais ou menos fixa” distribuindo-a entre: fraseologia, expressões idiomáticas, expressões usuais, idiotismos, colocações, máximas, clichés, adágios, ditados, provérbios, expressões feitas, fórmulas estereotipadas, unidades poliléxicas.
- b) **Bevilacqua** (2005, p. 245) menciona as seguintes denominações: colocação, colocação das linguagens especializadas, coocorrente, entidade fraseológica, frasea, fraseologismo, unidade fraseológica, unidade fraseológica especializada.
- c) **Gries** (2008, p. 15) denomina a unidade fraseológica como: agrupamento de palavras (word cluster) / n-gramas, elementos combinados, colocações, coocorrentes, fraseologismo. Também explica que não há flexibilidade entre os elementos do agrupamento de palavras/n-gramas e que pode haver alguma nas colocações. Ainda de acordo com o autor (2008, p. 16), a noção de coligação foi alterada, passando de coocorrente de padrões gramaticais para um tipo de fraseologismo em que uma ou mais palavras habitualmente co-ocorrem com um padrão gramatical.
- d) **Sabino** (2011, p. 386) menciona que os autores se referem ao objeto de estudo da seguinte maneira: lexias complexas; sequências, combinações ou combinatórias cristalizadas; sintagmas cristalizados; sequências, combinações ou combinatórias fixas; sintagmas fixos; expressões fixas; combinações estáveis; expressões pluriverbais; unidades pluriverbais lexicalizadas; unidades lexicais pluriverbais; unidades polilexemáticas; frases feitas; fraseolexemas; unidades fraseológicas; locuções

---

<sup>66</sup> La enorme cantidad de investigaciones dedicadas a la fraseología en las últimas décadas ha impulsado el desarrollo de esta disciplina, pero no ha resuelto el tema de la diversidad terminológica empleada para la designación del objeto de estudio.

fraseológicas; fraseologismos; unidades terminológicas – sendo estas últimas, fraseologias específicas da linguagem especializada.

- e) **Orenha-Ottaiano; Hasmann; Camargo** (2013, p. 174) preferem utilizar a expressão “colocações especializadas” ao estudarem o texto técnico e científico.
- f) **Moreno** (2014, p. 71) revela que a diversidade denominativa traz confusão e não contribui para a homogeneização de critérios. Afirma que este fenômeno pode ser chamado de colocação, frasema, fraseologismo, solidariedade lexical, entidade fraseológica, unidade fraseológica, etc., de acordo com a perspectiva de estudo adotada<sup>67</sup> (Tradução minha).

Importante mencionar que empreender esforços em visualizar nitidamente as fronteiras entre as unidades fraseológicas pode desapontar o interessado nesse procedimento. Mesmo imersas em opacidade, as sequências lexicais podem ser organizadas em um eixo gradativo de fixação, momento em que, de acordo com Aguado de Cea (2007, p. 56), as unidades fraseológicas

“[...] entram em um *continuum* combinatório, estando as combinações livres em um extremo e, no outro, as combinações fixas e restritas, isto é, as expressões idiomáticas. Este eixo representado por fixação se combina com eixo de frequência, que é outra característica que normalmente se aplica às unidades fraseológicas [...]”<sup>68</sup> (Tradução minha).

Monteiro (2012, p. 06) relata que outros autores já abordaram a questão do *continuum* dentro do qual se podem estipular os graus de convencionalidade atribuídos às unidades fraseológicas, como a seguir:

- **Benson, Benson, Ilson** (1997): em um extremo da gradação estão localizadas as *combinações livres*, ou seja, sequências em que uma palavra pode se combinar com diversas outras, restringidas apenas pela gramaticalidade; no extremo oposto estão as *expressões idiomáticas*, sequências totalmente fixas em que o entendimento não pode ser atingido ao se analisar o significado de seus elementos individualmente; e

<sup>67</sup> [...] la profusión terminológica para designar las unidades especializadas refleja la diversidad de puntos de vista de los estudios fraseológicos y es un factor que aporta confusión y contribuye poco a la homogenización de criterios. A pesar del enorme interés suscitado por la fraseología especializada no se ha alcanzado un consenso en lo relativo a la denominación; así se conoce este fenómeno como *colocación, frasema, fraseologismo, solidaridad léxica, entidad fraseológica, unidad fraseológica*, etc., según la perspectiva de estudio que se adopte.

<sup>68</sup> [...] entran dentro de un *continuum* combinatorio, estando las combinaciones libres en un extremo y, en el otro, las combinaciones fijas y restringidas, es decir, los idioms o modismos. Este eje, representado por el de fijación, se combina con el eje de frecuencia, que es otros de los rasgos que suele caracterizar las unidades fraseológicas [...]

*colocações*, que se encontram numa posição medial, ou seja, não são completamente livres e nem completamente fixas;

- **Lewis** (2000, p. 08): a gradação se inicia com *colocações livres*, passa pelas *colocações relativamente fixas* e chega às *colocações totalmente fixas*, como as expressões idiomáticas.
- **Sinclair** (2003, p. 28): o grau de fixação está atrelado ao sentido emanado pela combinação: no extremo *livre* estão as de sentido literal, que permitem substituição de elementos, enquanto no extremo oposto estão as combinações *fixas*, de sentido metafórico, resistentes a substituições;

Segundo Tagnin (1999, p. 41, 42), para ser considerada uma colocação, ou seja, uma combinação natural de palavras associadas umas com as outras (O'Dell e McCarthy, 2011, p. 4), é necessário que apresente: 1. Recorrência; 2. Significado composicional; 3. Forte ligação entre os elementos; 4. Probabilidade de ocorrência dentro de um contexto específico; 5. Ausência de razão semântica que explique tal ocorrência. De acordo com a mesma autora (1999, p. 41-2; 2002a, p. 195; 2002b, p. 738), as colocações podem ser classificadas a partir da classe gramatical a que pertencem. Sob essa ótica, cada classe recebe denominação de acordo com seu desempenho semântico. Ao apresentar mais carga semântica, um item lexical é denominado *base*, e o outro, por agregar características à base, é denominado *colocado*. Essa classificação pode ser espelhada de acordo com a formulação a seguir:

- **VERBAL**

Verbo <sub>colocado</sub> + Substantivo <sub>base</sub> = set up a business

Substantivo <sub>base</sub> + Verbo <sub>colocado</sub> = the economy shrinks

Verbo <sub>colocado</sub> + Preposição + Substantivo <sub>base</sub> = drop out of a business

Verbo <sub>colocado</sub> + Adjetivo <sub>base</sub> = grow serious

- **NOMINAL**

Substantivo <sub>base</sub> + Substantivo <sub>colocado</sub> = stock brokers

Substantivo <sub>base</sub> + Preposição + Substantivo <sub>colocado</sub> = a decline in value

- **ADVERBIAL**

Advérbio <sub>colocado</sub> + Adjetivo <sub>base</sub> = outrageously high

Verbo <sub>base</sub> + Advérbio <sub>colocado</sub> = to invest heavily

Advérbio <sub>colocado</sub> + Verbo <sub>base</sub> = sharply attack

- **ADJETIVA**

Adjetivo <sub>colocado</sub> + Substantivo <sub>base</sub> = a declining economy

A respeito do descrito no parágrafo anterior, Tagnin (2002a, p. 195) afirma que as colocações nominais e adjetivas perfazem o maior número de ocorrências do “inventário fraseológico”, assegurando que “[h]á milhares delas e a cada dia surgem outras, pois são empregadas para nomear novas tecnologias, processos, teorias, etc. [...] e novos objetos e produtos [...]”. Diante dessa constatação, é correto inferir que os outros tipos de colocação, as verbais, listados acima não ocorrem com tanta frequência.

Monteiro (2012, p. 1) explica que existem duas maneiras com as quais a colocação pode ser abordada. Segundo a autora, chama-se *estatística* a vertente postulada por Sinclair (2003) que examina o número de ocorrências de itens lexicais em certo contexto. A segunda vertente, postulada por Cowie (1998) e Howarth (1998) é denominada *fraseológica*, visto que “verifica o tipo de combinação entre palavras, sua tipificação e a convencionalidade.” Diante do exposto neste parágrafo, é preciso informar que, frente aos parâmetros de colocabilidade mencionados por Monteiro (2012, p. 1), esta pesquisa se dedica a elencar as colocações encontradas no contexto das certidões de nascimento, casamento e óbito, e seus equivalentes tradutórios, no intuito de facilitar o trabalho do tradutor ao lidar com esse gênero textual. Tendo isso em mente, pode-se afirmar que este estudo tem como uma de suas bases os postulados de Sinclair (2003), relacionado à Linguística de Corpus, portanto, evidencia a vertente *estatística*.

Com os argumentos discutidos até aqui, foi possível verificar que a fraseologia está diretamente ligada a uma área ainda em construção, dentro da qual muitos aspectos ainda precisam ser pesquisados e desvendados. De acordo com Sabino (2011, p. 235), basta uma consulta à literatura específica para

[...] perceber que não há limites rígidos capazes de estabelecerem e delimitarem, com precisão, a diversidade de fraseologismos existentes. Não há, ainda, consenso sobre quais unidades são objetos de estudos da fraseologia, nem tampouco sobre as denominações que elas devam receber, quer dizer, não há uma concordância de ideias entre os teóricos no que diz respeito aos termos genéricos que nomeiam os diversos tipos de combinação de palavras.

Como se pode perceber, estamos diante de uma disciplina que engloba número elevado de elementos, a respeito dos quais não há o que se pode chamar de alinhamento, uniformidade, unanimidade ou consenso, problema que, de acordo com Sabino (2011, p. 385), está centrado “na profusão terminológica gerada pelos próprios estudiosos da área, na dificuldade de estabelecimento de critérios precisos para o seu reconhecimento, bem como nas diferentes classificações empregadas nas pesquisas científicas desenvolvidas até o presente momento.” Mesmo diante da mencionada profusão, para Weinert (1995, p. 182), independentemente de

qual seja o nome utilizado para se referir à ocorrência de aproximação entre palavras, “[...] parece que os pesquisadores têm o mesmo fenômeno em mente<sup>69</sup>” (Tradução minha).

#### 2.1.4.1. Unidade Fraseológica Especializada (UFE)

Com a crescente demanda por textos traduzidos, motivados pelas aproximações comerciais, educacionais, ou técnico-científicas, verificou-se constante imposição de fluidez de informações nos textos de chegada, intimando o tradutor a agir com desenvoltura e eficiência diretamente proporcionais à rapidez com que essas informações transitam pelos canais midiáticos. Como se pode perceber, novamente as questões práticas impulsionaram a interesse nos estudos em determinada área: inicialmente esse movimento foi verificado com a terminologia e, neste momento, com a fraseologia, ou seja, com “[...] as necessidades de tradutores e redatores especializados para resolver os problemas semântico-lexicais apresentados pelos termos ao aparecerem em combinação com outras palavras<sup>70</sup>” (Aguado de Cea, 2007, p. 54) (Tradução minha).

Com o intuito de auxiliar o tradutor, nascem o estudo e o interesse em fraseologia especializada, buscando minimizar, ou solucionar, dificuldades semânticas na tradução e utilização adequada de combinações de palavras (Moreno, 2014, p. 64). Nesse setor, segundo Bevilacqua (2002, p. 139), “[...] não se encontram os equivalentes adequados dos co-ocorrentes de determinado termo de uma língua fonte para uma língua alvo, ou seja, não se conhece as UFEs de um âmbito específico.” O consultante, ao procurar por um termo, possivelmente não terá dificuldades para encontrá-lo e nem para chegar ao significado ou ao equivalente. Os termos não ocorrem isoladamente, mas em combinações, e o problema aparecerá quando houver a tentativa de localizar o exato coocorrente a ser utilizado com o referido termo, já que podem existir vários deles, e encaixá-lo na correta situação comunicativa para imprimir naturalidade ao texto científico, pode ser algo bastante complexo (Bevilacqua, 2002, p. 137; 2001b, p 106).

Muitas propostas de definições para UFE, de acordo com Moreno (2014, p. 66), são baseadas em propostas da fraseologia geral e compartilham propriedades semelhantes. Segundo o autor, a “[...] fraseologia especializada apresenta as mesmas variações terminológicas,

<sup>69</sup> [...] it seems that researchers have very much the same phenomenon in mind.

<sup>70</sup> [...] las necesidades de traductores y redactores de textos especializados por resolver los problemas léxico-semánticos que plantean los términos, al aparecer combinados con otras palabras.

taxonômicas e delimitativas que a fraseologia geral com uma diferença importante é que os provérbios são raros nas taxonomias de fraseologia especializada<sup>71</sup>” (Tradução minha). Diante do exposto, é possível afirmar que as instabilidades que acometem as fraseologias em âmbito geral são as mesmas que incidem sobre as encontradas em âmbito especializado. Mesmo com ambiente favorável aos estudos de fraseologia, a falta de simetria mencionada anteriormente acerca de aspectos das UFs pode, neste momento, ser replicada, dificultando a conceituação de UFEs e garantindo a manutenção da opacidade entre as fronteiras. A esse respeito, Aguado de Cea (2007, p. 55) esclarece que

[...] este panorama tão frutífero e esperançoso é ofuscado pela falta de consenso na hora de delimitar o conceito de fraseologia, caracterizar, extrair e estruturar as UFEs e sua relação com outras unidades sintagmáticas, de estabelecer uma fronteira nítida para diferenciar as possíveis combinações restritas das combinações livres. Este desacordo se manifesta também em como registrar nos dicionários, bases de dados terminológicos, etc. A partir disso é que há diferentes critérios para o estudo da UFEs<sup>72</sup> (Tradução minha).

Visando a apresentar uma definição satisfatória para a UFE, a seguir, cito alguns autores que manifestaram seu entendimento sobre a questão.

- **Bevilacqua** (2001, p. 106) afirma que são “[...] unidades sintagmáticas que incluem um termo entre seus elementos, que possuem um determinado grau de fixação e frequência relevante em um conjunto de textos ou em um âmbito especializado.”
- **Desmet** (2002, p. 41) menciona que UFEs são estruturas que parecem apresentar:
  1. No mínimo um termo, simples ou complexo;
  2. Certo grau de fixação determinado pelas relações sintático-semânticas;
  3. Frequência significativa nos textos em que surgem
- **Krieger e Finatto** (2004, p. 84) definem essas unidades fraseológicas como “fórmulas ou frases feitas, próprias de determinados âmbitos especializados: *Revoguem-se as disposições em contrário.*”

---

<sup>71</sup> La fraseología especializada presenta las mismas variaciones terminológicas, taxonômicas y delimitativas que la fraseología general con una diferencia importante y es que los refranes son poco frecuentes en las taxonomías de la fraseología especializada.

<sup>72</sup> [...] este panorama tan frutífero y esperanzador se ve ensombrecido por la falta de consenso a la hora de delimitar el concepto de fraseología, de caracterizar, extraer y estructurar las UFE y su relación con otras unidades sintagmáticas, de establecer una frontera nítida para diferenciar las posibles combinaciones restringidas de las combinaciones libres. Este desacuerdo se manifiesta asimismo en cómo se han de registrar en los diccionarios, bases de datos terminológicos, etc. De ahí que haya diferentes criterios en el estudio de las UFE.

- **Aguado de Cea** (2007, p. 57) explica que UFEs “[...] são combinações léxicas que incluem, além de um termo, um verbo ou uma categoria derivada de verbo (nome ou adjetivo)<sup>73</sup>” (Tradução minha). A partir disso, a realização linguística da relação estabelecida pelo verbo pode ser representada como V+N, N+N, N+Adj., V+Adv. Na verdade, trata-se de ‘colocações’ ou ‘frasemas’.
- **Zilio** (2012, p. 155) define UFE como “[...] um sintagma, contíguo ou não, não necessariamente limitado pela oração ou mesmo pela pontuação, que apresenta um determinado nível de fixação e que ocorre, necessariamente, em textos especializados.”

Uma rápida consulta à literatura terminológica pode evidenciar que não há unanimidade no que se refere à delimitação de critérios para o reconhecimento de UFE. Diante dessa constatação, é necessário estabelecer um parâmetro sobre como os autores se articulam na tentativa de chegar a um formato que possa abarcar, com segurança, o que se pode chamar de UFE. Tal proposição será discutida a seguir:

- a) **Bevilacqua** (2001, p. 98) explica que “apenas os *critérios linguísticos* não são suficientes para identificar unidades fraseológicas, é necessário levar em conta os aspectos pragmáticos, visto que as unidades fraseológicas se definem no nível de discurso.” É preciso considerar a **estereotipia** e a **frequência**.
- b) **Desmet** (2002, p. 44): toda unidade fraseológica especializada comporta um termo simples ou complexo.
- c) **Krieger e Finatto** (2004, p. 88) postulam que a estereotipia e a variabilidade são características que permitem o reconhecimento das UFEs. A **estereotipia** se relaciona com a área, tipo de texto e repetição, enquanto a **variabilidade** diz respeito à possibilidade de variação dos elementos que se agregam à base.
- d) **Bevilacqua** (2005, p. 237, 242) aborda a questão mediante a divisão do assunto em duas perspectivas: a **primeira** de caráter lexicológico e lexicográfico, que divide as unidades fraseológicas em “colocado” e “elemento coocorrente”. O colocado (ou base) é semanticamente autônomo, ao passo que o coocorrente é restringido semanticamente pela base. O seu significado só pode ser deduzido pelo conjunto desses elementos. Estruturas que facilitam o reconhecimento: substantivo + substantivo; substantivo + adjetivo; verbo + substantivo; adjetivo + advérbio; verbo + advérbio. A **segunda** de

---

<sup>73</sup> [...] combinaciones léxicas que incluyen, además de un término, un verbo o una categoría de verbal (nombre o adjetivo).

caráter terminológico. “[...] as UFE são unidades sintagmáticas, formadas por dois ou mais elementos, sendo que um elemento que as compõem é um “termo”, a partir do qual se identifica ou seleciona a unidade.” Entre os padrões morfossintáticos, cito: termo + adjetivo; verbo + termo; termo + verbo; substantivo + preposição + artigo + termo.

e) **Aguado de Cea** (2007, p. 57): segundo a autora não é simples fazer a distinção entre termo composto e unidade fraseológica porque o trânsito de unidades da língua geral em ambiente especializado impede o imediato reconhecimento. Mas, entre os critérios para a delimitação de uma UFE, cita:

- São estruturas sintagmáticas com um termo ou uma unidade terminológica.
- Incluem um verbo ou um elemento derivado de verbo (deverbal): *formatear un disco, el formateado del disco*;
- Têm alto grau de fixação, mas, podem aceitar que outros elementos sejam inseridos, como um quantificador ou um pronome: *gastar tóner, gastar mucho tóner, consumir energía, consumir toda la energía*;
- Apresentam regularidade composicional. Por exemplo, com o binômio adjetivo *computer-aided* encontram-se muitos casos: *computer-aided design/ engineering/ instruction / learning/ management/ manufacturing/ testing/ training/ translation*;
- Admitem a substituição de seus elementos. Assim, observa-se que o sentido especializado de um verbo está delimitado por um conjunto limitado de nomes, como em: *table a motion/ bill/ amendment* ou vice-versa, *adopt/ asume/ take on a role*, ou aparece uma categoria menos restrita, quando há um conjunto de verbos que se pode alternar com um conjunto de nomes: *carry out/ conduct an experiment/ a test/ a survey*;
- Mantém um significado específico dentro de um domínio, adquirindo uma estabilidade semântica: *volcar datos, navegar por la rede, localizar la impresora, capturar una pantalla, etc.*;
- São de uso frequente dentro de um domínio ou subdomínio de especialidade.

f) **Coimbra e Bevilacqua** (2008, p. 3): proposta de critérios para o reconhecimento de fraseologia especializada:

- Caráter sintagmático: deve ser formada por mais de uma unidade léxica;
- Estabilidade semântica: possibilidade de comutação entre os elementos da fraseologia, levando-se em conta o grau de fixação, ou seja, a amplitude de possibilidades combinatórias;

- Estabilidade sintática: identificação a partir de padrões morfossintáticos pré-estabelecidos: V + N, V + Art + N, N + Prep + N, entre outros.
- Grau de fixação: possibilidade de substituição ou inserção de elementos;
- Presença de termo: considerado núcleo para que seja considerada fraseologia especializada;
- Uso em âmbito específico: pertencer a um âmbito, o que determina sua prototipicidade;
- Frequência relevante: deve ter uso recorrente ou elevado em uma área temática. Critério que depende do tamanho do corpus e o objeto de estudo em questão.

Como visto, a partir do ponto de vista terminológico, alguns dos estudiosos listados afirmam que as fraseologias, para que sejam consideradas especializadas, precisam contar com pelo menos um termo, enquanto outros consideram que, por estarem presentes em textos especializados, as fraseologias já podem ser consideradas como tal. As fraseologias consideradas pertencentes à linguagem comum transitam livremente pelos ambientes especializados, e podem acarretar os mesmos tipos de problemas e dificuldades daquelas existentes a partir do ponto de vista terminológico. Com isso, neste estudo, não pretendo contribuir com um debate que elimine a falta de unanimidade entre estudiosos sobre como deve ser a composição de uma fraseologia para que seja considerada especializada, mas sim manter o foco no objetivo em questão, qual seja, o de tentar facilitar o trabalho tradutório no momento em que o tradutor se depara com sequências lexicais, independentemente da presença ou ausência de um termo.

## 2.2. LINGUÍSTICA DE CORPUS (LC)

A crescente conversão de pesquisas com base em dados empíricos em lançamentos impressos e digitais evidencia o maior espaço ocupado pela Linguística de Corpus (LC) como fonte de contribuições a respeito do funcionamento da língua e, como consequência disso, o interesse de editoras em disponibilizar esse material aos usuários. A mudança de ponto de vista pode ser verificada na quantidade de livros acadêmicos, livros didáticos, dicionários, gramáticas e glossários, para mencionar alguns exemplos, oferecidos no mercado. É evidente que isso tudo conta com grande aliada: a tecnologia em constante avanço. O aumento da capacidade de armazenamento em computadores e o desenvolvimento de ferramentas de processamento cada vez mais eficientes permitem que certo corpus possa crescer

indefinidamente e, ao mesmo tempo, prover respostas atualizadas de maneira rápida e inequívoca, elevando a confiabilidade dos produtos.

A atmosfera de prosperidade da Linguística de Corpus, descrita no parágrafo anterior, é muito recente e, portanto, não pode ser atribuída a toda sua trajetória. O início, passou por duras críticas teóricas e práticas, fatos discutidos a seguir.

### 2.2.1. Percurso da LC

Por mais surpreendente que possa parecer, a realização de buscas por informações em grande quantidade de textos teve início no século XIII, momento em que a Bíblia Cristã obteve, manualmente, indexação de palavras de cada frase, de todas as páginas (O’Keeffe, McCarthy, 2010, p. 3). Diante disso, Anthony of Padua, escreveu *Concordantiae Morales* (passagens concordantes), com base na Vulgata, o que pode ser considerado a pedra fundamental das concordâncias de que se tem conhecimento (O’Keeffe, McCarthy, 2010, p. 3; Peck, 1847, p. 452). Entretanto, um frei dominicano chamado Hugh of Saint-Cher, da França, auxiliado por quinhentos monges, foi quem completou a *Correctorium Bibliarum* (Correção dos Textos da Bíblia): dividiu o texto nos capítulos até hoje conhecidos e preparou um índice para se chegar às passagens em que certa palavra seria encontrada (O’Keeffe, McCarthy, 2010, p. 3). Esse projeto escrito em latim, iniciado em 1230 e publicado em 1238, é conhecido como o fundador dos concordanciadores (Peck, 1847, p. 453). Como era apenas um índice de palavras, foi avaliada como inútil por não facilitar o trabalho dos pregadores, criando a necessidade de ser revista, atividade que foi implementada entre os anos de 1250 e 1253, por freis Dominicanos ingleses, que adicionaram as citações completas das passagens indicadas (Reese, 2011, p. 17).

Na primeira metade do século XX, momento em que a coleta de dados para estudos era realizada manualmente, “[...] foi a linguística estrutural americana que lançou as bases da linguística de corpus como metodologia empírica na observação de dados [...] (Llamazares, 2008, p. 331, tradução minha<sup>74</sup>).” Evidentemente, uma atividade exaustiva como essa não estava isenta de falhas: o grande volume de textos sob análise era elemento facilitador para que dados passassem despercebidos, visto que total atenção humana era necessária para que nada escapasse aos olhos (Johansson, 2011, p. 17).

