

EQUIVALENTES LEXICAIS EM PORTUGUÊS E INGLÊS: CERTIDÕES BRASILEIRAS E AMERICANAS EM PERSPECTIVA COMPARADA

Vanessa Hagemeyer Burgo, Eduardo Francisco Ferreira, Guilherme Magri da Rocha

vanessa.hagemeyer@ufms.br, eduardo.ferreira@ifc.edu.br, g.magri@ufms.br,

Instituto Federal de Mato Grosso do Sul

V Seminário de Pós-graduação do IFMS – SEMPOG IFMS 2025

Resumo. *Este estudo, fundamentado na Linguística de Corpus, tem como objetivo realizar uma análise comparativa dos equivalentes lexicais identificados em certidões brasileiras e americanas, considerando os documentos originais e suas respectivas traduções juramentadas no par linguístico português/inglês. Propõe-se um percurso para a extração e a análise desses termos, com o intuito de evidenciar estratégias que favoreçam a escolha de equivalentes mais adequados na língua de chegada, contribuindo para a precisão e a qualidade da tradução. A análise evidencia como a tradução desses documentos pode preservar ou transformar a estrutura lexical do texto de origem, levando em conta os efeitos da simplificação textual e os desafios específicos da tradução juramentada. Os resultados desta pesquisa podem contribuir para reflexões acerca da prática tradutória e do uso de ferramentas de análise lexical em contextos oficiais.*

Palavras-Chave. *equivalentes lexicais; certidões; português/inglês.*

Abstract. *This study, based on Corpus Linguistics, aims to conduct a comparative analysis of the lexical equivalents identified in Brazilian and American certificates, taking into account the original documents and their respective sworn translations in the Portuguese/English language pair. We propose a methodology for extracting and analyzing these terms, with the purpose of highlighting strategies that support the selection of the most appropriate equivalents in the target language, thereby contributing to translation accuracy and quality. The analysis demonstrates how the translation of these documents can either preserve or transform the lexical structure of the source text, considering the effects of textual simplification and the specific challenges of sworn translation. The findings of this research may contribute to reflections on translation practice and the use of lexical analysis tools in official contexts.*

Keywords. *lexical equivalents; certificates; Portuguese/English.*

Resumen. *Este estudio, fundamentado en la Lingüística de Corpus, tiene como objetivo realizar un análisis comparativo de los equivalentes léxicos identificados en certificados*

brasileños y estadounidenses, considerando los documentos originales y sus respectivas traducciones juradas en el par lingüístico portugués/inglés. Se propone un procedimiento para la extracción y el análisis de estos términos, con el fin de evidenciar estrategias que favorezcan la elección de equivalentes más adecuados en la lengua de destino, contribuyendo a la precisión y calidad de la traducción. El análisis demuestra cómo la traducción de estos documentos puede preservar o transformar la estructura léxica del texto de origen, teniendo en cuenta los efectos de la simplificación textual y los desafíos específicos de la traducción jurada. Los resultados de esta investigación pueden contribuir a reflexiones sobre la práctica traductológica y el uso de herramientas de análisis léxico en contextos oficiales.

Palabras clave. *equivalentes léxicos; certificados; portugués/inglés.*

1. Introdução

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.

O desenvolvimento tecnológico e o incremento das pesquisas científicas têm propiciado a criação de múltiplas interfaces, as quais abrem novos campos de investigação e, conseqüentemente, suscitam a demanda por produtos terminológicos que auxiliem o tradutor na interpretação de textos especializados. Inserido nesse contexto, o presente estudo, ancorado na Linguística de Corpus, tem por objetivo realizar uma análise comparativa dos equivalentes lexicais identificados em certidões brasileiras e norte-americanas, considerando tanto os documentos originais quanto suas traduções juramentadas no par linguístico português/inglês.

Os registros documentais que dão origem às certidões apresentam ampla diversidade — abarcando desde certidões positivas e negativas até as de propriedade, de tributos ou de filiação, entre outras. Para fins desta investigação, entretanto, delimitou-se o corpus às certidões de nascimento, casamento e óbito emitidas no Brasil e nos Estados Unidos, com foco específico no par linguístico português brasileiro/inglês americano. Almeja-se, assim, oferecer subsídios que contribuam para a prática tradutória, em especial no enfrentamento

das especificidades terminológicas que caracterizam esse gênero documental.

