

Linguística de Corpus

Perspectivas

Organizadoras:

Maria José Bocorny Finatto

Rozane Rodrigues Rebechi

Simone Sarmiento

Ana Eliza Pereira Bocorny



INSTITUTO
DE LETRAS
UFRGS



UNIVERSIDADE FEDERAL
DO RIO GRANDE DO SUL



UNIVERSIDADE
FEDERAL DO RIO
GRANDE DO SUL

Reitor

Rui Vicente Oppermann

Vice-Reitora e Pró-Reitora
de Coordenação Acadêmica

Jane Fraga Tutikian



INSTITUTO
DE LETRAS
UFRGS

Universidade Federal
do Rio Grande do Sul
Instituto de Letras

Diretor

Sérgio de Moura Menuzzi

Vice-diretora

Beatriz Cerisara Gil

Linguística de Corpus

Perspectivas

Organizadoras:

Maria José Bocorny Finatto

Rozane Rodrigues Rebechi

Simone Sarmento

Ana Eliza Pereira Bocorny

© dos autores
1ª edição: 2018

Direitos reservados desta edição:



Esta licença permite que outros distribuam, remixem, adaptem e criem a partir dos trabalhos aqui publicados, mesmo para fins comerciais, desde que lhes atribuam o devido crédito pela criação original.

Capa: Ethel Kawa
Preparação de originais: Carlos Batanoli Hallberg
Revisão: Lia Cremonese
Editoração eletrônica: Fernando Piccinini Schmitt

Esta coletânea foi publicada graças ao apoio recebido da FAPERGS, processo 17/0399-3, Edital 06/2016 EDITAL FAPERGS 06/2016 – AOE, que apoiou o XVI Encontro de Linguística de *Corpus* (ELC 2017) e IX Escola Brasileira de Linguística Computacional (EBRALC 2017). Esta coletânea é um livro derivado do evento, reúne uma seleção de trabalhos gerados a partir de diferentes atividades de ambos os eventos. Todos os trabalhos aqui publicados foram avaliados por Comissão Científica especialmente convidada. Os Anais do evento correspondem a uma outra publicação denominada “Caderno de Resumos do ELC-EBRALC 2017”, ISBN: 9788561424183.

O direito autoral dos textos deste livro foi liberado por seus autores e organizadores, visto que é proibida a sua comercialização, sendo seu acesso livre e gratuito através do *site* do PPG-LETRAS-UFRGS, na guia E-BOOKS. A edição é do Instituto de Letras da UFRGS.

Versão DIGITAL gratuita disponível em:
PPG-LETRAS-UFRGS:
<https://www.ufrgs.br/ppgletras/ebooks.html>

Site do evento:
<http://www.ufrgs.br/elc-ebralc2017>



L755 Linguística de *corpus* : perspectivas [recurso eletrônico] / Organizadoras: Maria José Bocorny Finatto, Rozane Rodrigues Rebechi, Simone Sarmento, Ana Eliza Pereira Bocorny. — Porto Alegre: Instituto de Letras - UFRGS, 2018.
575 p.

Requisitos do sistema: Adobe Reader.
Modo de acesso: World Wide Web

1. Linguística. 2. Linguística de *corpus*. I. Finatto, Maria José Bocorny. II. Rebechi, Rozane Rodrigues. III. Sarmento, Simone. IV. Bocorny, Ana Eliza Pereira.

CDD 410

Catálogo na publicação: Vladimir Luciano Pinto – CRB 10/1112

ISBN 978-85-64522-36-7

O uso de *corpus* paralelo e comparável para descrever padrões de uso na tradução de abreviaturas e acrônimos de termos médicos

The use of parallel and comparable *corpus* to describe patterns of use in the translation of abbreviations and acronyms of medical terms

Márcia Moura da Silva
Gabriele Paparelli

Resumo: O objetivo do presente artigo é apresentar pesquisa em andamento que analisa o comportamento tradutório de abreviaturas e acrônimos de termos médicos no par linguístico português-inglês para então propor um glossário *on-line* que sirva como fonte para tradutores, revisores e pesquisadores. A pesquisa tem como base teórica a tradução técnico-científica e os Estudos da Tradução baseados em *corpus* e segue alguns princípios e técnicas desenvolvidas pela Linguística de *Corpus* para compilar um *corpus* paralelo e dois *corpora* comparáveis. Ainda que para facilitar a divulgação do conhecimento científico haja uma tendência de manter esses elementos em sua forma em língua inglesa, resultados preliminares nos mostram um número estatisticamente expressivo de abreviaturas e acrônimos que seguem os padrões de suas respectivas línguas.