---

<sup>74</sup> [...]fue la lingüística estructural americana la que sentó las bases de la lingüística de corpus como metodología empírica basada en la observación de datos [...].

Llamazares (2011, p. 331) esclarece que os estudiosos dessa época como, por exemplo, John Firth, e Michael Halliday, seu seguidor, acreditavam em linguagem como atividade social e que, com isso, o corpus era a única ferramenta capaz de fornecer os dados necessários para que as línguas fossem descritas minuciosamente.

No final da década de 1950, mais precisamente no ano de 1957, entram em cena os postulados de Chomsky, relegando a palavra “empirismo” para segundo plano e conferindo protagonismo à Gramática Transformacional-gerativa nas décadas de 1960 e 1970 (Rajagopalan, 2007, p. 24). Nesse período, Chomsky rechaça as descrições de linguagem em uso e se posiciona contra a observação de dados em geral, especialmente a partir da linguística de corpus, fato que motivou a instauração de pesquisas com o objetivo de identificar universais linguísticos (Tognini-Bonelli, 2010, p. 14). O racionalismo do novo modelo privilegiava as construções sintáticas consideradas “gramaticais”, mesmo que o resultado fosse algo não produzido naturalmente pelos falantes, distanciando-se cada vez mais da realidade da vida (Rajagopalan, 2007, p. 31).

É sabido que a condução de pesquisas sob os preceitos da linguística de corpus seria mais laboriosa sem a utilização de computadores e de programas de processamento de dados, visto que o tamanho a que os atuais corpora podem chegar é algo surpreendente. Tudo começou, relata Randall 5th (2006, p. 1), com o ENIAC (*Electrical Numerical Integrator And Calculator*), o primeiro computador eletrônico de grande porte, lançado em 13 de fevereiro de 1946 na *Moore School of Electronics*, da *Pennsylvania University*. Eram 30 toneladas com 180m<sup>2</sup> de área construída, que ocupavam uma sala de 93m<sup>2</sup>. Funcionava com de 17.468 válvulas e realizava 5.000 adições ou 300 multiplicações por segundo. Não se sabe exatamente quem inventou qual parte do computador, mas à frente do projeto estavam J. Presper Eckert and John W. Mauchly, configurando a primeira das 5 gerações de computadores, sistematizadas a seguir<sup>75</sup>:

- **1ª geração:** 1946 a 1954 – Utilização de Válvulas eletrônicas.
- **2ª geração:** 1955 a 1964 – Utilização de Transistores.
- **3ª geração:** 1965 a 1977 – Utilização de Circuitos Integrados/Microchips de Grande Escala de Integração.
- **4ª geração:** 1977 a 1991 – Utilização de Microprocessador com Circuitos Integrados/Microchips de Escala Muito Grande de Integração.

<sup>75</sup> <http://producao.virtual.ufpb.br/books/camyle/introducao-a-computacao-livro/livro/livro.chunked/ch01s02.html>  
<http://www.ctec.ufal.br/professor/cfs/introducao%20nocoas%20de%20logica%20e%20algoritmos.pdf>

- **5ª geração:** 1991 a atual – Supercomputadores com utilização de circuitos integrados/microchips de Escala Ultra Grande de Integração.

Com esse mapeamento das gerações dos computadores, pode-se verificar a capacidade dos recursos tecnológicos disponíveis em cada época, fator necessário para alocar a segunda geração da linguística de corpus, que, mesmo com o ambiente desfavorável das décadas de 1960 e 1970, passa a contar com esses equipamentos como forte aliado em estocar, acessar, transportar, ler e analisar eletronicamente textos/dados (Johansson, 2009, p. 33).

De acordo com o mencionado anteriormente, nos anos de 1960 a Gramática Gerativa conquistou a adesão dos linguistas, que passaram a rejeitar iniciativas de estudos de outras abordagens. Ainda assim, nesse mesmo período, mais especificamente em 1964, W. Nelson Francis e Henry Kučera criaram o primeiro corpus eletrônico de linguagem escrita, compilado para análise linguística, O Brown Corpus, fato de extrema relevância para o desenvolvimento da linguística de corpus por estabelecer um padrão de compilação e uso de corpus eletrônico (Garcia, 2012, p. 1; Johansson, 2011, p. 18, 2009, p. 33; Tognini-Bonelli, 2010, p. 14; Meyer, 2009, 2004, p. 1; Fernández, 2007, p. 06).

A partir da década de 1970, os estudos em tecnologia aceleraram o desenvolvimento dos computadores, aparelhando a linguística de corpus com ferramentas adequadas para o desenvolvimento de pesquisas confiáveis. É evidente que o tamanho exagerado, baixas capacidade e velocidade de processamento somados ao preço elevado dos computadores da primeira geração eram fatores que não agradavam a ninguém. Porém, todas essas oposições de compra foram mitigadas nas gerações seguintes, principalmente a partir do lançamento dos computadores pessoais, os PCs (Fernández, 2007, p. 9).

A partir dos anos 1980, os avanços tecnológicos contribuíram com rápidas ampliações na quantidade de estudiosos, variedade de textos e diversificação nos propósitos de pesquisas, no âmbito da linguística de corpus (Johansson, 2009, p. 33). O aumento, inclusive, da capacidade de estocagem de dados, possibilitou que John Sinclair e seus colaboradores, em 1987, dessem andamento ao projeto COBUILD *English Language Dictionary*, desenvolvido a partir da compilação de um corpus grande (Johansson, 2011, p. 17; Stubbs, 2011, p. 8).

A diminuição do tamanho do computador em oposição ao aumento da capacidade de armazenamento dos circuitos integrados, associados ao lançamento de ferramentas de análise de corpora digitais, motivaram, já na década de 1990, a proliferação dos computadores pessoais e o alavancamento dos estudos da linguística de corpus (Garcia, 2012, p. 11). Esses elementos permitiram que fossem construídos corpora orais e escritos extremamente grandes, com milhões de palavras, com a possibilidade de utilização de ferramentas de buscas rápidas e

precisas e sem erros, configurando, assim, o início da derrubada das críticas elaboradas pelos seguidores da Linguística Gerativa (Fernández, 2007, p. 9).

O estudo de dados reais por meio de corpus processados em computador é sintetizado por Tognini-Bonelli e Sinclair (2006, p. 208), da seguinte maneira:

1. Os 20 anos iniciais (1960 a 1980): aprender como compilar e armazenar um corpus de milhões de palavras; não há material em formato eletrônico, tudo tem que ser digitado;
2. Os 20 anos seguintes (1980 a 2000) são divididos em 2 décadas:
  - a) 1980: com scanner, é possível compilar 20 milhões de palavras;
  - b) 1990: aumento da possibilidade de viabilizar corpora ainda maiores.
3. Novo milênio: disponibilização de textos na internet.

Depois de todo esse percurso, é possível verificar a ascensão e a parceria de duas vertentes: uma tecnológica e outra linguística. Ambas perfeitamente estabelecidas e à disposição dos usuários. A **primeira**, por meio do computador, era algo de difícil alcance que passou a ser indispensável no âmbito das investigações acadêmicas, deixando de ser uma calculadora enorme para se tornar provedor de respostas e de conectividade. A **segunda**, a linguística, mais especificamente a linguística de corpus, que deixou de ser questionável para assumir posição de reveladora de informações significativas a partir de um dado corpus. Juntas construíram alicerces concretos e, apesar de ainda muito jovens, podem viabilizar descobertas altamente valiosas. A partir disso, parcerias entre a Linguística de Corpus e outras áreas linguísticas são criadas de maneira exponencial, continuamente ampliando o escopo de possíveis interfaces nas pesquisas. A respeito dessa prosperidade, Garcia (2012, p. 12) relata esses avanços ao mencionar que

[a] diversidade de enfoques que se pode privilegiar a partir daí é incomensurável. Vemos aqui novos corpora, novas abordagens de codificação, ferramentas de análise inovadoras, discussões sobre conceitos básicos e pesquisas específicas envolvendo metáforas, expressões fixas, textos históricos, linguagens especializadas, linguagem de aprendizes, linguagem oral, tradução, lexicografia, terminologia, análise do discurso e ensino de línguas.

### 2.2.2. Estatuto da LC

Para se chegar à conclusão de que existe debate em torno do estatuto da Linguística de Corpus não é preciso empenhar muito esforço. Para tanto, segundo Shepherd (2009, p. 150), é suficiente verificar a seção inicial de diversos livros da área para se comprovar a alta frequência

com que, há muito tempo, esse assunto entra em pauta para ser discutido. Pelas características que tem e pelo profundo entendimento de fenômenos linguísticos a que pode levar, a exata localização da Linguística de Corpus, entre teoria e metodologia, ainda é algo nebuloso. De acordo com Taylor (2008, p. 180), há várias definições para LC, assim como várias propostas para tentar caracterizar o que ela é, o que inclui: “[...] LC é uma ferramenta, um método, uma metodologia, uma abordagem metodológica, uma disciplina, uma teoria, uma abordagem teórica, um paradigma (teórico ou metodológico), ou uma combinação disso” (Tradução minha<sup>76</sup>). Zilio (2010, p. 134) esclarece que essa dissensão acontece porque alguns autores acreditam se tratar somente metodologia, enquanto outros defendem ser disciplina, visto que também possui aporte teórico além das ferramentas que utiliza. Na tentativa de justificar seus entendimentos a respeito da LC, podemos citar alguns autores:

1. Meyer (2004, p. 2, tradução minha<sup>77</sup>): “[...] é errado supor que a análise de corpora em nada pode contribuir com a teoria linguística: corpora podem ter inestimáveis recursos para testar hipóteses linguísticas com base em teorias de gramática mais funcionalmente embasadas, i.e. teorias de linguagem mais interessadas em explorar linguagem como ferramenta de comunicação” C
2. Hunston (2006, p. 244, tradução minha<sup>78</sup>): “[...] linguística de corpus é essencialmente uma metodologia ou conjunto de metodologias, em vez de ser uma teoria de descrição da linguagem.”
3. McEnery, Xiao, Tono (2006, p. 7, tradução minha<sup>79</sup>): “É principalmente considerada como metodologia que tem desenvolvido seus métodos sistemáticos e princípios para a aplicação de corpora nos estudos de linguagem em uso; então, é uma metodologia com “status teórico” usada em muitas áreas e teorias linguísticas.”
4. Oliveira (2009, p. 48): “[...] a Linguística de Corpus é uma área que permite o aprofundamento sobre o conhecimento empírico de diferentes línguas estudadas, levando a novas concepções teóricas sobre a linguagem, não podendo ser considerada, portanto, apenas como uma metodologia de análise.”

---

<sup>76</sup> [...] corpus linguistics is a tool, a method, a methodology, a methodological approach, a discipline, a theory, a theoretical approach, a paradigm (theoretical or methodological), or a combination of these.

<sup>77</sup> A collection of texts held in a machine-readable form and capable of being analysed automatically or semi-automatically in a variety of ways [...] it is put together for a particular purpose and according to explicit design criteria in order to ensure that it is representative of the given area or sample of language it aims to account for.

<sup>78</sup> [...] corpus linguistics is essentially a methodology or set of methodologies, rather than a theory of language description.

<sup>79</sup> It is mainly regarded as a methodology which has developed its own systematic methods and principles for the application of corpora to studies of language use; hence, it is a methodology with a “theoretical status” used in many areas and theories of linguistics.

5. Finatto, Evers, Alle (2010, p. 158): “Mostra-se como uma metodologia e também como uma abordagem diferenciada dos estudos da linguagem que se distingue, fundamentalmente, dos estudos de herança gerativista.”

De acordo com a discussão apresentada anteriormente, nesta seção, é possível depreender que a condição “híbrida” da LC seja o fator determinante para que haja pouca clareza a respeito de seu status. Shepherd (2009, p. 150) “[...] sugere que o status de ‘ramo da linguística’ ou ‘metodologia de tratamento de dados’ pode ser estabelecido ao se decidir o ponto de entrada em dados oriundos de textos eletrônicos [...].”

### 2.2.3. O Conceito de Corpus

A característica que diferencia o corpus da LC dos corpora de outros estudos linguísticos, apontam Aluísio e Almeida (2006, p. 157), é o formato basilar dos textos, ou seja, enquanto estes aceitam que as pesquisas sejam realizadas em ambiente concreto, aquela depende, exclusivamente, que o corpus esteja em formato digital. Essa distinção é nítida e passou a integrar as definições de corpus sem dispensar a presença do computador, como apontado pelos estudiosos mencionados a seguir:

1. Baker (1995, p. 225, tradução minha<sup>80</sup>): “Uma coleção de textos em formato legível por computador, capaz de ser analisado automaticamente ou semiautomaticamente em uma variedade de maneiras [...] é construído por um propósito em particular e de acordo com critérios explícitos para assegurar que é representativo de uma certa área ou amostra de linguagem que objetiva estudar.”
2. Hunston (2006, p. 234, tradução minha<sup>81</sup>): “Um corpus é uma coleção de amostras de linguagem que ocorrem naturalmente, armazenadas eletronicamente.”
3. Bennet (2010, p. 14, tradução minha<sup>82</sup>): “Um corpus é uma coleção de textos autênticos, selecionados por meio de critérios, armazenados eletronicamente.”
4. Reppen (2010a, p. 02, tradução minha<sup>83</sup>): “No mundo da linguística de corpus, corpus é uma grande quantidade de textos (escritos ou falados) que ocorrem naturalmente, coletados por meio de critérios, armazenados automaticamente.”

---

<sup>80</sup> A collection of texts held in a machine-readable form and capable of being analysed automatically or semi-automatically in a variety of ways [...] it is put together for a particular purpose and according to explicit design criteria in order to ensure that it is representative of the given area or sample of language it aims to account for.

<sup>81</sup> A corpus is an electronically stored collection of samples of naturally occurring language.

<sup>82</sup> A corpus is a principled collection of authentic texts stored electronically.

<sup>83</sup> In the world of corpus linguistics, corpus is a large, principled collection of naturally occurring texts (written or spoken) stored automatically.

Com isso, é possível dizer que as definições de corpus para a LC sofreram importantes modificações ao longo do percurso de sua utilização, já que, como discutido anteriormente, no início das pesquisas desse âmbito, o corpus somente era possível em formato concreto. Com o crescimento da LC, mediante o avanço da tecnologia em máquinas eletrônicas, explica Baker (1995, p. 225), a definição de corpus é pode ser alterada de 3 maneiras, ao revelar que ele: 1. deve ter formato possível de ser lido por computador; 2. passa a contar com textos orais; 3. passa a contar com variedade de textos e autores.

Diante da ampliação verificada nos conceitos mencionados anteriormente, há evidências dos requisitos que um corpus deve apresentar. De acordo com Finatto, Evers e Alle (2010, p. 158), Reppen (2010b, p. 31), Llamazares (2008, p. 340-1), Sinclair (2005), Biber, Conrad e Reppen (1998, p. 04): formato digital, dados autênticos, critérios de seleção, tamanho e representatividade, exigências que, a seguir, passo a esclarecer.

1. **Formato digital:** para extração de informações de um corpus, os dados devem estar em formato digital, o que possibilita processamento de dados com ferramentas avançadas, velozes e confiáveis. Como tudo na área de informática, os “*software*” passam por atualizações e ficam cada vez mais simples na utilização e mais eficientes ao fazer buscas e prover resultados.
2. **Dados autênticos:** os textos a serem compilados devem ter origem empírica, ou seja, devem ser amostras reais de uso.
3. **Critério de seleção:** toda seleção de textos deve ser realizada a partir de critérios adequados e atrelados a um objetivo concreto. Alguns critérios gerais, listados por Sinclair (2005, p. 1) para a escolha de textos são:
  - i. modo: menciona se a origem da linguagem é escrita, falada ou em modo eletrônico;
  - ii. tipo: se escrito, nota, livro, carta, jornal, etc.;
  - iii. domínio: acadêmico, popular, etc.;
  - iv. língua, línguas ou variação de língua do corpus;
  - v. localização: o Português do Brasil, de Portugal, do Nordeste do Brasil, etc.;
  - vi. data dos textos.
4. **Tamanho:** por maior que seja, um corpus sempre vai ser considerado um recorte, uma amostra, ao ser comparado com a possível infinidade de dados a que a linguagem pode chegar. Entretanto, deve ser representativo para o objetivo a que se

propõe estudar. De acordo com Kübler (2002, P. 44, tradução minha<sup>84</sup>), “Corpora menores, explorados com ferramentas simples, produzem resultados interessantes em uma escala mais individual.” O’Keeffe e McCarthy (2010, p. 06, tradução minha<sup>85</sup>) também concordam que um corpus pequeno com objetivo claro é “[...] ferramenta poderosa para investigação de usos especiais de linguagem, em que o linguista pode se aprofundar nos dados detalhadamente usando um arsenal de programas e descobrir usos particulares de linguagem.” Como visto, o propósito da construção do corpus é influente na decisão de seu tamanho, e a questão da autenticidade é uma das grandes vantagens nos pequenos, capazes de revelar conexões entre o padrão linguístico e o contexto de uso (Koester, 2010, p. 66; Gilquin e Granger, 2010, p. 359).

5. **Representatividade:** garantia de que os textos selecionados possam representar a língua ou a variedade linguística selecionada para estudo. É conceito ligado ao tamanho do corpus e do objetivo da pesquisa. Olohan (2004, p. 44, tradução minha<sup>86</sup>) afirma que “é difícil confirmar, com algum grau de convicção, que seus dados são representativos de determinada língua ou gênero.” Como ambos, tamanho e representatividade são conceitos que alojam muitas variáveis, para atendê-los deve-se recorrer ao bom senso do pesquisador. Quanto à representatividade de um corpus pequeno, de determinado gênero, como é o caso do utilizado neste estudo, Koester (2010, p. 69, 71, tradução minha<sup>87</sup>) esclarece que

[s]e um tipo de gênero muito específico está sendo investigado, então pode ser muito simples estabelecer representatividade situacional, visto que todas as amostras compiladas vão corretamente representar esse gênero. [Isso acontece porque] corpora especializados são frequentemente desenhados para responder a específicas perguntas de pesquisa. [...] Se todas as amostras vierem de apenas uma organização, então o corpus será representativo do gênero como usado naquela organização, mas não do gênero como um todo.

<sup>84</sup> Smaller corpora, exploited with simple tools, produce interesting results on a more individual scale.

<sup>85</sup> [...]powerful tool for the investigation of special uses of language, where the linguist can ‘drill down’ into the data in immense detail using a full armoury of software and shed light on particular uses of language.

<sup>86</sup> [...] it is often difficult to confirm with any degree of conviction that their data is ‘representative’ of a particular language or genre.

<sup>87</sup> If a very specific type of genre is being investigated, then it may be quite straightforward to establish situational representativeness, as all the samples collected will accurately represent that genre. [...] specialised corpora are often designed to answer specific research questions.

6. **Visão probabilística** da linguagem. Deve ser possível a verificação da frequência da aparição de uma palavra, ou sequência de palavras (padrões) em um sistema de probabilidades.

Verificada a conceituação e os requisitos necessários em torno do corpus, é possível propor uma definição. Sinclair (2005, p. 1, tradução minha<sup>88</sup>), por exemplo, afirma que “[u]m corpus é uma coleção de pedaços de textos linguísticos em formato eletrônico, selecionados de acordo com critérios externos para representar, como possível, uma língua ou variedade de língua como fonte de dados para pesquisa linguística.” (Para acesso a outras definições, ver Llamazares, 2008, p. 342<sup>89</sup>).

#### 2.2.4. Baseado em Corpus ou Dirigido pelo Corpus

Segundo Tognini-Bonelli (2001), o estudo de corpus está submetido a dois tipos de abordagem no que diz respeito à utilização do corpus, a saber: 1. baseado em corpus (*corpus-based*) ou 2. dirigido pelo corpus (*corpus-driven*). Segundo a autora (2001, p. 65, 85, tradução minha<sup>90</sup>), o primeiro tipo busca “[...] detalhar, testar ou exemplificar teorias e descrições que foram formuladas antes que corpora grandes fossem disponibilizados para uniformizar o estudo linguístico [...]”. Diferentemente do que acontece no segundo tipo de abordagem em que “[...] as afirmações teóricas são consistentes e refletem diretamente a evidência fornecida pelo corpus.”

Sobre isso, Biber (2012, p. 5, tradução minha<sup>91</sup>) assevera que “[...] a pesquisa baseada em corpus mostra que qualquer característica ou variante linguística é distribuída e usada de maneiras muito diferentes em diferentes registros”, apontando que não é possível haver uma descrição linguística competente sem considerar o registro em que está inserida. Por outro lado, ainda de acordo com Biber (2012, p. 10), o tipo dirigido pelo corpus, identifica palavras, categorias e variantes de uma mesma palavra em ambiente empírico, investigando seus padrões e variações de significado.

<sup>88</sup> A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research.

<sup>89</sup> <https://dialnet.unirioja.es/descarga/articulo/3332675.pdf>

<sup>90</sup> [...] expound, test or exemplify theories and descriptions that were formulated before large corpora became available to inform language study [...].

[...] the theoretical statements are fully consistent with, and reflect directly, the evidence provided by the corpus.

<sup>91</sup> [...] corpus-based research has shown that almost any linguistic feature or variant is distributed and used in dramatically different ways across different registers.

Diante dos elementos expostos nesta parte do capítulo, é possível sintetizar os postulados das abordagens mencionadas, como mencionado no Quadro 1 a seguir:

**Quadro 1 - Distinção entre com Base e Dirigido.**

<b>Abordagem baseada em corpus (<i>Corpus-based</i>)</b>	<b>Abordagem dirigida pelo corpus (<i>Corpus-driven</i>)</b>
Enxerga LC como ferramenta.	Enxerga LC como teoria.
Corpus como metodologia.	Corpus como ponto inicial para se desenvolver uma teoria.
Corpus para validar ou corrigir outras teorias.	Corpus fornece novas hipóteses e teorias.
Focaliza a classe gramatical e a sintaxe.	Focaliza a palavra.
Tende a utilizar marcações.	Tende a utilizar “ <i>raw corpus</i> ”.
Corpus pra verificar o conhecimento intuitivo.	Não confia na intuição do pesquisador.
Corpora de tamanhos variados.	Corpora muito grandes.
Corpus representativo do gênero em estudo.	Corpus representativo.

**Fonte:** Biber (2012); Cheng (2012); Candlin e Carter (2005); Tognini-Bonelli (2001).

### 2.2.5. Tipos de Corpus

A grande quantidade de estudos verificada na área de LC evidencia certa multiplicidade nos tipos de corpus colocados em aproximação. O ponto de entrada na pesquisa pode variar, já que essa tipologia está associada ao objetivo que a pesquisa pretende atingir. No entanto, existem alguns parâmetros eficazes no sentido de enquadrar a tipologia a que determinado corpus pertence.

Baker (1995, p. 229, tradução minha<sup>92</sup>) estipula alguns critérios de seleção sob os quais um corpus pode ser compilado. “[...] os mais importantes são:

- (i) Língua geral vs domínio restrito;
- (ii) Língua escrita vs falada
- (iii) Sincrônico vs diacrônico
- (iv) Tipicamente em termos de fontes (escrito/falantes) e gêneros (por exemplo: editoriais de jornais, entrevistas em rádio, ficção, artigos de jornal, audiências em tribunais)

<sup>92</sup> [...] the most important of which are: (i) General language vs restricted domain; (ii) Written vs spoken language; (iii) Synchronic vs diachronic; (iv) Typically in terms of range of sources (written/speakers) and genres (e.g. newspapers editorials, radio interviews, fiction, journal articles, court hearings); (v) Geographical limits, e.g. British vs American English; (vi) Monolingual vs bilingual or multilingual.

- (v) Limites geográficos, por exemplo: Inglês britânico vs americano
- (vi) Monolíngue vs bilíngue ou multilíngue.”