2. Linguística de Corpus

De acordo com Taylor (2008, p. 180), há várias definições para a Linguística de Corpus (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”¹ (tradução nossa). 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 nossa²): “[...] 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 nossa³): “[...] 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 nossa⁴): “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.”

¹ [...] 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.

² 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.

³ [...] corpus linguistics is essentially a methodology or set of methodologies, rather than a theory of language description.

⁴ 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.

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.”

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 [...]”

2.1 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 nossa⁵): “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 nossa⁶): “Um corpus é uma coleção de amostras de linguagem que ocorrem naturalmente, armazenadas eletronicamente.”

⁵ 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.

⁶ A corpus is an electronically stored collection of samples of naturally occurring language.

3. Bennet (2010, p. 14, tradução nossa⁷): “Um corpus é uma coleção de textos autênticos, selecionados por meio de critérios, armazenados eletronicamente.”

4. Reppen (2010a, p. 02, tradução nossa⁸): “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.”

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.

Sinclair (2005, p. 1, tradução nossa⁹), 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.”

3. Discussão dos dados

Nas pesquisas em Linguística de Corpus, o uso de ferramentas digitais é essencial para identificar padrões linguísticos em diferentes tipos de textos. Entre essas ferramentas, destaca-se o *Concord*, do pacote *WordSmith Tools* ⁹¹⁰ (Scott, 2020), que permite realizar buscas a partir da definição de nódulos específicos. No caso deste estudo, ao retomar a tipologia das certidões e das traduções juramentadas, já foi observado que ambos os documentos carregam a noção de “fê pública”, expressa em fórmulas redacionais utilizadas pelos cartorários e, posteriormente, pelos tradutores juramentados em suas versões traduzidas, sobretudo nos trechos finais.

Nessa parte do trabalho foi descrito como esse item lexical ocorre nos documentos,

⁷ A corpus is a principled collection of authentic texts stored electronically.

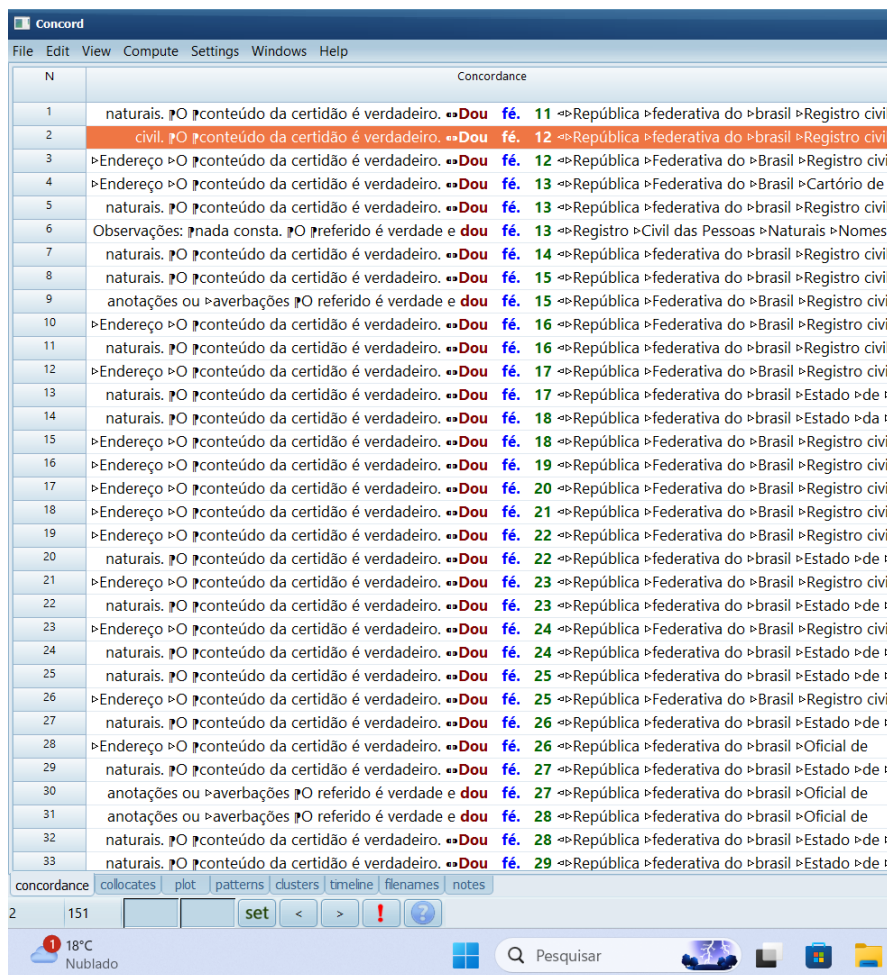
⁸ In the world of corpus linguistics, corpus is a large, principled collection of naturally occurring texts (written or spoken) stored automatically.