Em seguida, Baker (1995, p. 154, 230), revela que, para a pesquisa em tradução, há 3 tipos de corpora: Paralelo (textos em determinada língua, alinhados com suas respectivas traduções para outra), Multilíngue e Comparável (dois arquivos de textos na mesma língua, um de documentos originais e outro de traduções, similares em domínio, variedade linguística, período de tempo e possível de ser comparado). A respeito do corpus, Kenning (2010, p. 493) explica que o comparável tem a vantagem de conter somente linguagem não mediada, isto é, apenas naturalmente corrente, o que aumenta a confiabilidade. A autora menciona que também nesse tipo de corpus pode haver exame da influência do texto fonte no texto alvo, indicando, para isso, o trabalho de Baker (1993 e 1995) podem auxiliar a destacar as características distintivas do texto traduzido, ou seja: **1) Simplificação** do texto para fácil e rápido entendimento; **2) Explicação**: retira subentendidos, elipses e provê explicações; **3) Normalização**: retira diversificação lexical, homogeneíza o texto, busca atender expectativas do leitor-alvo; **4) Nivelamento**: retira seleção lexical que possa dificultar compreensão, tenta evitar ambiguidade e se utiliza de elementos da normalização ao neutralizar o estilo do texto fonte no texto traduzido. Sinclair (1996) também utiliza os parâmetros postulados por Baker (1995), acrescentando que um corpus comparável pode conter corpora com os mesmos critérios de compilação, mas que não são traduções. Ainda a respeito do corpus comparável, Tognini-Bonelli (2010, p. 21) e McEnery e Xiao (2007, p. 3) ressaltam que os subcorpora desse tipo de corpus não são traduções um do outro, do contrário seriam paralelos, e devem conter iguais proporções, gênero textual, domínio e período de tempo.

Verificando a proliferação de tipologia de corpus, Fernandes (2006, p. 87) relata a possibilidade de haver estudos que não se enquadravam nos critérios estabelecidos por Baker (1996, 1995) e propõe parâmetros calcados em aspectos distintivos para suprir essa necessidade, como no Quadro 2 a seguir:

**Quadro 2 - Proposta de parâmetros para classificação de corpora.**

<b>Crítérios</b>	<b>Tipologia Correspondente</b>	<b>Por haver distinção entre:</b>
1. Alinhamento	Paralelo ou Comparável?	paralelo: apresenta correspondências tradutórias e comparável: apresenta correspondências textuais;
2. Objetivo	Linguístico ou Tradutório?	estudo com base em corpus para realizar estudos linguísticos ou para investigar processos e produtos tradutórios;

3. Domínio	Geral ou Especializado?	geral: estuda a linguagem de material traduzido como um todo e especializado: investiga a linguagem de gêneros específicos traduzidos;
4. Modo	Escrito ou Falado?	falado: a partir de transcrições e escrito: compilado de uma revista, por exemplo;
5. Tempo	Sincrônico ou Diacrônico?	sincrônico: o objeto de estudo está focalizado em um período de tempo; e diacrônico: dedicado ao desenvolvimento histórico do objeto ao longo do tempo.
6. Nº de Línguas	Mono ou Bi ou Multilíngue?	
7. Direção	Uni ou Bi ou Multidirecional?	

**Fonte:** Fernandes (2006, p. 91-3)

Com o passar do tempo, a evolução rapidamente ampliou os parâmetros de classificação para estabelecer a tipologia de corpus. Hunston (2010, p. 14-16) e Llamazares (2008, p. 342) relatam considerar os seguintes elementos para se chegar ao tipo de corpus:

- (i) a **modalidade** da língua: escritos, orais, mistos;
- (ii) o **número** de línguas: monolíngues, bilíngues ou multilíngues (duas ou mais línguas).
- (iii) os **limites** do corpus: fechados e abertos ou monitor (destinados a verificar modificações na linguagem);
- (iv) a **especialidade**: gerais (muitos tipos de textos. Escritos, falados ou os dois) ou especializados (textos de um tipo determinado);
- (v) o período **temporal** que abarcam: sincrônicos e diacrônicos ou históricos (em diferentes períodos de tempo);
- (vi) o **tamanho** dos textos: referência (Quando um corpus ser um repositório confiável e razoável de todas as características de uma língua (Tognini-Bonelli, 2010, p. 20, tradução minha<sup>93</sup>) e textuais (textos inteiros);
- (vii) o **tratamento** aplicado: simples e codificados e anotados.

McEnery and Hardy (2012, p. 19) mencionam que, com a inclusão de duas ou mais línguas, há 3 tipos de corpora: Tipo A – textos em uma língua com traduções em outras duas ou mais; Tipo B – corpora em pares ou grupos ou monolíngues compilados com o mesmo tipo de enquadramento (Corpus Comparável); Tipo C – combinação de A e B.

Por ser utilizado nesta pesquisa, é importante informar que o corpus especializado não precisa ser tão grande quanto o de língua geral, mas de tamanho compatível com o que pretende

<sup>93</sup> When a sample corpus claims to be a reasonably reliable repository of all the features of a language, it can be called a reference corpus.

desvendar. Sinclair (2005, p. 18) menciona que o geral precisa ser grande por ter estreita relação com representatividade da língua, fator que exige a inclusão de muitos e variados textos, enquanto o especializado tende a ser representativo apenas de determinada área. Hunston (2006, p. 235) informa que isso pode ser explicitado facilmente ao se comparar as listas de frequência do léxico dos dois e ver que o especializado apresenta vocabulário mais concentrado.

### 2.2.6. Colocação

Estudos de corpora se ampliam e passam a investigar as frases, ou seja, o comportamento de sequências de palavras contidas em textos escritos de linguagem natural. Römer (2013, p. 4), Bennett (2010, p. 7) e Flowerdew (2010, p. 444) esclarecem que, com a padronização fraseológica como elemento central da LC, Sinclair (1991) buscava determinar que o significado de uma palavra pode ser encontrado na sequência das palavras em que está localizada. Esse estudo fraseológico envolve, segundo Palumbo (2009, p. 19), Baker, Hardie e McEnery (2006, p. 36, 135) e Olohan (2004, p. 62), quatro categorias:

1. **colocações:** relação natural uma palavra e outras que a acompanham;
2. **coligações:** forma de colocação em nível gramatical;
3. **prosódia semântica:** preferência de algumas palavras em ocorrer em ambientes positivos ou negativos;
4. **preferência semântica:** preferência das palavras em ocorrer em contextos específicos.

A terminologia em torno dessa atividade também passa pela falta de consenso entre estudiosos, visto que vários nomes têm surgido para se referirem à fraseologia/sequência lexical. Greaves e Warren (2010, p. 213) e McCarthy (2006, p. 11) citam alguns: Frases lexicais, Padrões pré-fabricados, Sequência Formulaica. Chunks / Cluster, Expressões multipalavras e n-gramas entre várias outras. Entretanto, como já mencionado no capítulo anterior, neste trabalho adoto o nome “colocação”.

A pesquisa a respeito de fraseologia foi iniciada há tempo considerável, mais precisamente na década de 1950. Interessado nas coocorrências e visando a estudos de significado, Firth é visto como o pioneiro dessa área ao introduzir o termo “colocação” (Bartsch, Evert, 2013, p. 31; Evert, 2009, p. 1212). Seu trabalho inicial visava a investigar o uso de fraseologias em cartas dos séculos XVIII e início do XIX para listar as colocações que permaneciam e as que não mais eram utilizadas (Greaves, Warren, 2010, p. 215). Segundo Firth

(1957, p. 179, tradução minha<sup>94</sup>) “[v]ocê pode conhecer uma palavra pela companhia que ela mantém”, ou seja, o significado e o uso de uma palavra podem ser classificados mediante a verificação das palavras que estão ao redor dela. Depois disso, na década de 1960, refinando a noção de Firth, Haliday também analisava os padrões de coocorrências, mas em busca de contribuir com descrições linguísticas (Vincze, 2015, p 8).

Também na década de 1960, Sinclair passa a “[...] estudar o léxico e a maneira com que ele interage com a sintaxe na construção de unidades de sentido” (Palumbo, 2009, p. 19, tradução minha<sup>95</sup>). Sinclair (1957, p. 109) postulava que, para se chegar ao significado, é necessário implementar dois modelos de interpretação: um denominado *Princípio da Livre Escolha* e o outro, *Princípio Idiomático*. No primeiro, “[...] linguagem é o resultado de um grande número de escolhas complexas”, ou seja, “[...] quando uma unidade é completada, [...] um grande número de escolhas se abre, limitado apenas pela gramaticalidade” (Tradução minha<sup>96</sup>). No segundo, as palavras têm certa organização ao formarem sequências, indicando “[...] que o usuário tem, ao seu dispor, um grande número de frases pré-construídas que constituem escolhas únicas, mesmo que parecendo ser analisáveis em segmentos” (Sinclair, 1957, p. 110, tradução minha<sup>97</sup>). De acordo com esse princípio, explica Hunston (2010, p. 143, tradução minha<sup>98</sup>), “cada palavra no texto é usada em uma fraseologia comum, em que o significado é atrelado à frase inteira mais do que a partes individuais dela, e o ouvinte ou leitor entende a frase como frase e não como um quadro gramatical com itens lexicais.”

Com a utilização de corpus é possível identificar sequências léxicas preferenciais e, por serem fenômenos probabilísticos (McCarthy, 2006, p. 9), é perceptível a dificuldade em haver uma definição rigorosa de colocação, conforme menciono algumas a seguir:

1. Sinclair (1991, p. 170, tradução minha<sup>99</sup>): “[...] ocorrência de duas ou mais palavras em um curto espaço entre uma e outra em um texto.”

---

<sup>94</sup> You shall know a word by the company it keeps.

<sup>95</sup> [...] study of lexis and the way lexis interacts with syntax in the construction of units of meaning.

<sup>96</sup> [...] language text as the result of a very large number of complex choices. At each point where a unit is completed (a word or a phrase or a clause), a large range of choice opens up and the only restraint is grammaticalness.

<sup>97</sup> [...] that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments.

<sup>98</sup> [...] each word in the text is used in a common phraseology, meaning is attached to the whole phrase rather than to the individual parts of it, and the hearer or reader understands the phrase as a phrase rather than as a grammatical template with lexical items in

<sup>99</sup> [...] the occurrence of two or more words within a short space of each other in a text

2. Baker, Hardie, McEnery (2006, p. 36, tradução minha<sup>100</sup>): “[...] colocação é o fenômeno ao redor do fato de que certas palavras são mais propensas a ocorrer em combinação com outras palavras em certos contextos.”
3. Palumbo (2009, p. 19, tradução minha<sup>101</sup>): “O termo colocação se refere à tendência de certos segmentos de palavras em ocorrer regularmente juntas, de maneira que o significado de uma palavra pode ser identificado em relação às palavras com que ‘coloca’.”
4. O’Dell, McCarthy (2008, p. 04, tradução minha<sup>102</sup>): “Colocação significa uma combinação natural de palavras; refere-se a maneira [...] com que palavras são proximamente associadas umas com as outras.” De acordo com os autores (2008, p. 06), as colocações podem ser fixas, isto é, não aceitam alterações, ou dispostas em uma gradação entre fraco, compostas por palavras que colocam com muitas outras, e forte, composta por palavras que colocam com poucas outras.
5. Hunston (2010, p. 68, tradução minha<sup>103</sup>): “Colocação é a propensão das palavras a serem tendenciosas na maneira em que coocorrem.”

A colocação pode contar com variadas classes gramaticais que, de acordo com O’Dell e McCarthy (2008, p. 10), podem ser categorizados como a seguir:

- a. Verbo + Substantivo = Our lawyer *drew up a contract* for us to sign.
- b. Substantivo + Verbo = An *opportunity arose* for me to work abroad.
- c. Substantivo + Substantivo = By *a stroke of luck* I found my keys.
- d. Adjetivo + Substantivo = The *simple/plain truth* is that nobody was aware of it.
- e. Advérbio + Adjetivo = She was a *stunningly attractive* woman!
- f. Verbo + Advérbio ou Frase Preposicional = As soon as the singer came on the stage she *burst into song*.

Outro desafio enfrentado pelos linguistas é o de determinar o tamanho exato de uma fraseologia, ou seja, o total de elementos lexicais necessários para serem entendidos como colocação. Para decidir que se trata de um padrão, quantas palavras devem ser atraídas em direção a outra que, inicialmente, ocupe a posição central? Qual a melhor saída? Arbitrariamente decidir por certo número ou simplesmente verificar a capacidade de

<sup>100</sup> [...] collocation is the phenomenon surrounding the fact that certain words are more likely to occur in combination with other words in certain contexts.

<sup>101</sup> The term collocation refers to the tendency of certain sets of words to occur regularly together, in such a way that the meaning of an individual word can be identified in relation to the words it “collocates with”.

<sup>102</sup> Collocation means a natural combination of words; it refers to the way [...] words are closely associated with each other.

<sup>103</sup> Collocation is the tendency of words to be biased in the way they co-occur.

processamento oferecida nas linhas de concordâncias de determinado software? Em busca de respostas e esse questionamento, Sinclair, Jones e Daley (2004, p. 05) mencionam que a maioria das palavras que ocupam posição central, também conhecidas como nódulos ou bases, não atrai muitas outras ao redor, chamadas colocados, mais do que quatro ou cinco posições além da base, algo que, segundo os autores ainda deveria ser verificado com mais rigor. Em seguida, chegam ao veredito, revelando que investigações “[...] mostraram que extensão bem sucedida (*span*) possui quatro palavras em cada lado da base. Mais curta perderia valiosas evidências e mais longa sobreporia padrões relevantes com material mais distante” (Tradução minha<sup>104</sup>).

---

<sup>104</sup> [...] showed that the optimum extent (called span) was four words on either side of the node. A shorter span would miss valuable evidence, and a longer one would overlay the relevant patterns with more distant material.

### 3. METODOLOGIA

Cabré (1999, p. 121) estabelece os limites dentro dos quais há o início do trabalho terminológico, ao esclarecer que a “[c]omunicação oral e escrita, técnica e científica é a fonte básica do material para extração de termos” (Tradução minha<sup>105</sup>). Nessa comunicação, especialistas da respectiva área criam designações a partir dos conceitos que, em seguida, são analisadas pelos terminólogos que, ao aceitá-las sem ressalvas, começam a extraí-las de seu habitat e passam a organizá-las em listas preliminares. Nessa etapa, as fontes devem ser **representativas**, de modo a prover grande quantidade de listas iniciais; **atuais**, contendo designações realmente utilizadas pelos especialistas; e **explícitas**, para permitir facilmente a recuperação de dados (Cabré, 1999, p. 121). O próximo passo trata do registro dos dados em fichas, tendo em mente organizar as informações colhidas. São quatro tipos de fichas: 1) **de extração**, contendo o termo e informações adicionais; 2) **terminológica**, com vistas a cumprir os objetivos do trabalho; 3) **de correspondência**, com as designações que se referem a um mesmo conceito; e 4) **de dúvida**, contendo busca terminológica a partir de questionamento ou imprecisão.

Definidos os textos, estabelecidos os objetivos, estipulados os modelos de fichas, começa a etapa de busca de termos, o que, segundo Cabré (1999, p. 29, 130, 151), obedece a um dos quatro métodos descritos a seguir: 1) **busca sistemática monolíngue**: o terminólogo se dedica a grande quantidade de termos de uma área em uma língua, nas etapas a seguir: a) delimitar o que vai ser feito; b) preparar-se para a tarefa; c) preencher as fichas mencionadas anteriormente; d) apresentar o projeto; e) supervisionar o projeto; f) resolver os pontos problemáticos; 2) **busca sistemática multilíngue**: os mesmos passos acima para cada uma das línguas envolvidas; 3) **busca monolíngue ad-hoc**: tentativa de encontrar o termo para expressar o conceito ou o equivalente. Dedicar-se a somente um termo, para solucionar um problema; ou 4) **busca multilíngue ad-hoc**.

Também no que se refere à extração terminológica, podemos mencionar a proposta por Krieger, Maciel e Finatto (2000, p. 149-152) ao expor sequência detalhada para a criação de produtos impressos e digitais a partir de textos sobre Leis do Meio Ambiente. É possível afirmar que o trabalho foi dividido em duas etapas consecutivas, como segue: **Primeira etapa: 1)** Estabelecimento da árvore de domínio, que funciona como plano referencial para identificar os termos de uma especialidade; **2)** Constituição da nomenclatura a partir do reconhecimento da

---

<sup>105</sup> Oral and written, technical and scientific communication is the basic source material for extracting terms.

natureza dos termos; **3)** Utilização do dicionário como obra de referência; **4)** Atendimento ao critério de pertinência pragmática; **5)** Consulta a especialistas; **6)** Formulação de definições; **7)** Categorização das definições em: a) definição extraída do corpus; b) definição extraída de fontes bibliográficas; c) definição elaborada pela equipe de trabalho; **8)** Redação das definições. **Segunda etapa:** **9)** Constituição de um corpus em língua estrangeira; **10)** Alinhamento de conceitos para aplicação de equivalência entre termos. Não havendo identificação de termos equivalentes, recorre-se à paráfrase ou à omissão e para nomes de entidades, organismos, associações, programas e planos, recorre-se à tradução literal.

Também sob o escopo da Terminologia, Almeida e Vale (2008) e Almeida, Aluísio e Oliveira (2007), reforçam a ideia de como devem ser os procedimentos metodológicos sob esse viés, e estipulam a sequência de etapas que consideram constitutivas de qualquer projeto terminológico com fins terminográficos, quais sejam: **1)** Seleção das fontes; **2)** Extração manual ou automática dos termos; **3)** Elaboração de mapa conceitual (ontologia); **4)** Inserção de termos na ontologia e validação pelos especialistas; **5)** Elaboração e preenchimento de fichas terminológicas; **6)** Elaboração e incremento da base definicional; **7)** Elaboração das definições e informações enciclopédicas; e **8)** Edição dos verbetes. Como se pode perceber, trata-se de uma metodologia para a criação de verbetes em direção a privilegiar as **definições** como foco do trabalho.

Ao lado da Terminologia, a Linguística de Corpus (LC) também oferece encaminhamentos metodológicos para o processamento de um corpus, expediente muito utilizado por vários profissionais, incluindo tradutores. Linhas de concordanciadores, listas de palavras-chave, cálculos estatísticos, verificação das colocações e anotações são exemplos de recursos proporcionados pela LC, gerados de maneira automática por programas como o *Wordsmith tools*<sup>106</sup> (Scott, 1999), que podem ajudar o analista a chegar a resultados e respostas interessantes a respeito das diferenças entre os textos dos corpora sob análise, principalmente quando se tratar de comparações entre os especializados (Hunston, 2010, p. 67-68).

Tendo em mente a constituição de um corpus para a extração de fraseologias de várias áreas, Tagnin (2002a) decide contar com a utilização dos recursos tecnológicos apregoados pela LC. Inicia um projeto fraseológico tendo como meta auxiliar os tradutores no desempenho de sua função profissional. As etapas do processo que explicitam a maneira como os dados foram extraídos do corpus bilíngue comparável são as seguintes: **1)** busca de textos, originais ou traduções, tradicionais ou digitais, em português e em inglês, de várias áreas, inclusive as

---

<sup>106</sup> <http://www.lexically.net/wordsmith/>

altamente especializadas; **2)** escolha de um tópico dentro de cada área; **3)** identificação de fraseologias e tentativa de traduzi-las com utilização de recursos tecnológicos como os programas Word, Excel, *Simple Concordance Program* e o *IntraText*<sup>107</sup>; **4)** realização de buscas no *WebCorp*<sup>108</sup> e no BNC<sup>109</sup> para confirmar colocações; **5)** organização dos verbetes contendo termos equivalentes e exemplos autênticos em ambas as línguas, sem a presença de definições.

As etapas descritas no parágrafo anterior representam a primeira parte do desenvolvimento da pesquisa. O passo seguinte teve a seguinte organização: **6)** redelimitação da área de pesquisa; **7)** processamento do corpus com a utilização do programa *WordSmith Tools* demo; **8)** introdução de um cabeçalho para a identificação de cada texto; **9)** organização dos textos em três grupos: inglês, português e paralelos. O corpus paralelo foi subdividido de acordo com a direção da tradução; **10)** inclusão de termos monolexêmicos; e **11)** apresentação do glossário em Access, disponibilizado no site do Centro Interdepartamental de Tradução e Terminologia (CITRAT<sup>110</sup>).

Ainda na busca de equivalentes tradutórios a partir de corpora comparáveis, Tagnin (2008, 2007) expõe outra possibilidade metodológica, com a utilização de ferramentas desenvolvidas para esse fim, evidenciando novos encaminhamentos na viabilização de chegada ao objetivo proposto, conforme segue: **1)** composição de corpus comparável; **2)** utilização do CorTec<sup>111</sup>, *site* com três ferramentas: gerador de listas por ordem de frequência ou alfabética; concordanciador, que apresenta o contexto de uso; gerador de N-gramas, que apresenta combinações de 2, 3 ou 4 palavras; **3)** obtenção de uma lista de frequência do vocabulário de cada corpus; **4)** realização de concordâncias com as palavras pretendidas, na busca de contexto de uso; **5)** confirmação do contexto de utilização por meio da ferramenta de busca GOOGLE e confirmação dos colocados para se chegar a resultados concretos; **6)** utilização do *WordSmith Tools* 3.0 para a ordenações não possíveis no CorTec. **7)** localização de candidatos a equivalentes por meio de lista de frequência; **8)** utilização de concordanciador, caso o item anterior não tiver sido suficiente para estabelecer equivalência; **9)** confecção do produto terminográfico.

A partir da utilização da sequência de etapas descrita anteriormente, Tagnin (2007) pretende enfatizar a possibilidade de se estabelecer equivalentes tradutórios para termos técnicos a partir de corpus comparável bilíngue. Para isso, considera que o processamento

<sup>107</sup> <http://www.intratext.com/>

<sup>108</sup> <http://wse1.webcorp.org.uk/home/>

<sup>109</sup> <http://www.natcorp.ox.ac.uk/>

<sup>110</sup> <http://citratt.fflch.usp.br/>

<sup>111</sup> <http://www.fflch.usp.br/dlm/comet/>

automático dos corpora é práxis eficiente nessa atividade, visto que uma busca leva a outras, enriquecendo o trabalho de buscar equivalentes, na tentativa de auxiliar tradutores por meio da confecção de glossários ou dicionários bilíngues.

É possível entender que existe, nos procedimentos metodológicos mencionados anteriormente, de um lado, a preocupação de se ligar a **definição** à sua **designação**, movimento que inventaria as unidades que podem assumir a posição de termos, acompanhadas de seu possível **conceito**. Terminólogos fazem o referido alinhamento, visando a atualizar o consulente a respeito do termo em questão, não importando a frequência com que essas unidades aparecem no corpus, mas sim o seu status diante da área a que supostamente pertence. Por outro lado, a Linguística de Corpus pode evidenciar justamente as unidades que, mesmo não consideradas terminológicas, são frequentes e de igual importância ao tradutor. Concordanciadores, por exemplo, podem iluminar unidades inicialmente consideradas apenas lexicais, mas que combinadas podem revelar uma vertente especializada.

Diante das considerações discutidas até este momento, pretendo implementar um percurso, um passo-a-passo que traga à tona as unidades terminológicas, mediante a utilização dos postulados da Terminologia, associados aos da Linguística de Corpus, a partir de um corpus comparável de certidões originais brasileiras e americanas de nascimento, casamento e óbito, e traduções juramentadas desse mesmo tipo de certidões, no par linguístico português/inglês, realizadas bidireccionalmente por tradutores brasileiros, no intuito de prover informações para acomodar o texto na língua de chegada, a partir da escolha de equivalentes que possam ser considerados adequados para esse fim. Levando-se em conta a variação de tempo de exercício profissional entre os tradutores, é possível entender que existe oscilação de quantidade de conhecimentos acumulados. Portanto, uma obra terminográfica mais informativa, que contemple os elementos discutidos nos parágrafos anteriores, certamente atingirá maior quantidade de profissionais.

### 3.1. O PERCURSO NESTE ESTUDO

Diante das muitas searas a partir das quais um corpus possa ser coletado para estudos, para os pesquisadores interessados em fenômenos tradutórios, a atitude inicial possivelmente mais recomendável para a realização dessa coleta esteja na escolha de um dos vários âmbitos preconizados pelos Estudos da Tradução (Williams e Chesterman, 2002). Com isso, nesta tese, o interesse está focalizado nas certidões e traduções juramentadas de certidões, situados na área dos textos técnicos, terminologia e glossários. A partir dos relatos dos tradutores juramentados

participantes deste estudo, especificamente as de nascimento, casamento e óbito ocupam lugar de destaque por terem alto índice de recorrência no desempenho de seu trabalho, fator que justifica minha escolha e motiva minha pesquisa. Delimitados a área de estudo, a interface e os tipos de textos, tem início um percurso envolvido em critérios, descobertas e tomadas de decisão, o qual passo a descrever a seguir.

### 3.1.1. A Seleção

A tarefa de coleta de textos para a composição de um corpus não ocorre de maneira aleatória, mas sim, obedece a critérios de seleção preconizados pelos estudiosos da área em questão. Os utilizados neste estudo levam em conta os postulados por Hunston (2010, p. 14-16), Walter (2010, p. 430), Llamazares (2008, p. 342), Fernandes (2006, 87), Baker (1995, p. 229), conforme o **Quadro 3**, sistematizados a seguir.

**Quadro 3 - Critérios e Tipologias Escolhidas.**

<b>Critérios</b>	<b>Tipologia Correspondente</b>	<b>Utilizado neste estudo</b>
1. Alinhamento	Paralelo ou Comparável?	Comparável
2. Objetivo	Linguístico ou Tradutório?	Tradutório
3. Domínio	Geral ou Especializado?	Especializado, restrito
4. Modo, fonte	Escrito ou Falado?	Escrito
5. Tempo	Sincrônico ou Diacrônico?	Sincrônico
6. Limite geográfico	Línguas envolvidas?	Português/Brasil e Inglês/EUA
7. Autores dos textos	Nativos ou não nativos?	Nativos e não nativos.
8. Nº de Línguas	Mono ou Bi ou Multilíngue?	Bilíngue
9. Direção	Uni ou Bi ou Multidirecional?	Bidirecional
10. Limites	Aberto, Fechado ou Monitor?	Fechado
11. Tamanho	De Referência ou Textuais?	Textuais
12. Tratamento	Simple ou Codificado?	Simple

### 3.1.2. A Busca

De acordo com Kenning (2010, p. 489), o processo de compilação de corpora pode ser dividido em três etapas: *design*, *coleta* e *preparação*, e tudo deve estar de acordo com o

objetivo da pesquisa implementada e ser direcionado ao formato aceito pelo *software* selecionado para o processamento dos dados. Depois de estabelecidos os critérios a respeito dos textos a serem coletados, é necessário estabelecer como será o *design* e, em seguida dar início às buscas. Mesmo tendo a *web* como fonte, ainda segundo a autora, o fundamental a ser observado na composição de corpus comparável é a autenticidade, fator que atende ao quesito da representatividade dos textos, especialmente os técnicos. Durante a **coleta** de textos, é possível investigar questões a respeito de autoria e confidencialidade, visto que pode haver a necessidade de respeito tanto ao primeiro quanto ao segundo, protegendo autores, clientes e tradutores. Tudo precisa estar dentro de padrões de legalidade. Por último, ocorre a **preparação** do corpus, o que envolve atender o que foi detectado na etapa anterior sobre o que deve ser retirado como, por exemplo tabelas, imagens, menção de páginas da *web*, e também sobre decidir se o texto será etiquetado ou não. De acordo com Kübler e Aston (2010, p. 508), ao avaliar documentos para verificar a possibilidade de incluí-los, é preciso estar certo de que eles estão alinhados com o desenho almejado, o quanto têm autoridade para serem candidatos, assim como ter em mente o receptor do estudo, o leitor-alvo, no caso o tradutor juramentado brasileiro.

### 3.1.3. O Tratamento

Realizadas as coletas para a formação do corpus, haverá vários arquivos e pastas salvos em computador, ou seja, um corpus que pode ser chamado, segundo Baker (1993, 1995), de “*raw text*”, isto é, de texto cru, designação atribuída ao corpus sem nenhum tipo de tratamento ou, etiquetagem ou processamento. A partir daí, há a necessidade de tratamento dos textos adquiridos: **1)** para que sejam alinhados em uma extensão capaz de ser aceita pelo programa que vai realizar o processamento de dados, visto que, ao abrir as pastas em que os documentos estão guardados, o pesquisador pode encontrar diversos formatos de extensão de arquivos, principalmente os não editáveis, como é o caso dos JPG, PDF ou HTML, presentes em fotos, documentos e sites. **2)** Para descartar a presença de elementos que não vão ao encontro dos objetivos propostos no estudo, e para atender a âmbitos legais, respeitando direitos.

### 3.1.4. A Descrição

Interessa ao pesquisador entrar em contato com o conteúdo dos textos, verificar como elementos utilizados nos de partida se configuram nos de chegada, bem como as estratégias tradutórias empreendidas no sentido de acomodar informações transpostas de forma que

funcionem adequadamente no aspecto cultural. A esse respeito, Kenning (2010, p. 493) enfatiza que existem formas de estabelecer e avaliar comparabilidade no corpus e menciona que Tognini-Bonelli (2001, p. 132) propõe uma metodologia em três passos para isso: **1)** análise dos padrões do item L1, seguida pela identificação do significado/função de cada padrão; **2)** identificação de um equivalente L2 *prima facie* para cada significado/função (usando a própria experiência, obras de referência e/ou corpora paralelos, quando disponíveis); **3)** análise das realizações formais de cada significado/função em L2. Depois de analisar a relação tradutória entre os textos, o estudo pode continuar em direção a verificar padrões do texto traduzido, o que, ainda segundo Kenning (2010, p. 494) são resultantes das restrições do processo de traduzir.

Ainda a respeito de descrições dos textos, Kübler e Aston (2010, p. 509) explicam que critérios utilizados de maneira isonômica podem oferecer ao tradutor um meio de familiarização com o domínio e gênero em ambas as línguas, de origem e alvo, proporcionando formas de identificar diferenças e semelhanças interculturais.

### **3.1.5. A Ferramenta**

Existem ferramentas que auxiliam o pesquisador na tarefa de processar os corpora e cumprir o objetivo que se tem em mente. Algumas estão disponíveis gratuitamente na internet, como é o caso do AntConc<sup>112</sup>, outros estão atrelados ao corpus a que se referem, como é o caso dos grandes corpora de linguagem geral, como o BNC<sup>113</sup> e o COCA<sup>114</sup>, e também é possível adquirir um software mediante pagamento de licença, como o *WordSmith Tools 9.0*<sup>115</sup>. Independentemente da maneira de como obtê-los, Walter (2010, p. 432) preconiza a existência de requisitos que devem constar nas ferramentas de análise de corpora e destaca as mais importantes: **1)** Lematização: apresenta as inflexões de certa palavra-base, evidenciando tipo e similaridade; **2)** Marcação: marca automaticamente as classes gramaticais; **3)** Sensibilidade: busca ou descarta iniciais maiúsculas; **4)** Seleção: seleciona linhas que apresentam texto incomum; **5)** Busca unidades múltiplas palavras; **6)** Expansão do contexto, expande a linha verificada; **7)** Separa por ordem alfabética e exclui ocorrências inexpressivas; **8)** Filtra: busca por característica específica.

<sup>112</sup> <https://www.laurenceanthony.net/software/antconc/>

<sup>113</sup> <http://www.natcorp.ox.ac.uk/>

<sup>114</sup> <https://www.english-corpora.org/coca/>

<sup>115</sup> <https://www.lexically.net/wordsmith/>

### 3.1.6. As Listas de Frequência e de Concordâncias

Como afirma Para Hunston (2002, p. 3), o software que analisa corpora é capaz de evidenciar uma nova perspectiva sobre o familiar por meio da submissão dos textos a dois processos interligados: **1)** a produção de listas de **frequência**, em ordem de classificação ou alfabética e **2)** a geração de **concordâncias**, que apresentam exemplos de itens específicos em contexto. De acordo com Evison (2010, p. 122), o empreendimento não acaba na simples geração dessas listas, e esclarece que funcionam como ponto de partida para se chegar a descobertas tanto quantitativas quanto qualitativas que podem levar a novos encaminhamentos. A lista de frequência é uma ferramenta automática e agrupa todos os itens de um corpus, mostrando a porcentagem que oscila entre os mais comuns até os utilizados apenas uma vez, destacando, com isso, a quantidade de *tokens* e de *types*. Está interligada à análise de concordância, visto que pode mostrar os itens a partir dos quais **um** possa ser eleito para ser o investigado pelo pesquisador. De acordo com Tribble (2010, p. 175) o item-alvo, também chamado de *node*, palavra-nó, ou nódulo, aparecerá centralizado na tela, ao mesmo tempo em que evidencia as ocorrências tanto à sua direita quanto à sua esquerda, o que pode suscitar mais descobertas e novas buscas.

### 3.1.7. A Leitura da Concordância

De acordo com o mencionado anteriormente, somente gerar listas de frequência e de concordância pode ser considerado pouco frutífero. É preciso entender como funcionam e como podem contribuir para desvendar nuances do corpus, levando o pesquisador a descobertas. A esse respeito, Tribble (2010, p. 178-79; Olohan, 2004, p. 74) esclarecem que os postulados de Sinclair (2003: xvi–xvii) podem funcionar como norteadores eficientes na realização de análise dos dados em listas de concordância, a partir de 7 passos, resumidos a seguir: **1)** Iniciar - processo de buscar padrões à direita ou à esquerda da palavra-nó que tenham algum tipo de destaque e que possam valer a pena focar para avaliar sua possível relevância para a análise em questão. Segundo Olohan (2004, p. 82), utilizar a quantidade de até 5 palavras de cada lado já é o suficiente para expandir o co-texto na busca de entendimento do contexto primário. **2)** Interpretar - observar as palavras repetidas e tentar formular uma hipótese que possa vinculá-las à maioria delas. **3)** Consolidar – examinar, com mais distanciamento, a palavra-nó para avaliar se há padrões adicionais ou outras variações no padrão. **4)** Relatar – ao esgotar os

padrões que pode observar e revisar a hipótese para que ela seja tão flexível quanto necessário e tão forte quanto possível, escreva-a para ter uma versão explícita e testável. **5) Reciclar** - consideração rigorosa dos contextos ampliados em que o nó é encontrado. **6) Resultado** - Essas observações podem ser registradas como um foco para estudos futuros e serão incorporadas em um relatório mais completo que contém uma segunda hipótese de trabalho. **7) Repetir** - repetir o processo com mais dados. Isso permite ao pesquisador testar e, em seguida, expandir, refinar ou revisar a hipótese para torná-la tão robusta e útil quanto possível para seus objetivos específicos.

### **3.1.8. A Fraseologia, a Colocação e a Polissemia**

A investigação de palavras de forma individualizada chega a determinados resultados, os quais podem ser ampliados mediante a observação do contexto em que ocorrem, isto é, as vizinhanças com a quais desenvolvem certa interdependência, elucidando padrões fraseológicos que podem incidir sobre significados tanto quanto sobre condições de uso. Moon (2010, p. 199) esclarece que os estudos de corpus se debruçam em descobrir a medida em que as palavras ocorrem como parte de **fraseologias**, ou seja, da maneira com que obedece à **colocação**, mediante a observação do seu entorno. Outro aspecto a ser levado em conta no decorrer da análise fraseológica se refere à **polissemia**, ou seja, aos sentidos e usos e suas quantidades em relação ao significado das palavras e como esses são distinguidos no contexto.

A existência e o comportamento dos itens que acomodam múltiplas palavras são amplos e variados (Greaves e Warren, 2010, p. 201), no entanto, um rastreamento de como circulam pelo corpus aqui descrito será empreendido com vistas a revelar seu funcionamento nesse âmbito. Aqui, verificarei o que o corpus em análise tem a dizer sobre esse assunto.

### **3.1.9. A Composição do Corpus**

Para a composição do corpus deste estudo, optei pelas certidões de nascimento, casamento e óbito por serem, de acordo com tradutores juramentados consultados, as mais requisitadas para serem traduzidas no modo juramentado, fato que despertou a minha atenção, além de, após buscas, não ter encontrado um produto terminográfico que provesse informações relativas a essa seara.

A composição desse corpus teve seu início por meio de pesquisas na internet, utilizando o buscador *Google*. No entanto, a partir do material encontrado, ficou entendido que seria

necessário percorrer outros caminhos que pudessem levar à posse dessas certidões, o que passo a esclarecer, conforme a seguir.

Inicialmente, houve a procura por certidões de nascimento, casamento e óbito na internet. Existem fontes disponíveis que fornecem algumas delas, mas as que apresentam somente modelos são preponderantes, havendo a necessidade de buscas em outras fontes. Para isso, entrei em contato, via *e-mail*, com alunos que frequentavam o curso de especialização em que eu era docente, explicava a minha necessidade e sua finalidade, e pedia que levassem seus documentos para serem digitalizados, no dia de nosso encontro. Tudo foi digitalizado e salvo em uma pasta. Para completar o estoque e chegar a um número satisfatório de certidões no banco de dados, me dirigi a alguns cartórios e solicitei o que precisava. Para isso, como mencionado, basta pagar os emolumentos.

Em seguida, ocorre a busca por certidões de nascimento, casamento e óbito americanas. Diferente do primeiro caso, a internet não é somente a fonte preliminar, mas sim a principal. Como nos Estados Unidos as certidões não são padronizadas como no Brasil, para atender ao critério de autenticidade, precisei verificar como realmente eram. Com isso, entrei em contato com parentes e amigos que moram lá e solicitei que me enviassem modelos que pudessem funcionar como norteador de como seriam, de fato, esses documentos. Depois disso, houve a necessidade de contratar *sites* de genealogia que, por um certo custo, disponibilizam documentos civis de muitas famílias americanas para a construção de árvores genealógicas.

O terceiro movimento foi em direção a compor o subcorpus de traduções juramentadas de certidões de nascimento, casamento e óbito brasileiras em inglês e de certidões americanas do mesmo tipo em português, realizadas por tradutores juramentados brasileiros. Por questões de legalidade, privacidade, ou confidencialidade, esses documentos não estão presentes em *sites* da internet e, por conseguinte, houve a necessidade de entrar em contato diretamente com os tradutores pelos endereços de *e-mail* disponíveis na página da Junta Comercial de cada estado, ou na página da Associação de Tradutores Juramentados, disponíveis na internet.

Com arquivos e pastas salvas em computador, passo a informar como foi a seleção e composição do corpus destinado a prover respostas nesta pesquisa, bem como seus recortes, ou seja, a divisão realizada para dar prosseguimento ao estudo.

É possível afirmar que existem várias maneiras de um pesquisador abordar um corpus na intenção de trazer informações à tona. Nesse âmbito, entra em jogo um percurso cheio de elementos que vão exigir diversas tomadas de decisão, tendo como fio condutor o tipo de textos selecionados, o desenho de como corpus foi compilado, o objetivo que se deseja atingir, e tudo

isso entre outros elementos, sem perder de vista o leitor implícito, que se beneficiará dos resultados desse empreendimento.

Como anteriormente mencionado neste estudo, o corpus em questão é comparável, isto é, não há alinhamento direto entre texto fonte e texto traduzido, mas sim, a presença de documentos que atendem ao mesmo gênero textual, o das certidões, mais especificamente as do registro civil, ou seja, as de nascimento, casamento e óbito brasileiras e americanas, e traduções desse mesmo tipo de textos, realizadas bidirecionalmente no par linguístico português/inglês por tradutores brasileiros.

A partir deste momento, passo a descrever a maneira com a qual são extraídas as informações, tendo em mente o que foi proposto para ser investigado, um relato das descobertas e elucidação do percurso.

### 3.2. PROCEDIMENTOS, DADOS E ANÁLISES

Com o propósito de prover ao tradutor juramentado o acesso a um glossário de fraseologias que povoam o âmbito das certidões civis em questão, é necessário encontrar equivalentes lexicais bidireccionalmente entre línguas. O escopo regional do estudo está nas relações entre Brasil e Estados Unidos, portanto, português brasileiro e inglês americano. Com foco na homogeneidade do corpus, foi determinada a mesma quantidade de unidades para todos os tipos de documentos que comporiam o total. A quantidade de documentos em cada subcorpus foi mantida em 50 unidades devido a motivos extratextuais e um motivo textual: **1)** para compor os subcorpora de certidões de nascimento, casamento e óbito brasileiros, houve custos com emolumentos; **2)** para compor o de certidões americanas, também houve custos, com *sites* de genealogia; **3)** para compor o de traduções juramentadas houve dificuldade para encontrar as de óbito. Quando as de óbito atingiram 50 unidades, esse número funcionou como parâmetro para a composição dos outros subcorpora, finalmente alinhando todos os documentos no quesito “quantidade”. Outro motivo, de ordem textual, é que, ao coletar certidões, **4)** foi possível verificar que o conteúdo se repete em vários momentos no texto, o que pode indicar uma estabilidade na relação *type/token*, isto é, ao aumentar a quantidade de textos não se observa aumento significativo na diversidade vocabular (Baker, 1992; Kübler e Aston, 2010), essa ocorrência é mencionada por McEnery e Wilson (2001, p.148) como *closure*, ou seja, o texto se torna saturado. Segundo os autores, há uma relação diretamente proporcional à especialidade do texto, ou seja, quanto mais especializado, mais probabilidade de ser saturado. A relação *type/token* será retomada mais adiante quando o foco estiver na frequência dos dados.

Diante do exposto, os documentos foram alinhados como demonstra detalhadamente o **Quadro 4** a seguir:

**Quadro 4 - O corpus comparável deste estudo.**

<b>CORPUS COMPARÁVEL</b>			
<b>Documentos em Português</b>		<b>Documentos em Inglês</b>	
50 Certidões Brasileiras de Nascimento	50 Trad. Juramentadas de Certidões Americanas de Nascimento	50 Certidões Americanas de Nascimento	50 Trad. Juramentadas de Certidões Brasileiras de Nascimento
50 Certidões Brasileiras de Casamento	50 Trad. Juramentadas de Certidões Americanas de Casamento	50 Certidões Americanas de Casamento	50 Trad. Juramentadas de Certidões Brasileiras de Casamento
50 Certidões Brasileiras de Óbito	50 Trad. Juramentadas de Certidões Americanas de Óbito	50 Certidões Americanas de Óbito	50 Trad. Juramentadas de Certidões Brasileiras de Óbito
150	150	150	150
300		300	
<b>Total: 600</b>			

O comportamento lexical dos itens que circulam pelos textos não será verificado separando os documentos em categorias unitárias como, por exemplo, somente as certidões de nascimento brasileiras, ou somente as traduções de certidões de casamento americanas, visto que os fatores que regulam essa circulação lexical funcionam da mesma forma pra todas, mesmo que agrupadas e, por outro lado, ao tradutor interessa encontrar no glossário equivalentes tradutórios para dar conta de uma dificuldade, e informar que essa dificuldade está nesse ou naquele documento não parece relevante, já que ele já está debruçado sobre ele. Os textos, então serão analisados por língua e categoria da seguinte forma: gerar listas de frequência dos documentos:

- a. Verificação de 150 certidões brasileiras;
- b. Verificação de 150 traduções juramentadas em português;
- c. Verificação de 150 certidões americanas;
- d. Verificação de 150 traduções juramentadas em inglês.
- e. Análise comparativa entre os itens anteriores.

### 3.2.1. A Utilização do *Software*

Inicialmente, é necessário tornar um arquivo não editável em arquivo editável. Para isso, existem conversores gratuitos na internet, mas há um problema: nem todo tipo de arquivo não editável responde adequadamente ao passar pela conversão. Geralmente, a certidão é fotografada ou digitalizada e, em seguida, colocada em um documento PDF ou em uma página da internet. Parece ser uma extensão dentro de outra, o que inviabiliza o processo de conversão. Diante disso, tudo tem que ser conduzido de forma manual, ou seja, é preciso digitar cada um dos documentos em *word*, que é editável, isto é, 50 arquivos não editáveis tornam-se apenas 1 arquivo editável, já que todos os documentos são digitados um após o outro no mesmo lugar. Isso foi feito com todas as certidões brasileiras, todas as americanas e quase todas as traduções juramentadas, visto que havia algumas já em arquivos editáveis.

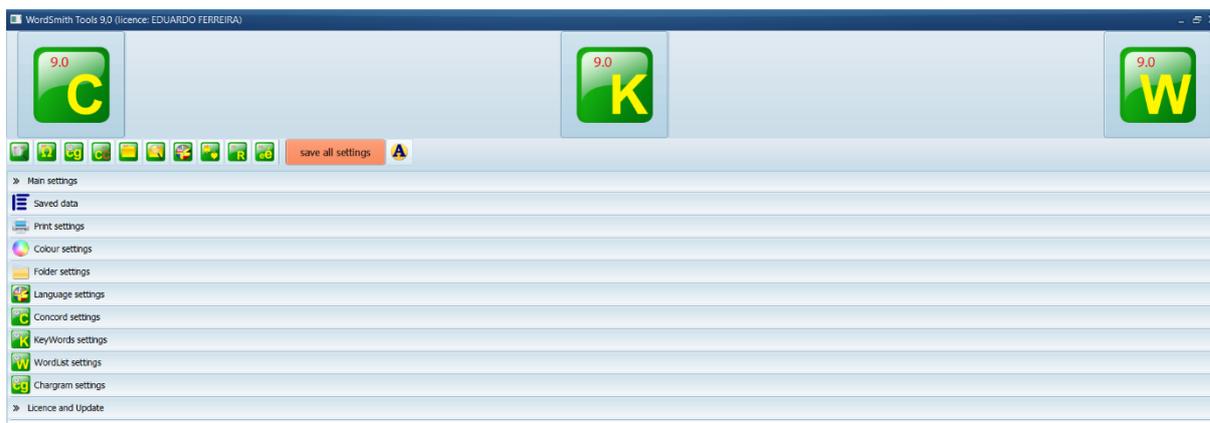
Em seguida, é necessário “limpar” o texto. Retirar elementos que não se alinham com a pesquisa em pauta e retirar todos os que esbarram em questões legais. Com isso, foram deletados **1)** nomes de pessoas: dos nascidos, dos pais, dos avós, dos declarantes, dos noivos, das testemunhas, dos padres, dos pastores, dos falecidos, dos médicos, dos agentes funerários, dos escrivães e dos tradutores; **2)** nomes de localidades: estados, condados, cidades, hospitais e cemitérios; **3)** datas: do nascimento, do casamento, do óbito, do registro, da tradução. Em resumo, todos os nomes, numerais e símbolos gráficos como, por exemplo, parênteses, aspas, colchetes, asteriscos, ou qualquer outro, foram retirados.

O terceiro movimento é o de juntar arquivos de acordo com o objetivo do estudo proposto e convertê-los em uma extensão aceita pelo *software* de processamento, no caso, o *Wordsmith Tools 9.0*. Uma dessas extensões é a *.txt*. Nesta etapa, 3 arquivos com extensão *.docx* tornam-se apenas 1 arquivo *.txt*, movimento que é replicado nos outros, resultando, portanto, em 4 arquivos *.txt*: 2 em português (certidões e traduções) e 2 em inglês (certidões e traduções).

Para que se possa dar encaminhamentos em LC, é necessário contar com apoio tecnológico: *hardware*, *software* e um corpus para ser investigado. Existem vários tipos de computadores, de várias marcas, cores e tamanhos. São leves, silenciosos, com processadores eficientes e com grande espaço para armazenamento que pode ser feito na própria máquina, em periféricos como *pen drives* e *HDs* externos, ou em nuvem, que é um servidor remoto disponível na internet. Além da parte física, é importante ter acesso a um *software* versátil de processamento de dados linguísticos. Basta entrar na internet, acessar o *Google* e buscar um programa que atenda às necessidades. Depois de encontrar o adequado, paga-se a licença de

uso com cartão de crédito e faz-se o download do programa, a menos que a preferência seja por um gratuito, disponível *online*. Com máquina, programa e corpus alinhados, os trabalhos podem ser iniciados. Como, nesta pesquisa, o corpus foi desenhado por mim, considere interessante possuir uma ferramenta de processamento instalada no computador pessoal. Optei por baixar o programa *WordSmith Tools 9.0* (Scott, 2020) porque apresenta todos os recursos necessários e pode ser adquirido rapidamente via rede e possui valor acessível, conforme **Figura 9** a seguir.