⁹ 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.

¹⁰ <https://www.lexically.net/wordsmith/>

contando com a presença de exemplos. Coincidentemente, a palavra “fé” figura entre as 10 palavras mais recorrentes nos textos originalmente escritos em português. 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 1, a seguir:

Figura 1 - 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 nódulo. No primeiro caso, à direita, nas certidões, o nódulo 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 nódulo 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 nódulo 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.

4. Considerações Finais

O registro de atos civis, como nascimento, casamento ou óbito, constitui uma obrigação legal a ser cumprida em cartório. Trata-se de um procedimento jurídico regulamentado, no qual se define não apenas quais informações devem ser incluídas, mas também a forma de sua redação, assegurando a padronização dos documentos. A transposição dos dados do registro para a certidão segue parâmetros igualmente determinados, assim como ocorre com o assento original. No caso específico das traduções juramentadas, o trabalho do tradutor está sujeito a normas e restrições próprias, que delimitam sua atuação e estabelecem exigências adicionais de caráter jurídico. Considerando esse conjunto de regulamentações, percebe-se que a liberdade criativa tende a ser limitada, conduzindo o tradutor a adotar uma postura mais contida e objetiva em seu trabalho.

A tradução juramentada transmite a ideia de uma cópia fiel do documento de origem, legitimada para uso na cultura de chegada por meio de selo, carimbo, assinatura e declaração

de veracidade emitida pelo tradutor. Nos corpora analisados, a prática tradutória manteve-se próxima à língua fonte, privilegiando a literalidade e a fidelidade ao texto original. Essa estratégia, ainda que resulte em construções distintas do padrão da língua-alvo, revelou-se adequada, pois assegurou a transparência do processo tradutório e preservou a função social do documento.

Em síntese, a análise indica que a tradução juramentada se estrutura a partir de um equilíbrio entre o rigor normativo e a responsabilidade do tradutor, em que a literalidade constitui um recurso essencial para garantir a legitimidade, a autenticidade e a eficácia comunicativa do texto em contextos oficiais. Observa-se que a tradução pode, simultaneamente, preservar e reorganizar a estrutura lexical do documento de origem, considerando os efeitos da simplificação textual e os desafios próprios desse tipo de prática tradutória. Os resultados desta investigação oferecem subsídios para reflexões críticas acerca da prática da tradução juramentada e do emprego de ferramentas de análise lexical em contextos que exigem precisão e conformidade jurídica.

5. Referências

- 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.
- BAKER, M. Corpora in Translation Studies: An Overview and Some Suggestions for Future Research. **Target** 7(2), 1995. p. 223-243.
- BAKER, P.; HARDIE, A.; McENERY, T. **A Glossary of Corpus Linguistics**. Edinburgh: Edinburgh University Press, 2006.
- 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.
- 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.

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.

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

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

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.

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/matraga/matraga24/matraga.html>. Acesso em: 10 dez. 2022.

OLOHAN, M. **Introducing Corpora in Translation Studies**. London, New York: Routledge, 2004.

PALUMBO, G. **Key Terms in Translation Studies**. London, New York: Continuum International Publishing Group, 2009.

REPPEN, R. **Using Corpora in the Language Classroom**. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi, Tokio, Mexico City: CUP, 2010.

SCOTT, M. **WordSmith Tools: Version 9.0**. Stroud: Lexical Analysis Software Ltd., 2020.

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.pgletras.uerj.br/matraga/matraga24/matraga.html>. Acesso em: 10 dez. 2022.

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.: **Reading concordances: an introduction**. New York: Pearson/Longman, 2003.

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.

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.