**Figura 9 - WordSmith Tools 9.0**



**Fonte:** Scott (2020)

Corpus compilado, tratado e dividido em subgrupos, bem como ferramenta adequada ao processamento de dados instalada, dá-se início ao primeiro passo: verificação do corpus de 150 certidões em português como língua de partida, verificação de 150 traduções em português como língua de chegada, o mesmo movimento com os documentos em língua inglesa, e posterior análise comparativa.

De início, foram geradas **listas de frequência** a partir das quais foi possível verificar algumas informações ligadas a quantidades, ou seja, o *type/token ratio* (TTR), que se refere à relação entre o número total de palavras (*tokens*) e a quantidade de palavras diferentes, sem repetições (*types*). Como essa relação leva em conta o tamanho do corpus que pode ser grande, a divisão é feita de mil em mil e, na base 10, quanto menor for o resultado, considera-se que haja mais repetição de palavras e, por conseguinte, menos variedade lexical (Kenning, 2010; Olohan, 2003). A **Tabela 1** a seguir evidencia essa relação.

**Tabela 1 - Quantidades Lexicais.**

<b>EM PORTUGUÊS</b>	
<b>150 Certidões</b>	<b>150 Traduções</b>
<i>Tokens:</i> 21.493	<i>Tokens:</i> 68.311
<i>Types:</i> 529	<i>Types:</i> 2.870
TTR: 2,46%	TTR: 4,20%
<b>EM INGLÊS</b>	
<b>150 Certidões</b>	<b>150 Traduções</b>
<i>Tokens:</i> 29.629	<i>Tokens:</i> 39.437
<i>Types:</i> 618	<i>Types:</i> 1.679
TTR: 2,09%	TTR: 4,26%

Algumas considerações podem ser pontuadas a partir da observação da tabela anteriormente apresentada. Ao serem comparados os números relativos ao que ocorre dentro de cada subcorpus, observa-se que, diante da maneira com a qual os itens se repetem para organizar os textos, é possível afirmar que a quantidade de itens lexicais elencados (*Types*) frente à maneira com que se efetivam ao longo do texto (*Tokens*), a relação TTR é baixa. Levando-se em conta as línguas de partida, verifica-se que os *types* certidões americanas ocorrem em maior número do que o contido nas brasileiras. Em ambos os casos, a porcentagem de utilização lexical gira em torno de apenas 2%, o que, de acordo com Olohan (2003, 2002) indica ocorrência de alta repetição de itens e, por conseguinte, menos variedade lexical. A similaridade nos dois casos, no entanto, é mais marcante nas brasileiras.

Voltando o olhar para as línguas de chegada, os subcorpora de traduções, observa-se a mesma ocorrência na relação entre *Types* e *Tokens*, ou seja, a porcentagem também é baixa. A possível explicação para o aumento lexical pode estar justamente nos quesitos necessários para a realização do modo juramentado. Nesse âmbito, conforme explica Nascimento (2006), o tradutor inclui elementos pré e pós-textuais, de acordo com o estipulado pela Junta Comercial dos estados, e marcações, explicações, notas no texto e notas de rodapé relativas ao próprio texto submetido à tradução, ou seja, toda sorte de esclarecimentos que tornem as informações aceitáveis ao serem acomodadas na língua de chegada.

Um terceiro movimento pode ser realizado ao alinhar a comparação entre língua de partida e língua de chegada. Nesse caso, observa-se que para realizar a transposição, a quantidade de *types* praticamente dobra, e a quantidade de *tokens* mais que triplica, o que ocorre

pelos motivos já descritos anteriormente, evidenciando a necessidade do tradutor juramentado em apresentar um trabalho realizado de maneira minuciosa.

Além da contabilidade total de itens lexicais, conforme apresentado anteriormente, as listas de frequência, podem mostrar a ordem com que ocorre cada item mostrando o número de vezes em que constam no corpus e a porcentagem que cada um obtém em um parâmetro de 100%. A lista pode ser colocada em ordem alfabética. No entanto, neste estudo, o interesse está nas porcentagens de frequência para localizar fraseologias. Com isso, converter a lista em ordem alfabética é uma manobra que visa a apenas facilitar essa busca e não, por exemplo, para investigar classes gramaticais, conjugação, flexões ou sufixação. As **Figuras 2, 3, 4 e 5** a seguir, ilustram como a ferramenta elenca os itens lexicais, do mais frequente para o menos. Para localizar um item ou obter um panorama, basta seguir a lista ou, para encontrar um item, basta clicar na ordem alfabética e se dirigir até a inicial de interesse.

**Figura 10 - Lista de frequência de certidões brasileiras.**

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
1	DE	2.060	9,51%	1	100,00%	0,76		
2	DO	791	3,65%	1	100,00%	0,77		
3	E	767	3,54%	1	100,00%	0,75		
4	O	583	2,69%	1	100,00%	0,72		
5	REGISTRO	451	2,08%	1	100,00%	0,74		
6	CIVIL	383	1,77%	1	100,00%	0,76		
7	QUE	381	1,76%	1	100,00%	0,75		
8	ESTADO	294	1,36%	1	100,00%	0,71		
9	A	290	1,34%	1	100,00%	0,74		
10	EM	275	1,27%	1	100,00%	0,71		
11	NO	271	1,25%	1	100,00%	0,74		
12	NATURAIS	246	1,14%	1	100,00%	0,76		
13	DA	188	0,87%	1	100,00%	0,74		
14	COMARCA	186	0,86%	1	100,00%	0,71		
15	FEDERATIVA	185	0,85%	1	100,00%	0,73		
16	PESSOAS	182	0,84%	1	100,00%	0,72		
17	CERTIDÃO	170	0,79%	1	100,00%	0,69		
18	NASCIMENTO	170	0,79%	1	100,00%	0,54		
19	#	160	0,74%	1	100,00%	0,76		
20	NESTA	157	0,73%	1	100,00%	0,63		
21	DOU	151	0,70%	1	100,00%	0,70		
22	É	151	0,70%	1	100,00%	0,70		
23	BRASIL	149	0,69%	1	100,00%	0,75		
24	REPÚBLICA	149	0,69%	1	100,00%	0,75		
25	DAS	147	0,68%	1	100,00%	0,74		
26	NÃO	147	0,68%	1	100,00%	0,71		
27	FILHO	142	0,66%	1	100,00%	0,74		
28	CARTÓRIO	139	0,64%	1	100,00%	0,65		
29	COM	137	0,63%	1	100,00%	0,68		
30	NACIONALIDADE	137	0,63%	1	100,00%	0,57		
31	RESIDENTE	132	0,61%	1	100,00%	0,64		
32	NADA	126	0,58%	1	100,00%	0,67		
33	ANOTAÇÕES	125	0,58%	1	100,00%	0,55		

**Figura 11 - Lista de frequência de certidões americanas.**

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
1	OF	2.389	8,02%	1	100,00%	0,74		
2	THE	1.295	4,35%	1	100,00%	0,74		
3	NAME	659	2,21%	1	100,00%	0,75		
4	AND	627	2,11%	1	100,00%	0,73		
5	OR	553	1,86%	1	100,00%	0,74		
6	TO	501	1,68%	1	100,00%	0,74		
7	DATE	476	1,60%	1	100,00%	0,73		
8	CITY	450	1,51%	1	100,00%	0,69		
9	BIRTH	443	1,49%	1	100,00%	0,64		
10	IF	416	1,40%	1	100,00%	0,74		
11	NUMBER	397	1,33%	1	100,00%	0,73		
12	IN	383	1,29%	1	100,00%	0,74		
13	DEATH	343	1,15%	1	100,00%	0,62		
14	ON	336	1,13%	1	100,00%	0,73		
15	MARRIAGE	331	1,11%	1	100,00%	0,50		
16	STATE	304	1,02%	1	100,00%	0,71		
17	BY	300	1,01%	1	100,00%	0,74		
18	THIS	300	1,01%	1	100,00%	0,64		
19	CERTIFICATE	294	0,99%	1	100,00%	0,71		
20	AT	270	0,91%	1	100,00%	0,75		
21	THAT	248	0,83%	1	100,00%	0,68		
22	GIVE	245	0,82%	1	100,00%	0,62		
23	PLACE	202	0,68%	1	100,00%	0,66		
24	WAS	193	0,65%	1	100,00%	0,71		
25	COUNTY	189	0,63%	1	100,00%	0,66		
26	ADDRESS	186	0,62%	1	100,00%	0,70		
27	IS	181	0,61%	1	100,00%	0,71		
28	FATHER'S	179	0,60%	1	100,00%	0,58		
29	LAST	179	0,60%	1	100,00%	0,74		
30	SIGNATURE	179	0,60%	1	100,00%	0,75		
31	COUNTRY	178	0,60%	1	100,00%	0,74		
32	I	169	0,57%	1	100,00%	0,75		
33	NOT	169	0,57%	1	100,00%	0,68		

**Figura 12 - Lista de frequência de traduções em português.**

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
1	DE	5.301	7,73%	1	100,00%	0,77		
2	DO	2.541	3,71%	1	100,00%	0,77		
3	A	1.463	2,13%	1	100,00%	0,77		
4	EM	1.419	2,07%	1	100,00%	0,76		
5	DA	1.290	1,88%	1	100,00%	0,76		
6	E	1.208	1,76%	1	100,00%	0,77		
7	O	1.206	1,76%	1	100,00%	0,77		
8	ESTADO	1.012	1,48%	1	100,00%	0,76		
9	QUE	965	1,41%	1	100,00%	0,77		
10	OU	918	1,34%	1	100,00%	0,76		
11	NOME	876	1,28%	1	100,00%	0,77		
12	DATA	726	1,06%	1	100,00%	0,76		
13	NASCIMENTO	718	1,05%	1	100,00%	0,68		
14	ASSINATURA	681	0,99%	1	100,00%	0,77		
15	NÃO	598	0,87%	1	100,00%	0,76		
16	PARA	564	0,82%	1	100,00%	0,76		
17	NO	552	0,81%	1	100,00%	0,77		
18	B	487	0,71%	1	100,00%	0,73		
19	CERTIDÃO	487	0,71%	1	100,00%	0,76		
20	LOCAL	476	0,69%	1	100,00%	0,75		
21	REGISTRO	474	0,69%	1	100,00%	0,76		
22	CONDADO	464	0,68%	1	100,00%	0,75		
23	ÓBITO	454	0,66%	1	100,00%	0,56		
24	Nº	425	0,62%	1	100,00%	0,76		
25	DOCUMENTO	406	0,59%	1	100,00%	0,74		
26	CASAMENTO	339	0,49%	1	100,00%	0,54		
27	CÓPIA	338	0,49%	1	100,00%	0,73		
28	OFICIAL	334	0,49%	1	100,00%	0,72		
29	CIDADE	332	0,48%	1	100,00%	0,76		
30	ENDEREÇO	328	0,48%	1	100,00%	0,75		
31	SE	308	0,45%	1	100,00%	0,75		
32	NA	307	0,45%	1	100,00%	0,76		
33	C	305	0,44%	1	100,00%	0,72		

**Figura 13 - Lista de frequência de traduções em inglês.**

N	Word	Freq.	%	Texts	%	Dispersion	Lemmas	Set
1	OF	3.211	8,14%	1	100,00%	0,77		
2	THE	2.368	6,00%	1	100,00%	0,76		
3	AND	1.375	3,49%	1	100,00%	0,77		
4	IN	978	2,48%	1	100,00%	0,76		
5	NO	673	1,71%	1	100,00%	0,77		
6	A	620	1,57%	1	100,00%	0,76		
7	THIS	574	1,46%	1	100,00%	0,76		
8	TO	560	1,42%	1	100,00%	0,76		
9	STATE	491	1,24%	1	100,00%	0,77		
10	ON	486	1,23%	1	100,00%	0,76		
11	BY	388	0,98%	1	100,00%	0,76		
12	STAMP	379	0,96%	1	100,00%	0,73		
13	AT	345	0,87%	1	100,00%	0,73		
14	IS	337	0,85%	1	100,00%	0,77		
15	CERTIFICATE	317	0,80%	1	100,00%	0,75		
16	DEATH	316	0,80%	1	100,00%	0,53		
17	WAS	315	0,80%	1	100,00%	0,74		
18	COMMERCIAL	309	0,78%	1	100,00%	0,71		
19	CIVIL	288	0,73%	1	100,00%	0,73		
20	REGISTRAR	282	0,72%	1	100,00%	0,75		
21	REGISTRY	278	0,70%	1	100,00%	0,77		
22	MY	273	0,69%	1	100,00%	0,77		
23	I	270	0,68%	1	100,00%	0,76		
24	PUBLIC	269	0,68%	1	100,00%	0,74		
25	CERTIFY	260	0,66%	1	100,00%	0,77		
26	TRANSLATOR	241	0,61%	1	100,00%	0,76		
27	TRANSLATION	236	0,60%	1	100,00%	0,76		
28	BIRTH	233	0,59%	1	100,00%	0,56		
29	BRAZIL	232	0,59%	1	100,00%	0,75		
30	BORN	224	0,57%	1	100,00%	0,75		
31	WITNESS	224	0,57%	1	100,00%	0,75		
32	THAT	221	0,56%	1	100,00%	0,76		
33	REPUBLIC	213	0,54%	1	100,00%	0,74		

Conforme pode-se verificar, todas as palavras são colocadas em sequência decrescente, revelando, mais detalhadamente, a quantidade de vezes que aparecem no corpus analisado. Observa-se que a palavra mais frequente é a preposição “de / of”, unanimemente empregada em todos os contextos, indicando que, mesmo existindo em língua inglesa, o caso genitivo, por exemplo, que utiliza apóstrofo seguido de ‘s’, pode não ter sido utilizado nas traduções, fato que pode ser confirmado mediante a observação das três palavras mais frequentes: a sequência é praticamente a mesma em todos os subcorpora. Como visto, isso pode revelar a predisposição do tradutor em manter-se fiel ao texto fonte (Venutti, 1998).

Ainda mantendo o foco no objetivo deste estudo, o caminho a ser seguido deve-se manter na busca por facilitar o trabalho do tradutor e, com isso, aproximar-se do que pode ser o ponto inicial para elencar nódulos que possam funcionar como candidatas a fornecer uma combinação fraseológica. Diante disso, focalizar a investigação, primeiramente nos substantivos pode ser frutífero, já que, para isso, basta seguir a lista de frequência e observar como se comportam. Depois disso, essa mesma atividade pode ser expandida para os verbos, por exemplo, desde que isso seja levado em conta como vantajoso ou não quando da confecção do glossário. Aqui, essa sugestão é apenas para esclarecer que a busca pode ser replicada e

ampliada, bastando haver previsão para isso. Para auxiliar nesse intento, uma lista inicial foi criada constando apenas os substantivos de cada grupo, conforme o **Quadro 5** a seguir:

**Quadro 5 - Lista de frequência de traduções em inglês**

Certidões em Português	Certidões em Inglês
1. Registro	1. Name
2. Estado	2. Date
3. Comarca	3. City
4. Pessoas	4. Birth
5. Certidão	5. Number
6. Nascimento	6. Death
7. Fé	7. Marriage
8. Brasil	8. State
9. República	9. Certificate
10. Cartório	10. Place
Traduções em Português	Traduções em Inglês
1. Estado	1. State
2. Nome	2. Stamp
3. Data	3. Certificate
4. Nascimento	4. Death
5. Assinatura	5. Registrar
6. Certidão	6. Registry
7. Local	7. Translator
8. Registro	8. Birth
9. Condado	9. Brazil
10. Óbito	10. Witness

O **Quadro 5** contempla os substantivos mais utilizados em cada subgrupo, iluminando algumas possíveis comparações. Dos 10 itens mais frequentes no grupo das certidões originariamente escritas em português, 7 continuaram entre as 10 mais frequentes nas traduções juramentadas em inglês. Dos 10 itens mais frequentes do grupo pertencente às certidões americanas, 6 deles também permaneceram com alta presença entre os 10 mais frequentes das traduções juramentadas em português. Por outro lado, dos 10 itens mais frequentes no grupo das certidões originariamente escritas em português, apenas 4 continuaram entre os 10 mais frequentes nas traduções juramentadas também em português. Dos 10 itens mais frequentes do grupo das certidões americanas, também tem frequência baixa, somente 4 deles permaneceram entre os 10 mais frequentes das traduções juramentadas em inglês. O que ocorre pode ser explicado por Baker (1993), ou seja, o texto pode ter sido simplificado para ser prontamente entendido e normalizado por meio de homogeneização textual, o que vai ao encontro dos anseios do leitor-alvo. Por serem os tradutores juramentados deste estudo todos brasileiros, partilham cultura, formação, legislação e ideologias, entre outros aspectos, são pressões que podem incidir no trabalho tradutório (Venutti, 1993). Ao comparar o texto originalmente

em português ao texto traduzido para português, bem como texto originalmente escrito em inglês ao texto transportado para inglês, observa-se que o número de coincidências lexicais entre as 10 mais frequentes é de apenas 4, enquanto que entre idiomas diferentes o número sobe para 6 e 7. Isso pode significar que o texto traduzido tende a se alinhar com os que estão na língua de partida, mas não com os que já circulam na língua de chegada, o que, mais uma vez de acordo com Venutti (1993), evidencia a tendência do tradutor em manter conexões com o texto fonte.

Depois de utilizadas as listas de frequência, o estudo se depara com nódulos capazes de serem fontes gerenciadoras de colocados que levem a fraseologias. Para isso, é necessário visitar os lugares que ocupam no texto e seus arredores, visto que, de acordo com Sinclair (2005), em determinados textos, um nódulo pode ter seu comportamento avaliado de acordo com a presença das palavras no seu entorno. Nesse caso, entra em cena outra ferramenta que realiza o mencionado anteriormente, o concordanciador.

O *Concord* também é uma ferramenta do *WordSmith Tools 9* e, para que faça buscas, é necessário elencar os nódulos para esse fim. Anteriormente neste estudo, na parte que descreve a tipologia das certidões e das traduções juramentadas, foi mencionado que tanto um quanto outro tipo de documento apresenta “fé pública”, algo escrito nas certidões pelo cartorário e pelo tradutor juramentado ao traduzi-as, principalmente como forma de encerramento. Nessa parte do trabalho foi descrito como esse item lexical ocorre nos documentos, contando com a presença de exemplos. Coincidentemente, a palavra “fé” figura entre as 10 palavras mais recorrentes nos textos originalmente escritos em português. Anteriormente, foram elencadas as possíveis maneiras com que pode ocorrer e, neste momento, a intenção de utilizá-la como nódulo visa a constatar como, de fato, é o seu comportamento.

Para abastecer a ferramenta de busca, utilizaremos inicialmente 1 arquivo contendo 150 certidões originalmente em português e, depois disso, 150 traduções juramentadas em português, traduzidas por profissionais juramentados brasileiros, tendo, portanto, o ponto de vista idiomático como parâmetro para tal investigação. O mesmo empreendimento será realizado com documentos em língua inglesa.

Depois de selecionar os arquivos para realizar a análise, o *Concord* lista todas as ocorrências na tela, como na **Figura 14**, a seguir

**Figura 14 - Concordanciador – o item “fé”.**

N	Concordance
1	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 11 República Federativa do Brasil Registro civil
2	civil. O conteúdo da certidão é verdadeiro. Dou fé. 12 República Federativa do Brasil Registro civil
3	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 12 República Federativa do Brasil Registro civil
4	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 13 República Federativa do Brasil Cartório de
5	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 13 República Federativa do Brasil Registro civil
6	Observações: nada consta. O preferido é verdade e dou fé. 13 Registro Civil das Pessoas Naturais Nomes
7	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 14 República Federativa do Brasil Registro civil
8	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 15 República Federativa do Brasil Registro civil
9	anotações ou averbações O referido é verdade e dou fé. 15 República Federativa do Brasil Registro civil
10	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 16 República Federativa do Brasil Registro civil
11	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 16 República Federativa do Brasil Registro civil
12	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 17 República Federativa do Brasil Registro civil
13	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 17 República Federativa do Brasil Estado de
14	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 18 República Federativa do Brasil Estado de
15	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 18 República Federativa do Brasil Registro civil
16	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 19 República Federativa do Brasil Registro civil
17	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 20 República Federativa do Brasil Registro civil
18	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 21 República Federativa do Brasil Registro civil
19	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 22 República Federativa do Brasil Registro civil
20	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 22 República Federativa do Brasil Estado de
21	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 23 República Federativa do Brasil Registro civil
22	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 23 República Federativa do Brasil Estado de
23	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 24 República Federativa do Brasil Registro civil
24	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 24 República Federativa do Brasil Estado de
25	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 25 República Federativa do Brasil Estado de
26	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 25 República Federativa do Brasil Registro civil
27	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 26 República Federativa do Brasil Estado de
28	Endereço O conteúdo da certidão é verdadeiro. Dou fé. 26 República Federativa do Brasil Oficial de
29	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 27 República Federativa do Brasil Estado de
30	anotações ou averbações O referido é verdade e dou fé. 27 República Federativa do Brasil Oficial de
31	anotações ou averbações O referido é verdade e dou fé. 28 República Federativa do Brasil Oficial de
32	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 28 República Federativa do Brasil Estado de
33	naturais. O conteúdo da certidão é verdadeiro. Dou fé. 29 República Federativa do Brasil Estado de

A primeira busca mediante a utilização do Concord com 150 certidões brasileiras, evidencia que a palavra “fé” ocorre em 151 linhas de concordância, portanto, é possível concluir que ocorre apenas 1 vez em cada certidão analisada. A partir deste momento, para prosseguir a análise dos dados dessa concordância, entram em cena os passos norteadores de Sinclair (2003). O primeiro, “iniciar”, é o processo que busca padrões à direita e à esquerda do nóculo. No primeiro caso, à direita, nas certidões, o nóculo é a última ocorrência. Depois dele vêm datas, carimbos e assinaturas. “República Federativa do Brasil”, como mostrado no concordanciador é, na verdade é o início da próxima certidão, lembrando que todas estão no mesmo arquivo. À esquerda ocorre uma frase que auxilia no entendimento do que carrega a fé pública em questão.

As ocorrências nas certidões brasileiras estão, em ordem decrescente de frequência, organizadas da seguinte maneira:

- O conteúdo da certidão é verdadeiro. Dou **fé**.
- O referido é verdade e dou **fé**.
- Nada mais, era o que continha do que dou **fé**.
- Era o que continha em dado assento do qual me reporto e dou **fé**.
- Nada mais. Está conforme o original, dou **fé**.
- Nada mais declarou, dou **fé**.
- Está conforme o original, dou **fé**.

No ambiente das traduções juramentadas, a palavra “fé” ocorre na mesma posição e condições que nas certidões originalmente em português. Ao submeter o arquivo contendo 150 traduções ao *Concord*, houve 153 linhas de concordância, como a seguir, também dispostas a partir da maior frequência no corpus.

- Era o que continha o documento, dou **fé**.
- Nada mais continha o documento, do que dou **fé**.
- Traduzi para o vernáculo, conferi, achei conforme e dou **fé**.
- Era o que continha o referido documento. Dou **fé**.

Esse estudo fraseológico do nóculo envolve, segundo Palumbo (2009, p. 19), Baker, Hardie e McEnery (2006, p. 36, 135) e Olohan (2004, p. 62), algumas categorias. Uma delas é em relação à **colocação**, posicionada à esquerda. Em todas as ocorrências está a palavra “dou” e mais à esquerda, em quase todas as ocorrências a palavra “documento”, ou outra que se refira a documento, e ainda mais à esquerda a palavra “continha”, revelando uma composição formada por “Verbo + Substantivo + Verbo + Substantivo” (O’Dell, McCarthy, 2008, p. 10), ou seja, há a preferência do nóculo em ser acompanhado por elementos que obedeçam a essa estrutura, e juntos figuram apenas em contextos semânticos de finalização do texto. Outro aspecto a ser levado em conta no decorrer da análise fraseológica se refere à **polissemia**. No contexto de utilização da palavra fé, é possível afirmar que, em todas as ocorrências, o sentido estava voltado a reforçar a veracidade do conteúdo abordado, e seu posicionamento atrelado ao final de todos documentos, fato que indica não só uma decisão tradutória, uma escolha, mas também obediência a orientações institucionais. Em português, não foi encontrada em outro contexto de uso.

Com relação aos documentos em direção a inglês, também foi mencionado neste estudo que, para finalizar uma tradução, o tradutor juramentado pode utilizar as seguintes estruturas, de acordo com Asensio (2003):

- In witness whereof I sent my hand.
- In witness hereof I sign this English translation of a \_\_\_\_\_ certificate.
- This is a complete and faithful translation witness my hand and official stamp.
- Such is the tenor of said document, which I hereby certify.
- In my capacity as a sworn translator, I hereby certify that the above text is a true and accurate translation from its original done to the best of my knowledge and ability on this [date].

Ao examinar os exemplos acima, foi possível observar que entre eles há uma palavra que consta entre as 10 mais frequentes do subgrupo das traduções: “witness”. Diante disso, houve o interesse em investigar como esse item se comporta, de fato, no corpus, mediante a utilização do concordanciador. Adianto que, ao verificar o arquivo das certidões originalmente escritas em inglês, nas 33 ocorrências, a palavra “witness” apresenta variação polissêmica, ou seja, a que designa uma pessoa que presenciou um fato social, não a pessoa que finaliza um texto afirmando que seu conteúdo é verdadeiro e o inbuindo de fé pública. O concordanciador encontrou o seguinte sobre as traduções em inglês realizadas por tradutores juramentados brasileiros:

- **Witness** my hand and seal.
- In **witness** whereof.

Diante do resultado que retorna dessa busca, pode-se afirmar que há preferência dos tradutores em finalizarem o trabalho com uma das frases acima, principalmente a primeira, elevando-a ao status de a mais frequente do corpus. Ao contrário do que ocorre com “fé”, “witness” se posiciona na parte inicial da colocação (Palumbo, 2009, p. 19; Baker, Hardie e McEnery, 2006, p. 36, 135; Olohan, 2004, p. 62;), indicando a preferência de seleção lexical a ser utilizada em seguida. O que ocorreu, de fato, incide em uma das categorias do padrão morfossintático fraseológico pré-estabelecido, descritas por Tagnin (2002a, p. 195; 2002b, p. 738; 1999, p. 41-2), isto é, pertence à categoria nominal, ou seja, (Substantivo<sub>base</sub> + Substantivo

colocado), o que indica seu grau de especialização. Segundo Walter (2010, p. 435), no contexto em questão, o especializado, as palavras co-ocorrem com alta frequência, muito maior do que seria por acaso, revelando a força das colocações, a probabilidade de itens ocorrerem juntos (Tribble, 1997).

### 3.3. A CONFECÇÃO DO GLOSSÁRIO

Corpus compilado, tratado, estudado, com fraseologias especializadas extraídas, o próximo passo é construir uma maneira de como organizar um produto terminográfico que atenda o objetivo proposto neste estudo: um caminho para chegar a um glossário de fraseologias que facilite o trabalho do tradutor juramentado, mais uma parte do percurso que passo a tratar.

#### 3.3.1. Macroestrutura

Segundo Wright (1997, p.19) e Cabré (1999a, p. 133), existem alguns aspectos que devem ser levados em consideração quando se tem em mente construir um produto terminográfico: **1)** para quem é destinado – o trabalho aqui desenvolvido se destina ao tradutor juramentado brasileiro; **2)** corpus como fonte de extração – foram utilizadas fontes autênticas; **3)** apresentação que atenda os destinatários – a apresentação não fugirá do que o consulente já está habituado, ou seja, **a)** uma parte de português para inglês, **b)** outra de inglês para português, e **c)** ordem alfabética de entradas; **4)** trabalho ágil, leitura e consulta fáceis – o verbete será identificado rapidamente; **5)** trabalho de natureza descritiva – inclusão de informações que preencham lacunas; **6)** fonte representativa de certo campo temático – vocabulário que circula no âmbito das certidões civis; **7)** coleta de listas apropriadas – utilização de *software* específico para esse fim; **8)** empreender uma pesquisa coerente – o percurso está descrito e fundamentado; **9)** estabelecer relações de equivalência entre duas ou mais línguas – construção de verbetes eficazes.

#### 3.3.2. Microestrutura

Os verbetes de interesse neste estudo serão montados a partir do ambiente especializado, o que não quer dizer que todos os itens lexicais que por ele transitam também sejam de especialidade, visto que, para adquirirem essa denominação precisam ser regidos por algumas orientações. Segundo Greaves e Warren (2010), a literatura terminológica pode evidenciar que

não há unanimidade no que se refere à delimitação de critérios para o reconhecimento de UFE, no entanto, é necessário confiar em algumas orientações para que possam compor a nomenclatura de um glossário especializado: **a)** Aguado de Cea (2007, p. 57), Bevilacqua (2001, p. 98): deve-se levar em conta aspectos pragmáticos – são frequentes e têm significado balizado pela área em que estão, específico dentro do domínio; **b)** Coimbra e Bevilacqua (2008, p. 3), Aguado de Cea (2007, p. 57), Bevilacqua (2005, p. 237, 242): devem ter um termo como colocado, um nóculo acrescido de itens ocorrentes ao mesmo tempo.

Diante do exposto, o verbete a ser implementado terá a seguinte sequência:

- 1) entrada pelo termo, escrito em caixa alta, no singular, e negrito;
- 2) exemplos de utilização em ordem decrescente de frequência;
- 3) exemplos de equivalentes na língua de chegada, em ordem decrescente de frequência;
- 4) NC: (notas e comentários), em caixa alta e negrito, utilizada para esclarecer ambiguidades ou fornecer informações adicionais, quando necessário.

Como visto, uma UFE possui significado único, o que dispensa um código de tratamento de remissivas e de polissemia, mas, se ainda assim, o pesquisador decidir incluir alguma informação a respeito, pode utilizar o item 4) acima.

Anteriormente, houve a utilização de *software* para a busca do item “fê” e, ao ser comparado diante do que se estipula para que seja considerado UFE, a resposta foi afirmativa. Com isso, o candidato foi aceito e o verbete ficará da seguinte maneira, conforme os Quadros 6, 7 e 8, sucedido por equivalentes:

#### Quadro 6 - Microestrutura do glossário em português.

<b>FÉ</b>
Era o que continha o documento, dou <b>fé</b> .
Nada mais continha o documento, do que dou <b>fé</b> .
Traduzi para o vernáculo, conferi, achei conforme e dou <b>fé</b> .
Era o que continha o referido documento. Dou <b>fé</b> .
Witness my hand and seal.
In witness whereof.
<b>NC:</b> Somente ocorre no final do documento.

Do lado em que as entradas ocorrem pelo termo em inglês, basta fazer o contrário, da seguinte forma:

### Quadro 7 - Microestrutura do glossário em inglês.

WITNESS
<b>Witness</b> my hand and seal. In <b>witness</b> whereof.
Era o que continha o documento, dou fé. Nada mais continha o documento, do que dou fé. Traduzi para o vernáculo, conferi, achei conforme e dou fé. Era o que continha o referido documento. Dou fé.
<b>NC:</b> Somente ocorre no final do documento.

Replicando os passos anteriormente descritos, podemos chegar a mais um exemplo, conforme a seguir:

### Quadro 8 - Microestrutura do glossário em português.

CARTÓRIO
<b>Cartório</b> do Registro Civil das Pessoas Naturais <b>Cartório</b> Registral das Pessoas Naturais
Bureau of Vital Records. Office of Vital Records Department of Public Health and Vital Records Civil Registry of Vital Records Judicial Zone of Vital Records
<b>NC:</b> Ocorre principalmente no início do documento.

Depois de se debruçar sobre o corpus e de completar o glossário com todas as possibilidades de UFE, resta apenas a preocupação de como tornar o trabalho ao alcance do consultante, a acessibilidade. Existem possíveis formatos para isso, e para cada um deles, diferentes providências a serem tomadas a depender se impresso, digital ou *online*.

O trajeto percorrido neste estudo, por meio de manipulação de corpora comparável, visando à criação de um glossário, foi capaz de trazer entendimentos sobre alguns aspectos do processo tradutório.

Em procedimentos em que a tradução precisa atender aos anseios dos seus recipientes, parece não haver lugar para liberdade de ação, resolução de desafios tradutórios, ou implementação criativa de soluções o que, paradoxalmente, são atitudes necessárias para que a fluência textual e semelhança à fonte sejam atingidas (Asensio, 2003). Pode-se então dizer que, com base em garantias institucionais e profissionais, confia-se que o tradutor produza “novos” textos totalmente simétricos com suas fontes em todas as intenções e propósitos, tranquilizando os receptores com a manutenção da ideia de cópia, equivalência, coincidência, ou de algo “criado” para ser idêntico como anteriormente explicado por Hermans (2009).

Considerando-se que não é possível haver equiparação ou alinhamento entre línguas e culturas e que lacunas, diferenças e impossibilidades tradutórias serão presenças constantes, haverá sentimento de frustração ao reconhecer que a esperada simetria é algo inexistente. Sobre esse assunto, Hermans (2009, p. 196) postula que “a tradução nunca coincide com sua fonte, não é idêntica ou equivalente em nenhum senso formal ou simples” e, apesar disso, ainda continua a ser vista como um discurso se encaixando a outro, algo puramente ilusório, mas que “[é] parte do que nós, na nossa cultura entendemos por *tradução*<sup>116</sup>” (Tradução minha). Essa pretensão de encaixe, transparência e coincidência pode ser verificada com um pouco mais de nitidez em textos escritos, lugar onde é possível rastrear as tomadas de decisões do tradutor.

Ainda de acordo com Hermans (2009, p. 198), é possível dizer que existem **três** tipos de ocorrências que podem motivar a implantação de ajustes, pressionando o tradutor a tomar providências que visem a mitigar ou resolver tais situações. O **primeiro** diz respeito ao leitor implícito e à preocupação com referências culturais “compartilhadas entre o emissor e o receptor para que sejam capazes de funcionar como veículo para comunicação” (Idem, p. 199). Nessa preocupação em acomodar o texto em um novo ambiente cultural, o tradutor terá que, de certa forma, “manipular” o texto, adicionando ou omitindo informações, para que, ao ser entregue aos novos leitores, não haja sobressaltos no percurso comunicativo.

Verificando o corpus de traduções juramentadas em língua portuguesa, provenientes de certidões originais em língua inglesa, é possível rastrear a intenção do tradutor em ajustar o texto para o leitor brasileiro. Quando se referem ao nome do órgão registrador e emissor de certidões, os Estados americanos o denominam de formas diferentes como, por exemplo, “State Board of Health”, “Department of Health”, “Department of Public Health”, ou “Bureau of Health”. Ao serem traduzidos como “Secretaria de Saúde”, ou “Secretaria de Saúde Pública”, notamos uma tentativa de incluir no texto uma expressão que se assemelhe à língua de partida,

---

<sup>116</sup> The translation never coincides with its source it is not identical or equivalent in any formal or straightforward sense [...]. It is part and parcel of what we, in our culture, have come to understand by *translation*.

ao mesmo tempo em que, supostamente, não cause estranhamento na língua de chegada. Isso pode ocorrer porque a mencionada secretaria de fato existe no Brasil e, apesar de não ser responsável por emitir certidões de nascimento, casamento, ou óbito, deixa a sensação de que não há lacuna tradutória. Manobra análoga pode ser constatada ao se traduzir o nome do departamento americano responsável por emitir certidões: “Division of Vital Records”, “Division of Vital Statistics”, “Bureau of Vital Statistics”. Ao serem acomodados na língua de chegada, são traduzidos como “Registro de Pessoas Naturais”, o nome do órgão equivalente ao americano, aqui no Brasil.

O **segundo** caso é a respeito de momentos em que o texto pode chegar a pontos de intraduzibilidade, ou seja, pode haver componentes significativos na língua fonte, mas que não encontram consonância na língua alvo. Esses becos tradutórios podem entrar em cena a qualquer momento e se transformarem em um problema, que pode ser resolvido de duas maneiras: a) discretamente para que não seja percebido pelo segundo leitor, ou b) por meio de uma intervenção explícita, abrindo colchete, incluindo a informação da maneira e em quantidade que achar adequadas, e fechando o colchete, ou colocando tudo em notas de rodapé. No caso em questão, é possível dizer que existem no corpus deste estudo, várias intervenções do tradutor como, por exemplo: **1)** com utilização de colchetes, da seguinte forma: [selo em relevo], [carimbo], [assinatura], [cancelado], [ilegível] [verso], [código de barras], etc., ou **2)** por meio de notas de rodapé: “N.T. Os itens deixados em branco no documento original não foram traduzidos por serem irrelevantes.” Ou “<sup>1</sup> O erro aparece no original.” Nesses casos, há a quebra da homogeneidade do texto por meio da inserção paratextual e um novo texto é implantado de maneira quase que intrusa diante do leitor implícito.

Por fim, o **terceiro**, que envolve casos de “sobredeterminação contextual”, que se refere a certas “características ou declarações que criam uma lacuna de credibilidade, que somente são resolvidas ao lembrar aos leitores de que se trata de uma tradução” (Hermans, 2009, p. 194). Esse caso também oferece exemplos retirados do corpus analisado neste estudo. Diferente do que ocorreu no primeiro caso, ao se deparar com “Division of Vital Records”, “Division of Vital Statistics”, “Bureau of Vital Statistics” o tradutor pode optar por “Setor de Estatísticas de Saúde Pública”, localizando o texto traduzido próximo à língua de partida. O mesmo ocorre com a palavra “county”, frequentemente traduzida como “condado<sup>117</sup>”, em vez de “comarca”, solução que atenderia o primeiro caso mencionado anteriormente.

---

<sup>117</sup> Esclarecemos que a noção de “**condado**” pode ser assemelhada à noção de “comarca”, visto que ambos dizem respeito à subdivisão do estado em regiões, para melhor fluência de fatores políticos, jurídicos ou administrativos. Diante disso, a depender da demanda, um condado pode conter diferentes quantidades municípios. De acordo com

## CONCLUSÃO

Ao entrar na seara dos Estudos da Tradução, uma nova perspectiva a respeito do processo tradutório começa a ser desenvolvida mediante a abertura de entradas para possíveis estudos e pesquisas, abordagens e métodos, revelando a amplitude de uma área à espera de contribuições. Na busca por uma das subáreas, encontrei os estudos com base em corpus, uma atividade calcada em investigações de dados empíricos, autênticos, ainda em construção, motivos pelos quais imediatamente a ela aderi. Nesse viés, houve apenas a necessidade de estabelecer uma interface teórico-metodológica a partir da qual os estudos pudessem ser iniciados. Nesse contexto, nasce um projeto de pesquisa, desenvolvido ao longo da realização desta tese, percurso que passo a revisar a partir de agora.

Este estudo está localizado na interface entre os Estudos da Tradução, Terminologia e Linguística de Corpus e, por meio de um corpus comparável de certidões de nascimento, casamento e óbito brasileiras e americanas, e traduções juramentadas desses mesmos tipos de certidões, no par linguístico português/inglês, realizadas por tradutores juramentados brasileiros, buscou subsídios para a implementação de um percurso metodológico que levassem à confecção de um glossário técnico de Unidades Fraseológicas Especializadas. Tudo isso com vistas a contribuir com o tradutor juramentado brasileiro no momento em que encontra dificuldades tradutórias nesse contexto, além de também contribuir com estudiosos da área que se interessem por investigações em tradução técnica.

A primeira seção, contou com três grandes partes que se aprofundaram em assuntos relacionados às certidões de nascimento, casamento e óbito produzidas no Brasil, ao mesmo tipo de certidões produzidas nos Estados Unidos, e às Traduções Juramentadas de certidões também do mesmo tipo, realizadas por tradutores regularmente credenciados nas Juntas Comerciais do estado em que exercem sua atividade profissional, fatores que atendem a critérios de autoridade e autenticidade do material compilado. Foram discutidas as bases de constituição dos documentos, instâncias notariais onde estão depositados os registros a que se referem, além de sua utilidade na obtenção de dados estatísticos institucionais e de direitos civis. O mesmo parâmetro foi utilizado para abordar os assuntos a respeito das certidões americanas, evidenciando em que se assemelham, ou não, às brasileiras. Também foram discutidos alguns desafios ao tradutor juramentado brasileiro na atividade de acomodar um

---

a Associação Nacional de Condados (<http://www.naco.org/Counties/learn/Pages/Overview.aspx>), há 3.069 (três mil e sessenta e nove) condados nos Estados Unidos, variando em tamanho e população. Os estados que menos os têm são Havaí e Delaware, com 3 (três) cada, e o que mais tem é o Texas, com 254 (duzentos e cinquenta e quatro).

texto desse tipo na língua de chegada. Foi exposto o caminho para que a um tradutor seja permitido exercer sua atividade no modo juramentado, dificuldades nesse âmbito e possíveis soluções.

A seção 2 abarcou a fundamentação teórica do estudo em pauta, ou seja, aspectos da Terminologia e da Linguística de Corpus em ambiente de Tradução. Houve um passo-a-passo no caminho teórico Terminológico, ou seja, teve início no termo, passando pela fraseologia e pela fraseologia especializada, chegando à Terminografia, um percurso alinhado aos preceitos da Linguística de Corpus que, por meio de ferramenta computacional, chegou a considerações quantitativas e qualitativas sobre dados concretos.

A terceira seção foi o espaço destinado à metodologia utilizada para realização de análises, tomadas de decisão e finalização do percurso estabelecido: chegar ao verbete de um possível glossário especializado com fraseologias de certidões civis, considerado eficaz para atender as dificuldades do tradutor em sua tarefa. Foram visitados encaminhamentos metodológicos de estudiosos das áreas em questão, o que funcionou para orientar as tomadas de decisão na pavimentação de um caminho desenhado para atender as ocorrências desse tipo de texto o que, de fato, foi mostrado detalhadamente. Ficaram claras as maneiras de como os textos foram buscados, selecionados, tratados, assim como os elementos relativos à ferramenta e sua utilização, por meio de listas de frequência e de concordância. Já na parte que se dirige à finalização, tudo que foi apreendido no decorrer do estudo contribuiu para a chegada ao produto terminográfico. Ainda aqui, foram informadas as restrições de cada etapa na construção, processamento, e disponibilização de produto terminográfico. Foram barreiras que envolveram, por exemplo, **1)** questões de privacidade e confidencialidade, aspectos que incidiram diretamente no acesso a alguns documentos e, conseqüentemente, na utilização dos conteúdos expressos nos textos, tendo em mente a legislação; **2)** questões monetárias, com aquisição de documentos, acesso a *sites* particulares e licenças de uso de genealogia e de *software*, correios, com valores que, dependendo do caso, podem inviabilizar o processo; **3)** questões de alcance das ferramentas utilizadas que, apesar de rápidas, não são capazes de fornecer todas as respostas, o que pode depender de certa habilidade investigativa do pesquisador para suprir essa lacuna; e **4)** questões que envolvem a veiculação do produto terminográfico, ou seja, que de acordo com formato da disponibilização para acesso do público, diferentes são as condições a serem adotadas, visto que pode-se utilizar o papel, o aplicativo, uma página na internet.

Foi também discutido como chegar à reta final do percurso investigativo: a proposição de macro e microestrutura de um possível glossário técnico. A respeito da primeira, foi decidido que a melhor forma de apresentação é aquela com que o consulente já esteja acostumado, visto

que pode empreender consultas de maneira intuitiva. A maneira considerada ideal para atender a esse critério foi a ordem alfabética. Não ficou estabelecido em que posição, ou direção, as línguas deveriam ocorrer na macroestrutura do glossário, se primeiro de português para inglês, ou o contrário, por serem apenas dois idiomas envolvidos. Em um glossário multilíngue, que pode ser futuramente desenvolvido a partir das etapas elencadas aqui, esse quesito entraria em análise. Já com relação à microestrutura, a construção do verbete, obedeceu ao que estudiosos manifestaram necessário incluir, a ideia de acesso rápido e fácil a soluções.

Ao visitar as referências bibliográficas da área da terminografia, ficou evidente que a área pode contar com incríveis avanços metodológicos ao se unir à Linguística de Corpus, que além de compilação e estocagem de corpora de toda sorte e tamanho, e investigá-los com parâmetros estatísticos com precisão e rapidez. Nos casos mencionados, para abastecer produtos lexicográficos e terminográficos, pela facilidade em alinhar os textos em nível frasal, entraram em cena os corpora paralelos, ou seja, textos originados em determinada língua fonte e suas respectivas traduções em língua de chegada. Neste estudo, optei pelo nível genérico, em que não havia a relação “homem a homem” entre os elementos dos textos fonte e de chegada e, portanto, depositando o crédito em um corpus comparável. No entanto, ainda que fugindo do procedimento tradicional para extração de termos, o entendimento foi o de que as respostas continuavam lá, mas precisavam de uma espécie de mapa, um percurso que chegasse até elas. Tendo isso em mente, o Objetivo Geral deste estudo foi o de evidenciar um percurso, detalhadamente descrito, para a extração de fraseologias, “alinhando”, de certa forma, certidões civis brasileira e americanas e traduções juramentadas no par linguístico português/inglês, produzidas por tradutores brasileiros, que pudesse funcionar como modelo para tradutores e pesquisadores.

Para cumprir o Objetivo Geral, foram propostos alguns Objetivos Específicos, que passo a considerar.

O primeiro é justamente a respeito de utilização de corpus comparável para a extração de itens candidatos a termos em um glossário. Como visto, a preferência é a de utilização dos paralelos, pelo alinhamento preciso que pode ser realizado, já que o texto na língua de chegada é o produto direto daquele da língua de partida, em nível lexical, com exemplificação em ambiente frasal. Basta colocá-los lado a lado e as ocorrências aparecem. No comparável, o alinhamento é ampliado para nível textual, abarcando elementos semânticos e pragmáticos, como estilo, registro, preferências, posicionamentos, colocação, nível de especialização, objetivo social de utilização, entre outros. Como mencionado, estão lá, mas a forma de encontrá-los é a partir de alinhamentos de outras naturezas.

O segundo, relacionado ao passo-a-passo foi especialmente levado em conta aqui. Na seção destinada à metodologia, foi colocado um roteiro detalhado para responder a esse questionamento. O passo que considerei o mais desafiador foi o de selecionar nódulos candidatos para a composição de fraseologias por meio de listas. De fato, estamos mais acostumados a alinhamentos mais diretos, conforme descrito anteriormente, e transitar entre frequências, concordâncias e textos em diferentes combinações de corpora em análise é, inicialmente, meio confuso. No segundo momento, o pesquisador também ficou no modo automático. No entanto, tudo isso contribuiu para a identificação dos procedimentos básicos para extrações nesse tipo de corpus.

A seguir, sobre a macro e microestrutura, considerei que a mais adequada não foge aos parâmetros de uso já padronizados e circulantes na comunidade, ou seja, rápido acesso, fácil entendimento, eficaz no provimento de respostas, e em ordem alfabética pelo termo. Existem outras possibilidades, mas ao fugir dos hábitos, a eficácia pode se diluir. Acredito que a proposta de macro e microestrutura deste estudo atendeu a todos os parâmetros a elas imputados.

O último objetivo específico diz respeito ao processo tradutório empreendido pelos tradutores juramentados que contribuíram com este estudo. Um ato civil, como o de nascer, casar-se ou morrer deve ser registrado em um cartório. É uma obrigação determinada por lei. Por ser algo jurídico há regulamentação do que e de que forma incluir dados, para que fiquem padronizados. Ao transferir esses dados do registro para a certidão, há parâmetros legais a seguir, da mesma forma que ocorre no assento depositado. No âmbito da certidão no modo juramentado, há restrições, e determinações de como devem se comportar o tradutor, assim como seu trabalho. São mais regulamentações jurídicas. Levando-se todos esses quesitos em consideração, é possível afirmar que, diante de tantas orientações, abrir mão da liberdade, por meio da criatividade nas traduções, pode parecer uma atitude inseguro, compelindo o tradutor a trabalhar de maneira isenta. Conforme mencionei aqui neste estudo, uma tradução juramentada passa a ideia de cópia do original que reassumiu a autoridade de funcionar na cultura alvo por ter recebido selo, carimbo, assinatura e afirmação de que tudo é verdade, pelo tradutor. A postura aqui nos corpora foi a de realizar a tradução posicionada mais perto da língua fonte, ou seja, a fidelidade ao original. Mesmo que fique “diferente”, a literalidade funciona bem, visto que só revelou se tratar de uma tradução, mas manteve ativo o propósito social do acessado pelo documento.

Sintetizando o resultado da pesquisa aqui desenvolvida ficou perceptível que a mescla de procedimentos propiciada pela interface escolhida no estudo se mostrou eficiente na tarefa de verificar, inventariar, e construir caminhos, verbetes e glossário. Teve todos os seus passos

revelados e analisados de maneira rigorosa, e respaldada por estudiosos e, em alguns momentos, por tradutores juramentados. O trajeto foi complexo, no entanto deixou claro que é possível e viável, deixando um rastro de novos conhecimentos como, por exemplo, as aproximações e alinhamentos dos subcorpora, de métodos, por meio de explicitação detalhada das etapas, de aplicações práticas, ao chegar a um produto terminográfico possível e viável.

Importante mencionar que este estudo é preliminar. Há espaço para que outros possam ser desenvolvidos tendo este como fonte e orientações e provimento de lacunas. Entendo que um glossário mais robusto poderia ter sido desenhado para melhor exemplificar a nomenclatura, no entanto, preferi manter o foco no objetivo proposto que estava em iluminar um caminho e não prover um produto completo, e isso foi cumprido.

A partir daqui, pode haver replicação do estudo em outros âmbitos, replicação do percurso metodológico para outras buscas, ampliação de escopo de interesse, ampliar a quantidade de idiomas nas buscas, enfim, o campo é grande e as necessidades e possibilidades são diretamente proporcionais ao seu tamanho. Diante disso, é possível afirmar que este trabalho contribui para futuras reflexões e discussões teórico-metodológicas, para tornar a interface utilizada ainda mais robusta e relevante como fonte estável de conhecimento, para a compilação de corpora especializados e produção de produtos terminológicos mais afinados e acessíveis aos consulentes.

## REFERÊNCIAS

- AGUADO DE CEA, G. La fraseología en las lenguas de especialidad. **In:** "Las lenguas profesionales y académicas". **Ariel**, Barcelona, 2007, pp. 53-65. ISBN 978-84-344-8122-0. Disponível em: <http://oa.upm.es/6536/> Acesso em: 15 mar. 2022.
- ALMEIDA, G. M. B. A Teoria Comunicativa da Terminologia e a Sua Prática. **Alfa**, São Paulo, 50 (2): 85-101, 2006.
- ALMEIDA, G. M. B. O Percurso da Terminologia: de Atividade Prática à Consolidação de uma Disciplina Autônoma. **TradTerm**, 9, 2003, p. 211-222.
- ALMEIDA, G. M. B.; ALUÍSIO, S. M.; OLIVEIRA, L. H. M. O Método em Terminologia: Revendo alguns procedimentos. **In:** ISQUIERDO, A. N.; ALVES, I. M. **As Ciências do Léxico** – Lexicologia, Lexicografia, Terminologia. Vol. 3. Campo Grande: Editora UFMS, 2007.
- ALMEIDA, G. M. B.; VALE, O. A. Do Texto ao Termo: Interação entre Terminologia, Morfologia e Linguística de Corpus na Extração Semi-Automática de Termos. **In:** ISQUIERDO, A. N.; FINATTO, M. J. B. (Orgs.) **As Ciências do Léxico: Lexicologia, Lexicografia e Terminologia**. 1ª ed. Campo Grande: Editora da UFMS, 2008, v. IV, p 483-499.
- ALUÍSIO, S. M.; ALMEIDA, G. M. B. O que é e como se constrói um *corpus*? Lições aprendidas na compilação de vários *corpora* para pesquisa linguística. **Calidoscópico**, Vol. 4, nº 3, p. 156-178, set/dez 2006. Disponível em: <http://revistas.unisinos.br/index.php/calidoscopio/article/view/6002>. Acesso em: 10 dez. 2022.
- ANTHONY, L. A critical look at software tools in corpus linguistics. **Linguistic Research**, 2013, 30(2), p. 141-161. Disponível em: <http://www.laurenceanthony.net/publications.html#papers>. Acesso em: 10 dez. 2022.
- ASENSIO, R. M. **Translating Official Documents**. Manchester, UK: St. Jerome Publishing, 2003.
- AUBERT, F. H. Dúvidas e Controvérsias. Cadernos de Terminologia nº 5, 2012, pp 5-43. Disponível em: [http://tradterm.vitis.uspnet.usp.br/images/revistas/ct05/03\\_-\\_Dúvidas\\_e\\_Controvérsias.pdf](http://tradterm.vitis.uspnet.usp.br/images/revistas/ct05/03_-_Dúvidas_e_Controvérsias.pdf) Acesso em: 1 out. 2022.
- AUBERT, F. H. Tipologia e procedimentos da tradução juramentada. Vl. 1 e 2, São Paulo: CITRAT/FFLCH/USP, 1998a.
- AUBERT, F. H. e TAGNIN, S. E. O. Um Corpus de Traduções Juramentadas – Material de Pesquisa Linguística, Sociológica e Histórica. **TradTerm**: 10: 163-178. Manchester, UK: St Jerome, 2004. Disponível em: <https://www.stjerome.co.uk/tsa/abstract/6500/> Acesso em: 5 mar. 2022.
- AUBERT, F. H. e TAGNIN, S. E. O. A corpus of sworn translations. **In:** Proceedings of the Corpus Linguistics 2003, Dawn Archer, Paul Rayson, Andrew Wilson e Tony

McEnery (eds.), **UCREL Technical Papers** Vol.16, Part 1, Special Issue, 2003.

BAKER, Mona. Corpus-based Translation Studies: The Challenges that Lie Ahead. **in** SOMERS, H. (ed.) **Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager**. Amsterdam: John Benjamins, pp. 175–86, 1996.

BAKER, M. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. **Target** 7(2), 1995. p. 223-243.

BAKER, Mona. Corpus Linguistics and Translation Studies: Implications and Applications. **in** BAKER, M.; FRANCIS, G.; TOGNINI BONELLI, E. (eds). **Text and Technology: In Honour of John Sinclair**. Amsterdam: John Benjamins, pp. 233–50, 1993.

BAKER, M. **In Other Words: A Coursebook on Translation**. London: Routledge, 1992.

BAKER, P.; HARDIE, A.; McENERY, T. **A Glossary of Corpus Linguistics**. Edinburgh: Edinburgh University Press, 2006.

BARROS, L. A. **Curso Básico de Terminologia**. São Paulo: Editora da Universidade de São Paulo (EDUSP), 2004.

BARROS, L. A.; BABINI, M.; AUBERT, F. H. Terminologia e Tradução Juramentada. Questões de tipologia textual e equivalência terminológica interlinguística Português-francês-italiano. **Filolol. Linguíst. port.**, n 12(2), p. 233-249, 2010. Disponível em: [http://www.fflch.usp.br/dlcv/lport/flp/images/arquivos/FLP12\\_2/Barros\\_etal.pdf](http://www.fflch.usp.br/dlcv/lport/flp/images/arquivos/FLP12_2/Barros_etal.pdf) Acesso em 5 mar. 2022.

BARTSCH, S.; EVERT, S. Exploring the Firthian notion of collocation. **In**: HARDIE, A.; LOVE, R. (Eds.) **Corpus Linguistics 2013 – Abstract Book**. Lancaster: UCREL, 2013. p. 31. Disponível em: <http://ucrel.lancs.ac.uk/cl2013/doc/CL2013-ABSTRACT-BOOK.pdf>. Acesso em: 10 dez. 2022.

BENNETT, G. R. **Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers**. Michigan ELT, University of Michigan Press, 2010. Disponível em: <https://www.press.umich.edu/pdf/9780472033850-part1.pdf>. Acesso em: 10 dez. 2022.

BENSON, M.; BENSON, E.; ILSON, R. **The BBI Combinatory Dictionary of English: a guide to word combinations**. Philadelphia, PA: John Benjamins. Revised edition. 1997. 386 p.

BEVILACQUA, C. R. Unidades Fraseológicas Especializadas: estado da questão em relação a sua definição, denominação e critérios de seleção. **TradTerm**, 11, 2005, p. 237-253

BEVILACQUA, C. R. Fraseologia: perspectiva da linguagem comum e da língua especializada. **Revista Língua e Literatura**, v. 6 e 7, nº 10/11, p. 73-86, 2004/2005. Disponível em

<http://revistas.fw.uri.br/index.php/revistalinguaeliteratura/article/viewFile/40/78> Acesso em: 10 out. 2022.

BEVILACQUA, C. R. Terminologia mono/bi/multilíngüe: algumas propostas e reflexões referentes às unidades fraseológicas especializadas. **TradTerm** 8, 2002. p. 135-147

BEVILACQUA, C. R. Do domínio jurídico-ambiental: proposta de critérios para seleção e tratamento de unidades fraseológicas. **In** KRIEGER, M.G.; MACIEL, A. M. B. (Orgs.) **Temas de Terminologia**. Porto Alegre/São Paulo: Ed. Universidade / UFRGS / Humanitas / USP, 2001. (p. 98 – 105)

BEVILACQUA, C. R. Unidades fraseológicas especializadas: novas perspectivas para sua identificação e tratamento. **In** KRIEGER, M.G.; MACIEL, A. M. B. (Orgs.) **Temas de Terminologia**. Porto Alegre/São Paulo: Ed. Universidade/UFRGS/Humanitas/USP, 2001b. (p. 106 – 117)

BIBER, D. Corpus-Based and Corpus-Driven Analyses of Language Variation and Use. **In**: HEINE, B.; NARROG, H. (Eds.) **The Oxford Handbook of Linguistic Analysis**. Oxford Handbooks Online, 2012. p. 1-34. Disponível em: <http://eclass.uoa.gr/modules/document/file.php/MEDIA139/%CE%91%CE%BD%CE%B1%CE%B3%CE%BD%CF%8E%CF%83%CE%BC%CE%B1%CF%84%CE%B1/Biber,%20Corpus-Based%20and%20Corpus-Driven%20Analyses%20of%20Language%20Variation%20and%20Use.pdf> Acesso em: 15 jan. 2022.

BIBER, D.; CONRAD, S.; REPPEN, R. **Corpus Linguistics: Investigating language structure and use**. Cambridge: Cambridge University Press, 1998.

BIX, B. H. **The Oxford Introductions to U.S. Law - Family Law**. New York: Oxford University Press, 2013. Disponível em: <https://books.google.com.br/books?id=wjVpAgAAQBAJ&pg=PR4&dq=The+Oxford+Introductions+to+U.S.+Law:+Family+Law&hl=pt-BR&sa=X&ei=h37iVPaxI-eIsQS5tILOBQ&ved=0CB4Q6AEwAA#v=onepage&q=The%20Oxford%20Introductions%20to%20U.S.%20Law%3A%20Family%20Law&f=false> Acesso em: 3 jun. 2022.

CABRÉ, M. T. El Principio de Poliedricidad: la Articulación de lo Discursivo, lo Cognitivo y lo Lingüístico en Terminología. **Organon**, v. 25, nº 50, 2011. Disponível em: <http://seer.ufrgs.br/organon/article/view/28343> Acesso em: 1 out. 2023.

CABRÉ, M. T. Theories of terminology: their description, prescription and explanation. **Terminology**, v.9(2), p.163-200, 2003.

CABRÉ, M. T. **Terminology – Theory, methods and applications**. Juan C. Sager (ed.). (Trad.: Janet Ann DeCesaris). Amsterdam/Philadelphia: John Benjamins Publishing Company, 1999.

CABRÉ, M. T. **Terminologia: Representación y Comunicación – elementos para una teoría de base comunicativa y otros artículos**. Barcelona: Univesitat Pompeu Fabra. Sèries Monografies, 3, 1999a.

CABRÉ, M. T. **Terminology – Theory, Methods and Applications**. Volume 1. (Edited by Juan Sager and Translated by Janet Ann DeCesaris). Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998.

CABRÉ, M. T.; FREIXA, J.; LORENTE, M.; TEBÉ, C. La Terminología Hoy: Replanteamiento o deversificación. Porto Alegre: **Organon**, nº 26, 1998.

CAMARGO, D. C. Tradução e tipologia textual. Tradução e Comunicação – **Revista Brasileira de Tradutores**, n. 16, p. 46-52, S. J. do Rio Preto, 2007. Disponível em: <http://sare.unianhanguera.edu.br/index.php/rtcom/article/viewArticle/133>  
Acesso em: 5 mar. 2022.

CAMPBELL, A. S. Tradutores Públicos e Tradutores Juramentados no Brasil. **In:** PORTINHO, W. M. (Org.) **A Tradução Técnica e seus Problemas**. São Paulo: Editora Álamo, 1984 (p. 107-120).

CANDLIN, C. N.; CARTER, R. **Corpus-based Language Studies: An Advanced Resource Book**. London: Routledge, 2005. Disponível em: <http://www.lancaster.ac.uk/fass/projects/corpus/ZJU/xCBLS/chapters/A01> Acesso em: 15 jan. 2023.

CHATMAN, S. **Coming to Terms – The Rhetoric of Narrative in Fiction and Film**. New York: Cornell University Press, 1990. Disponível em: [http://books.google.com.br/books?id=loD1JXOtmTYC&printsec=frontcover&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](http://books.google.com.br/books?id=loD1JXOtmTYC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false) Acesso em: 1 out. 2022.

CHENG, W. **Exploring Corpus Linguistics: Language in Action**. Londres e Nova Iorque: Routledge, 2012.

COIMBRA, S. A. C.; BEVILACQUA, C. R. Subsídios para o fazer tradutório: a problemática das unidades fraseológicas especializadas. **Anais do X Encontro Nacional de Tradutores & IV Encontro Internacional de Tradutores (ABRAPT-UFOP)**. Ouro Preto, 2009. Disponível em: <http://www.ichs.ufop.br/anaisdoentrad/images/stories/PDFs/19Coimbra.pdf> Acesso em: 20 out. 2022.

COIMBRA, S. A. C.; BEVILACQUA, C. R. Avaliação e proposta de critérios para identificação de unidades fraseológicas especializadas: um estudo exploratório. **Anais do CELSUL**, 2008. Disponível em: [http://www.ufrgs.br/termisul/biblioteca/artigos/artigo\\_CELSUL\\_2008\\_COIMBRA\\_BEVILACQUA.pdf](http://www.ufrgs.br/termisul/biblioteca/artigos/artigo_CELSUL_2008_COIMBRA_BEVILACQUA.pdf) Acesso em: 20 out. 2022.

COSTA, Sérgio Roberto. **Dicionário de gêneros textuais**. 3.ed. ver. ampl. Belo Horizonte: Autêntica, 2014, p.43.

DESMET, I. Teoria e Prática da Fraseologia de Especialidade: **Aplicações. Filol. Lingüíst. Port.**, nº 5, p. 27-56, 2002. Disponível em: <http://revistas.usp.br/flp/article/viewFile/59699/62797> Acesso em: 20 out. 2022.

DOLIVER, K.; FIGUERO, D. **Real Work at Home Opportunities**. United States of America: AuthorHouse, 2012.

EVERT, S. Corpora and collocations. **In:** LÜDELING, A.; KYTÖ, M. (Eds.) **Corpus Linguistics: An International Handbook**, Vol. 2. Berlin, New York: Walter de Gruyter, 2009. p. 1212-1248.

EVISON, Jane. What are the basics of analysing a corpus? In O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 122 – 135. London and New York: Routledge – Taylor and Francis Group, 2010.

FERNANDES, L. **Brazilian Practices of Translating Names in Children's Fantasy Literature: A Corpus Based Study**. Advanced Research in English Series. Pós-Graduação em Inglês – UFSC. Florianópolis: UFSC/CCE/DLLE, 2013.

FERNANDES, L. Design Issues in the Compilation of a Parallel Corpus to Investigate Brazilian Practices of Translating Children's Fantasy Literature. **Ilha do Desterro**, Florianópolis, nº 52, p. 151-166, jan. /jun. 2007.

FERNANDES, L. Corpora in Translation Studies: revisiting Baker's typology. **Fragmentos**, nº 30, p. 087/095. Florianópolis, jan./ jun. 2006. Disponível em: <https://periodicos.ufsc.br/index.php/fragmentos/article/viewFile/8217/7690>. Acesso em: 20 jun. 2022.

FENÁNDEZ, P. M. Aproximação à linguística de corpus como metodologia de base empírica. Compilação e anotação do Corpus Paralelo PALOP (português-espanhol) de Narrativa Pós-colonial. **Revista AGÁLIA**, n. 89-90. Compostela: AGAL, 2007. p. 9-80. Disponível em <http://www.agalia.net/estapaanterior/29-8990.html>. Acesso em: 10 dez. 2022.

FINATTO, M. J. B.; EVERS, A.; ALLE, C. M. O. Para Além das Terminologias: Estudos de Convencionalidade em Linguagens Científicas. In: PERNA, C. B. L.; DELGADO, H. O. K.; FINATTO, M. J. B. (Orgs.) **Linguagens Especializadas em Corpora: Modos de Dizer e Interfaces de Pesquisa**. p. 152-182. Porto Alegre: EDIPUCRS, 2010.

FIRTH, J. R. **Papers in Linguistics 1934-1951**. London: Oxford University Press, 1957.

FLACH, M. G. **Responsabilidade Civil do Notário e do Registrador** – De acordo com o novo Código Civil. Inclui legislação e jurisprudência. Porto Alegre: Editora AGE, 2004.

FLOWERDEW, L. Using corpora for writing instructions. In: O'KEEFFE, A.; McCARTHY, M. (Eds.) **The Routledge Handbook of Corpus Linguistics**. London and New York: Routledge, 2010. p. 444-457.

FINKBEINER, W. E.; URSELL, P. C.; DAVIS, R. L.: **Autopsy Pathology** – A Manual and Atlas. 2<sup>nd</sup> ed. Philadelphia: Saunders Elsevier, 2009.

FUNARO, P. **Legislative Services Agency** – Legislative Guide to Marriage Law. University of Iowa, 1991. Disponível em: <https://www.legis.iowa.gov/DOCS/Central/Guides/marriage.pdf> Acesso em: 20 jun. 2022.

GALINSKI, C.; BUDIN, G. Introducción. In: WÜSTER, E.: **Introducción a la Teoría General de la Terminología y a la Lexicografía Terminológica**. Responsable de la edición: M. Teresa Cabré. Barcelona: Institut Universitari de Lingüística Aplicada – Universitat Pompeu Fabra, 1998.

- GARCIA, A. F. Prefácio. In SHEPHERD, T. M. G.; SARDINHA, T. B.; PINTO, M. V. **Caminhos da Linguística de Corpus**. 1. ed. Campinas, SP: Mercado das Letras, 2012. p. 11-13.
- GILQUIN, Gaëtanelle; GRANGER, Sylviane. How can data-driven learning be used in language teaching? In O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 359 – 370. London and New York: Routledge – Taylor and Francis Group, 2010.
- GRANGER, S.; MEUNIER, F.: The many faces of phraseology. In: GRANGER, S.; MEUNIER, F. (Orgs.) **Phraseology – An interdisciplinary perspective**. Amsterdam, Philadelphia: John Benjamins Publishing Company, 2008. (XIX – XXVIII)
- GREAVES, Chris; WARREN Martin. What can a corpus tell us about multi-word units? In O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 212 – 226. London and New York: Routledge – Taylor and Francis Group, 2010
- GREEN, M. A.; BOWIE, M. J. **Essentials of Health Information Management – Principles and Practices**. United States of America: Cengage Learning, 2005.
- GRIES, S. T.: Phraseology and linguistic theory. In: GRANGER, S.; MEUNIER, F. (Orgs): **Phraseology – An interdisciplinary perspective**. Amsterdam, Philadelphia: John Benjamins Publishing Company, 2008. (3 - 26)
- HANKS, P. **Terminology, Phraseology, and Lexicography**. Oslo, EURALEX – European Association for Lexicography, 2010. Disponível em: [http://www.euralex.org/elx\\_proceedings/Euralex2010/122\\_Euralex\\_2010\\_9\\_HANKS\\_Terminology,%20Phraseology,%20and%20Lexicography.pdf](http://www.euralex.org/elx_proceedings/Euralex2010/122_Euralex_2010_9_HANKS_Terminology,%20Phraseology,%20and%20Lexicography.pdf) Acesso em: 3 mar. 2022.
- HERMANS, T. The Translator's Voice in Translated Narrative (1996). In: BAKER, M. (ed.) **Critical Readings in Translation Studies**. pp. 193-212. London and New York: Routledge, 2009.
- HERMANS, T. Paradoxes and aporias in translation and translation studies. In: RICCARDI, A. (ed.) **Translation Studies – Perspectives on an Emerging Discipline**. Cambridge: CUP, 2002.
- HERMANS, T. 1996a. *Translation's Other*. Inaugural Lecture Delivered at University College London on 19 March 1996. Disponível em: [http://www.ucl.ac.uk/dutch/about\\_us/staff/inaugural\\_theo\\_hermans](http://www.ucl.ac.uk/dutch/about_us/staff/inaugural_theo_hermans) Acesso em: 1 out. 2022.
- HOLANDA, Y. R.; CIELDINE ARRUDA, F. E. Contribuições das Ciências do Léxico Especializado aos Estudos dos Direitos Humanos pelas Assessorias Jurídicas Universitárias Populares. **Migulim – Revista Eletrônica do Netlli**. V. 2, nº 2, ago. 2013, p. 03-13.
- HOLMES, J. S. **Translated! Papers on Literary Translation and Translation Studies**. Amsterdam – Atlanta: Editions Rodopi, 1988.
- HOWARTH, P. **Phraseology and Second Language Proficiency**. London: Longman, 1998.

HUNSTON, S. **Corpora in Applied Linguistics**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Dubai, Tokio: CUP, 7<sup>th</sup> printing, 2010.

HUNSTON, Susan. How can a corpus be used to explore patterns? **In:** O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p.152 – 166. London and New York: Routledge – Taylor and Francis Group, 2010.

HUNSTON, S. **Corpus Linguistics**. Birmingham: Elsevier Ltd., 2006. p. 234-248.

Disponível em:

<http://www.hum.uu.nl/medewerkers/a.dimitriadis/courses/data2012/Hunston06-CorpusLinguistics.pdf>. Acesso em: 10 dez. 2022.

HUNSTON, S. **Corpora in Applied Linguistics**. Cambridge: Cambridge University Press, 2002.

JOHANSSON, S. Corpus, lexis, discourse: a tribute to John Sinclair. **In:** HERBST, T.; FAULHABER, S.; UHRIG, P. (Eds.) **The Phraseological View of Language: A tribute to John Sinclair**. Berlin / Boston: Walter de Gruyter & Co., 2011. p. 17-26.

JOHANSSON, S. Some aspects of the development of corpus linguistics in the 1970s and 1980s. **In:** LÜDELING, A.; KYTÖ, M. (Eds.) **Corpus Linguistics: An International Handbook**, Vol. 2. Berlin, New York: Walter de Gruyter, 2009. p. 33-52.

JONES-VASSEY, K. A. Revisions to the North Carolina Birth Certificate and Their Impact on Tracking Maternal and Infant Health Data. **In:** **Statistical Primer** – A Publication of the State Center for Health Statistics, nº 19, June 2012. Disponível em [www.schs.state.nc.us/schs/pdf/Primer19.pdf](http://www.schs.state.nc.us/schs/pdf/Primer19.pdf). Acesso em: 10 maio 2022.

KENNING, Marie-Madeleine. What are parallel and comparable corpora and how can we use them? *In* O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 487 – 500. London and New York: Routledge – Taylor and Francis Group, 2010.

KOESTER, A. Building small specialised corpora. **In:** O'KEEFFE, A.; McCARTHY, M. (Eds.) **The Routledge Handbook of Corpus Linguistics**. London, New York: Routledge, 2010. p. 66-79.

KRIEGER, M. G. Diversificação e Unificação em Terminologia: Fundamentos para a Terminografia do Mercosul. **TradTerm**, 7, 2001, p. 95-109.

KRIEGER, M. G. O termo: questionamentos e configurações. **In:** KRIEGER, M.G.; MACIEL, A. M. B. (Orgs.) **Temas de Terminologia**. Porto Alegre/São Paulo: Ed. Universidade/UFRGS/Humanitas/USP, 2001. (p. 62 – 81)

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à Terminologia** – Teoria e Prática. São Paulo: Contexto, 2004.

KRIEGER, M. G.; MACIEL, A. M. B. (Orgs.) **Temas de Terminologia**. Porto Alegre/São Paulo: Ed. Universidade/UFRGS/Humanitas/USP, 2001.

KRIEGER, M. G.; MACIEL, A. M. B.; FINATTO, M. J. B. Terminografia das Leis do Meio Ambiente: Princípios teórico-metodológicos. **TradTerm**, São Paulo, Humanitas FFLCH/USP, v. 6, p. 143-169, 2000. Disponível em <http://www.revistas.usp.br/tradterm/issue/view/4144> Acesso em: 18 ago. 2022.

KÜBLER, N. Crating a Term Base to Customise an MT System: Reusability of Resources and Tools from the Translator's Point of View. LREC Workshop #8. **Language Resources for Translation Work and Research**. 2002. p. 44-48. Disponível em: [https://www.academia.edu/277837/Comparable\\_Corpora\\_In\\_Translation\\_Research\\_Overview\\_of\\_Recent\\_Analyses\\_Using\\_the\\_Translational\\_English\\_Corpus](https://www.academia.edu/277837/Comparable_Corpora_In_Translation_Research_Overview_of_Recent_Analyses_Using_the_Translational_English_Corpus). Acesso em: 10 dez. 2022.

KÜBLER, Natalie; ASTON, Guy. Using corpora in translation. In O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 501 – 515. London and New York: Routledge – Taylor and Francis Group, 2010.

LARA, M. L. G. Terminologia e Ciência da Informação. In: ALVES, I. M. et al. (Orgs.) **Os Estudos Lexicais em Diferentes Perspectivas**. Vol. II. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, 2010.

LEWIS, M. (Ed.) **Teaching Collocations: Further Developments in the Lexical Approach**. Hove, UK: Language Teaching Publications. 2000. 248 p. Disponível em: <http://pt.scribd.com/doc/242459315/Michael-Lewis-Ed-Teaching-Collocation-Further-Development-in-Lexical-Approach-Thomson-Heinle-2000#scribd> Acesso em: 10 mar. 2022.

LLAMAZARES, M. V. Lingüística con Corpus. E.H. Filología 30, 2008. p. 329-349. Disponível em: <https://dialnet.unirioja.es/descarga/articulo/3332675.pdf>. Acesso em: 10 dez. 2022.

LORETTE, K.; SPILLANE, M. **How to Open & Operate a Financially Successful Notary Business: With Companion CD-ROM**. Ocala, Florida: Atlantic Publishing Group, 2011.

McCARTHY, M. **Explorations in Corpus Linguistics**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: CUP, 2006.

McENERY, T.; HARDIE, A. **Corpus Linguistics: Methods, Theory and Practice**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Tokyo, Mexico City: CUP, 2012.

McENERY, T.; WILSON, A. **Corpus Linguistics: An Introduction**. 2nd ed., Edinburgh University Press, 2001.

McENERY, A. M.; XIAO, R. Z. Parallel and comparable corpora: What are they au to? In: JAMES G.; ANDERMAN, G. (Eds.) **Incorporating Corpora: Translation and the Linguist**. Clevedon, UK: Multilingual Matters. 2007. Disponível em: [http://www.research.lancs.ac.uk/portal/en/publications/parallel-and-comparable-corpora-what-are-they-up-to\(a72e1d6e-550f-49aa-b041-f6ccb2ab4fd9\).html](http://www.research.lancs.ac.uk/portal/en/publications/parallel-and-comparable-corpora-what-are-they-up-to(a72e1d6e-550f-49aa-b041-f6ccb2ab4fd9).html). Acesso em: 10 dez. 2022.

McENERY, T; XIAO, R.; TONO, Y. **Corpus-Based Language Studies: An Advanced Resource Book**. London, New York: Routledge, 2006.

MEYER, C. F. Pre-electronic corpora. In: LÜDELING, A.; KYTÖ, M. (Eds.) **Corpus Linguistics: An International Handbook**, Vol. 2. Berlin, New York: Walter de Gruyter, 2009. p. 1-13.

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

MONTEIRO, L. M. T. A Importância das Colocações no Ensino de Vocabulário de Inglês como Língua Estrangeira. **Revista Escrita**, Número 15, 2012. Disponível em: <http://www.maxwell.vrac.puc-rio.br/20851/20851.PDF> Acesso em: 10 mar. 2022.

MOON, Rosamund. What can a corpus tell us about lexis? In O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 197 – 211. London and New York: Routledge – Taylor and Francis Group, 2010

MORENO, A. A. **Estudio Contrastivo de Unidades Fraseológicas Especializadas (UFE) en un Corpus Comparable Bilingüe de Contratos de Derecho Civil en Lengua Inglesa y Española**. Madrid, Universidad Complutense de Madrid, 2014. [Tese de Doutorado]. Disponível em: <http://eprints.ucm.es/24414/> Acesso em: 20 out. 2022.

NASCIMENTO, L. A. S. **Investigating Norms in the Brazilian Official Translation of Semiotic Items, Culture-Bound Items, and Translator's Paratextual Interventions**. Tese de doutorado em Letras, Língua Inglesa e Linguística Aplicada, Estudos da Tradução. Florianópolis: UFSC, 2006.

NATIONAL CENTER FOR HEALTH STATISTICS. Vital Statistics of The United States, 1987, Vol. III, Marriage and Divorce. DHHS Pub. N° (PHS) 91-1103. Public Health Service, Washington. U.S. Government Printing Office, 1991. Disponível em: <https://books.google.com.br/books?id=WoUJrnGjrK4C&pg=SA4-PA4&dq=us+marriage+certificate+-philippines&hl=pt-BR&sa=X&ei=RFbiVI36A7aJsQTtk4K4BA&ved=0CDEQ6AEwAw#v=onepage&q=us%20marriage%20certificate%20-philippines&f=false> Acesso em: 3 mar. 2022.

NATIONAL RESEARCH COUNCIL. **Vital Statistics: Summary of a Workshop**. Michael J. Siri and Daniel L. Cork, rapporteurs. Committee on Nation Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press, 2009.

O'DELL, F.; McCARTHY, M.: **English Collocations in Use – Advanced**. How words work together for fluent and natural English. Self-study and classroom use. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Dubai, Tokyo, Mexico City: CUP, 2011.

O'DELL, F.; McCARTHY, M. **English Collocations in Use: Advanced**. How words work together for fluent and natural English. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Tokio, Mexico City: CUP, 2008.

O'KEEFFE, A.; McCARTHY, M. Historical perspective: What are corpora and how have they evolved? In: O'KEEFFE, A.; McCARTHY, M. (Eds.) **The Routledge Handbook of Corpus Linguistics**. London, New York: Routledge, 2010. p. 3-13.

O'KEEFFE, A.; McCARTHY, M.; CARTER, R. **From Corpus to Classroom – Language Use and Language Teaching**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo: CUP, 2007.

OLIVEIRA, L. P. Linguística de Corpus: Teoria, Interfaces e Aplicações. **Matraga**, Rio de Janeiro, Vol. 6, nº 24, jan./jun. 2009. p. 48-76. Disponível em: <http://www.pgletras.uerj.br/matraca/matraca24/matraca.html>. Acesso em: 10 dez. 2022.

OLIVEIRA, Ulisses. Meme: gênero e prática de ensino pela sequência didática. **Entretextos**, Londrina, v. 19, n. 2, p. 63–89, 2020. DOI: 10.5433/1519-5392.2019 v19 n2 p63. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/entretextos/article/view/34469> Acesso em: 17 out. 2023.

OLOHAN, M. **Introducing Corpora in Translation Studies**. London, New York: Routledge, 2004.

OLOHAN, M. **Introducing Corpora in Translation Studies**. London: Routledge, 2003.

OLOHAN, M. Leave it out! Using a Comparable Corpus to Investigate Aspects of Explicitation in Translation. Universidade Federal de Santa Catarina. **Cadernos de Tradução**. V. 1, n. 9, 2002. Disponível em: <https://periodicos.ufsc.br/index.php/traducao/article/view/5984/5688> Acesso em: 3 mar. 2023.

ORENHA, A; CAMARGO, D. C. A Extração de Unidades Fraseológicas Especializadas a Partir de Corpora Paralelos e Comparáveis. **The ESpecialist**, Pesquisa em Línguas para Fins Específicos. Descrição, Ensino e Aprendizagem. ISSN 2318-7115. Vol. 30, nº 1, PUC SP, 2009. Disponível em: <http://revistas.pucsp.br/index.php/esp/article/view/6216> Acesso em: 16 ago. 2022.

ORENHA-OTTAIANO, A.; HASMANN, D. S.; CAMARGO, D. C. Análise das Colocações Especializadas na Área de Sensoriamento Remoto sob a Perspectiva da Fraseologia. **SIGNUM: Estud. Ling.**, Londrina, nº 16/2, p. 169-192, dez. 2013. Disponível em: <http://www.uel.br/revistas/uel/index.php/signum/article/viewFile/15578/14019> Acesso em: 20 out. 2022.

O'SULLIVAN, E. Narratology meets Translation Studies, or, The Voice of the Translator in Children's Literature. **ERUDIT. META**. Vol. 48, nº 1-2, pp. 197-207. Montreal, 2003. Disponível em: <http://www.erudit.org/revue/meta/2003/v48/n1-2/index.html> Acesso em: 1 out. 2022.

PALUMBO, G. **Key Terms in Translation Studies**. London, New York: Continuum International Publishing Group, 2009.

PAXECO, S. A. Configurações sintático-semânticas das unidades fraseológicas especializadas: o caso do léxico militar. **Anais do IX Encontro do CELSUL**, Palhoça, SC,

out. 2010. Disponível em:

<http://www.celsul.org.br/Encontros/09/artigos/Sabrina%20Pacheco.pdf> Acesso em: 20 out. 2022.

PECK, G. (Ed.) **The Methodist Quarterly Review**. Vol. XXIX. Third Series, Vol. VII. New York: Lane & Tippett, 1847.

PORTER, P. B.; CROW, A. J. **Online Roots** – How to Discover Your Family’s History and Heritage with the Power of the Internet. Nashville, Tennessee: Rutledge Hill Press, 2003)

POWELL, K. **The Everything Guide to Online Genealogy: A Complete source for Using the Web to trace your Family history**. Avon: FW Publications, 2008.

RAJAGOPALAN, K. A Lingüística de Corpus no tempo e no espaço: visão reflexiva. **In: GERBER, R. M.; VASILÉVSKI, V. (Orgs.) Um percurso para pesquisas com base em corpus**. Florianópolis: Editora da UFSC, 2007. P. 23-44.

RANDALL 5<sup>TH</sup>, A. From ENIAC to Everyone: Talking with J. Presper Eckert. 2006. p. 1-7. Disponível em: <http://www.cs.princeton.edu/~chazelle/courses/BIB/eniac.pdf> Acesso em: 10 dez. 2022.

REESE, N. **Dissecting the Bible**. UK: Transmission, 2011. p. 17-19. Disponível em: [https://www.biblesociety.org.uk/uploads/content/bible\\_in\\_transmission/files/2011\\_summer/BiT\\_Summer\\_2011\\_Rees.pdf](https://www.biblesociety.org.uk/uploads/content/bible_in_transmission/files/2011_summer/BiT_Summer_2011_Rees.pdf). Acesso em: 10 dez. 2022.

REISS, K. **Translation Criticism – The Potentials & Limitations**. Manchester: St. Jerome Publishing, 2000.

REMENCHE, M. L. R. Terminologia – Reconstrução Histórica dos Principais Paradigmas Epistemológicos da Ciência Terminológica. **TradTerm**, 16, 2010, p. 343-364.

REPPEN, R. **Using Corpora in the Language Classroom**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Tokio, Mexico City: CUP, 2010a.

REPPEN, R. Building a corpus: What are the key considerations? **In: O’KEEFFE, A.; McCARTHY, M. (Eds.) The Routledge Handbook of Corpus Linguistics**. London, New York: Routledge, 2010b. p. 31-52.

RIBEIRO, J. O. X. **Direito Notarial e Registral**. Rio de Janeiro: Elsevier Editora, 2008.

ROCHA, T. S. M.: **Lei de Registros Públicos Anotada** – Parte Geral e Registro Civil de Pessoas Naturais. Brasília: Thesaurus, 2007.

RÖMER, U. No corpus linguist is an island: Collaborative and cross-disciplinary work in researching phraseology. **In: HARDIE, A.; LOVE, R. (Eds.) Corpus Linguistics 2013**. Lancaster: UCREL, 2013. p. 03. Disponível em: <http://ucrel.lancs.ac.uk/cl2013/doc/CL2013-ABSTRACT-BOOK.pdf>. Acesso em: 10 dez. 2022.

SABINO, M. A. O Campo Árido dos Fraseologismos. **Revista SIGNÓTICA**. Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Goiás. V. 23, N. 2, 2011. Disponível em: <http://www.revistas.ufg.br/index.php/sig/article/view/15226/10474> (Acesso em: 3 mar. 2022).

SANTIAGO, M. S. **Unidades Fraseológicas Especializadas em Tutoriais de Ambientes Virtuais de Aprendizagem**: Proposta de um Sistema Classificatório com Base na Valência Verbal. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2013. [Tese de Doutorado] Disponível em: <http://www.lume.ufrgs.br/handle/10183/72745> Acesso em: 20out. 2022.

SARDONE, S. B. **Destination Weddings For Dummies**. Indianapolis: Wiley Publishing, 2007. Disponível em: <https://books.google.com.br/books?id=R7K4eVpbt0YC&pg=PA99&dq=us+marriage+certificate+-philippines+-immigration+-kenya&hl=pt-BR&sa=X&ei=uVbiVJCgFOvZsATnyYHwBg&ved=0CGAQ6AEwCTgK#v=onepage&q=us%20marriage%20certificate%20-philippines%20-immigration%20-kenya&f=true> Acesso em: 3 mar. 2022.

SCHIAVI, G. There is Always a Teller in a Tale. **Target**, Volume 8, Number 1, p. 1-21. John Benjamins Publishing Company, 1996. Disponível em: <http://www.ingentaconnect.com/content/jbp/targ/1996/00000008/00000001/art00001> Acesso em: 1 out. 2022.

SCOTT, M. **WordSmith Tools**: Version 9.0. Stroud: Lexical Analysis Software Ltd., 2020.

SCOTT, Mike. **WordSmith Tools 3.0**. Oxford: Oxford University Press, 1999.

SHEPHERD, T. M. G. O Estatuto da Linguística de Corpus: Metodologia ou Área da Linguística? **Matraga**, Rio de Janeiro, Vol. 16, Nº 24, Jan./Jun. 2009. p. 150-172. Disponível em: <http://www.pglettras.uerj.br/matraca/matraca24/matraca.html>. Acesso em: 10 dez. 2022.

SILVA, M. M. A. Aspectos Variacionais em Corpora Comparáveis na Terminologia da Nanociência/Nanotecnologia. In: ALVES, I. M. et al. (Orgs.) **Os Estudos Lexicais em Diferentes Perspectivas**. Vol. II. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, 2010.

SINCLAIR, J. M. Corpus and Text: Basic Principles. In: WYNNE, M. (Ed.) **Developing Linguistic Corpora**: a Guide to Good Practice. Oxford: Oxbow Books, 2005. p. 1-16. Disponível em: <http://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm>. Acesso em: 10 dez. 2022.

SINCLAIR, J. M. **Trust the Text**: Language, corpus and discourse. London and New York: Routledge, 2004.

SINCLAIR, J. M.: **Reading concordances: an introduction**. New York: Pearson/Longman, 2003.

SINCLAIR, J. M. The Search for Units of Meaning. **Textus**, v. 9, n. 1, p. 75106, 1996.

SINCLAIR, J. M. **Corpus, concordance, collocation**. Oxford: OUP, 1991.

SINCLAIR, J. M. **Preliminary Report on the Linguistic Survey of Scotland**. Edinburgh: University of Edinburgh, 1957.

SINCLAIR, J. M.; JONES, S.; DALEY, R. **English Collocation Studies: The OSTI Report**. London, New York: Continuum, 2004.

STAPLES, W. G. (ed.): **Encyclopedia of Privacy** – Volume 1, A-M. United States: Library of Congress Cataloging-in-Publication Data, 2007. Disponível em <https://books.google.com.br/books?id=mE9fo1rw1SkC&pg=PA58&dq=birth+certificate+pdf&hl=pt-BR&sa=X&ei=USfgVKPhJPHsATs14DoDQ&ved=0CCEQ6AEwAA#v=onepage&q=birth%20certificate%20pdf&f=false> Acesso em: 5 mar. 2022.

STARSKY, W. P.; CROSSETT, L. **The Texas Paralegal: Essential Rules, Documents, and Resources**. Delmar: Cengage Learning, 2010.

STUBBS, M. A tribute to John McHardy Sinclair (14 June 1933 – 13 March 2007). In: HERBST, T.; FAULHABER, S.; UHRIG, P. (Eds.) **The Phraseological View of Language: A tribute to John Sinclair**. Berlin / Boston: Walter de Gruyter & Co., 2011. p. 01-16.

STUPIELLO, E. N. A. A Relevância da Pesquisa Baseada em Corpora na Prática de Tradução Juramentada no Brasil: A Variação Semântica em Estudos Sociais. **Revista do GEL**, S. J. do Rio Preto, vol. 5, n. 1, pág. 81 a 94, 2008. Disponível em: [http://www.gel.org.br/revistadogel/volumes/5/RG\\_V5N1\\_05.pdf](http://www.gel.org.br/revistadogel/volumes/5/RG_V5N1_05.pdf) Acesso em: 5 mar. 2022.

TAGNIN, S. E. O. O Projeto COMET – Corpus Multilíngüe para Ensino e Tradução – na Universidade de São Paulo. Universidade Federal de Uberlândia. **Múltiplas Perspectivas em Linguística**, 2008. Disponível em: [http://www.filologia.org.br/ileel/artigos/artigo\\_065.pdf](http://www.filologia.org.br/ileel/artigos/artigo_065.pdf) Acesso em: 5 mar. 2022.

TAGNIN, S. E. O. A identificação de equivalentes tradutórios em corpora comparáveis. In **Anais do I Congresso Internacional da ABRAPUI: Belo Horizonte, 3 a 6 de junho de 2007**. Disponível em [http://www.fflch.usp.br/dlm/comet/Novo/Stella\\_Abrapui%202007\\_artigo.pdf](http://www.fflch.usp.br/dlm/comet/Novo/Stella_Abrapui%202007_artigo.pdf) Acesso em: 5 mar. 2022.

TAGNIN, S. E. O. Os corpora: instrumentos de auto-ajuda para o tradutor. **Cadernos de Tradução**, V. 1, nº 9. Florianópolis: Universidade Federal de Santa Catarina, 191-219, 2002a. Disponível em <https://journal.ufsc.br/index.php/traducao/issue/view/432> Acesso em: 5 mar. 2022.

TAGNIN, S. E. O. The Brazilian Lexicographic Road to Bilingual Verbal Collocations. **EURALEX** – European Association for Lexicography, 2002b. Disponível em: [http://www.euralex.org/elix\\_proceedings/Euralex2002/082\\_2002\\_V2\\_Stella%20Esther%20Ortwiler%20Tagnin\\_The%20Brazilian%20Lexicographic%20Road%20to%20Bilingual%20Verbal%20C.pdf](http://www.euralex.org/elix_proceedings/Euralex2002/082_2002_V2_Stella%20Esther%20Ortwiler%20Tagnin_The%20Brazilian%20Lexicographic%20Road%20to%20Bilingual%20Verbal%20C.pdf) Acesso em: 5 mar. 2022.

TAGNIN, S. E. O. **Convencionalidade e Produção de Texto: um dicionário de colocações verbais inglês/português - português/inglês**. Tese de Livre-Docência. São Paulo: Universidade de São Paulo. 1999.

TAYLOR, C. What is corpus linguistics? What the data says. **ICAME JOURNAL**, Computers in English Linguistics. Nº 32, April, 2008. p. 179-200. Disponível em: <http://clu.uni.no/icame/ij32/index.html>. Acesso em: 5 mar. 2022.

TEIXEIRA, E. D. **A linguística de Corpus a serviço do tradutor**: proposta de um dicionário de Culinária voltado para a produção textual. 2008. 400f. Tese (Doutorado) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008. Disponível em: <http://www.teses.usp.br/teses/disponiveis/8/8147/tde-16022009-141747/pt-br.php> Acesso em: 5 mar. 2022.

TOGNINI-BONELLI, E. **Corpus linguistics at work**. Amsterdam: John Benjamins, 2001.

TOGNINI-BONELLI, E. Theoretical overview of the evolution of corpus linguistics. **In**: O'KEEFFE, A.; McCARTHY, M. (Eds.) **The Routledge Handbook of Corpus Linguistics**. London, New York: Routledge, 2010. p. 14-27.

TOGNINI-BONELLI, E.; SINCLAIR, J. 'Corpora'. **In**: BROWN, K. (ed.) **Encyclopedia of Language and Linguistics**. 2. ed. Amsterdam: Elsevier, 2006. p. 206–19.

TOURY, Gideon. **Descriptive Translation Studies – and Beyond** (2nd ed.). Amsterdam: John Benjamins Publishing Company, 2012.

TRIBBLE, Christopher. What are concordances and how are they used? *In* O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p.167 – 183. London and New York: Routledge – Taylor and Francis Group, 2010.

UNICEF: **The Rights Start to Life** – A Statistical Analysis of Birth Registration. The United Nations Children's Fund, 2005. Disponível em: [http://www.unicef.org/protection/BirthReg10a\\_rev.pdf](http://www.unicef.org/protection/BirthReg10a_rev.pdf) Acesso em: 5 mar. 2022.

UNITED NATIONS PUBLICATIONS: **Principles and Recommendations for a Vital Statistics System** – Revision 2. Economic and Social Affairs, 2001. Disponível em: <http://unstats.un.org/unsd/default.htm> Acesso em: 5 mar. 2022.

VENUTI, L. **The Scandals of Translation**: Towards an Ethics of Difference. London: Routledge, 1998.

VENUTI, L. **The Translators Invisibility**: A history of translation. London: Routledge, 1995.

VINCZE, O. **Learning multiword expressions from corpora and dictionaries**. 2015. 397 f. Tese (Doutorado) - Departamento de Galego-Português, Francês e Linguística, Universidade da Coruña, Corunha, 2015. Disponível em: <http://ruc.udc.es/dspace/handle/2183/16191>. Acesso em: 10 jan. 2023.

WALTER, Elizabeth. Using corpora to write dictionaries. *In* O'KEEFFE, Anne; McCARTHY, Michael (Eds). **The Routledge Handbook of Corpus Linguistics**. p. 426 – 443. London and New York: Routledge – Taylor and Francis Group, 2010.

WEINERT, R. The Role of Formulaic Language in Second Language Acquisition:

a review. **Applied Linguistics**, v.16, p: 180 - 205. 1995. Disponível em:  
<http://pocketknowledge.tc.columbia.edu/home.php/viewfile/79646> Acesso em: 10 jan. 2023.

WILLIAMS, J.; CHESTERMAN, A. **The Map**: a beginner's guide to doing research in translation studies. Manchester: St. Jerome Publishing, 2002.

WILSON, S. H. (ed.) **The U.S. Justice System** – An Encyclopedia. Santa Barbara, California: ABC-CLIO, LLC, 2012.

WITHERS, J. L. **The Notable Notary**: The Complete Guide on How to Establish, Market, and Achieve Success with a Notary Public Business. 1<sup>st</sup> ed. The United States: Lulu Press, 2005.

WRIGHT, S. E. TermBase Management. **In**: WRIGHT, S. E., & BUDIN, G. (Eds.), **Handbook of Terminology Management**: Basic Aspects of Terminology Management (Vol. 1, pp. 147-168). John Benjamins Publishing Company. 1997.

WÜSTER, E. **Introducción a la Teoría General de la Terminología y a la Lexicografía Terminológica**. Institut Universitari de Lingüística Aplicada – Universitat Pompeu Fabra. Barcelona, 1998.

ZILIO, L. Colocações especializadas em alemão e português na área de Cardiologia. **TradTerm**, São Paulo, v. 20, dezembro/2012, p. 146-177

ZILIO, L. Terminologia Textual e Linguística de *Corpus*: Estudo em parceria. **In**: PERNA, C. B. L.; DELGADO, H. O. K.; FINATTO, M. J. B. (Orgs.) **Lingagens Especializadas em Corpora**: Modos de Dizer e Interfaces de Pesquisa. p. 128-151. Porto Alegre: EDIPUCRS, 2010.