Palavras-chave: Abreviatura. Acrônimo. Estudos da Tradução baseados em *corpus*. Tradução.

Márcia Moura da Silva – Professora adjunta do Instituto de Letras, Departamento de Línguas Modernas, da Universidade Federal do Rio Grande do Sul, doutora em Estudos da Tradução pela Universidade Federal de Santa Catarina – marcia.moura@ufrgs.br.

Gabriele Paparelli – Aluna de graduação do curso Bacharelado em Letras, habilitação português/inglês, da Universidade Federal do Rio Grande do Sul, bolsista de Iniciação Científica PROBIC FAPERGS UFRGS – gadifab@hotmail.com.

Abstract: The purpose of this paper is to present ongoing study that analyses the translation of abbreviations and acronyms of medical terms in the Portuguese-English pair in order to propose an on-line glossary to assist translators, revisers, and researchers. The study is based on Techno-Scientific Translation and Corpus-based Translation Studies. It borrows some principles and techniques developed within Corpus Linguistics to compile one parallel corpus and two comparable corpora. Although there is a tendency to keep these elements in English as means to facilitate the spread of scientific knowledge, preliminary results show a statistically significant number of abbreviations and acronyms that follow the patterns of their respective languages.

Keywords: Abbreviation. Acronym. Corpus-Based Translation Studies. Translation.

1 Introdução

O presente artigo tem por objetivo descrever pesquisa em andamento¹ que partiu da experiência de uma das autoras com tradução de textos médicos nos pares linguísticos português-inglês e inglês-português². Durante seu trabalho, frequentemente se deparou com a necessidade de lidar com abreviações e acrônimos, elementos bastante recorrentes nesse gênero textual. A falta de padronização resultou em pesquisas demoradas por textos paralelos ou *sites* de abreviaturas à procura de correspondentes. Mesmo quando encontrados, era preciso decidir, em meio a uma profusão de informação na internet (nem sempre confiável), se determinada abreviatura era mais comumente traduzida ou deixada na forma que aparecia em língua inglesa, por exemplo. Assim, pensou-se na criação de material de pesquisa que pudesse facilitar o trabalho de tradutores e outros profissionais e pesquisadores de tradução.

Academicamente falando, basta uma busca pelas dissertações e teses defendidas nos programas de pós-graduação das universidades brasileiras para constatar que, embora as pesquisas com tradução de textos médicos tenham aumentado nos últimos anos, ainda há pouco interesse acadêmico em investigar as escolhas tradutórias para abreviaturas e acrônimos.

Entre as pesquisas existentes na área, podemos citar alguns trabalhos nos pares linguísticos inglês-português e português-inglês com variados enfoques, que incluem metodologia de *corpus* (COULTHARD, 2005), tradução de resumos médicos (PASQUALI; PINTO, 2013) e tradução de legendagem (COLLET, 2012). Entre os trabalhos existentes, destacamos o de Paiva (2009), que tem por objeto as subáreas de cardiologia e cardiovascular. A pesquisadora vale-se dos Estudos

¹ *A tradução de abreviaturas e acrônimos de termos médicos no par linguístico português-inglês.*

² Por muitos anos, a professora Márcia M. da Silva traduziu material na área da saúde que incluía transcrições no par linguístico português-inglês e material impresso no par linguístico inglês-português. As traduções eram parte de grandes projetos que envolviam médicos, profissionais da saúde e representantes de instituições governamentais da área da saúde.

da Tradução baseados em *corpus* (ETC) como pressuposto teórico, tendo como objetivo observar as estratégias que evidenciam traços de simplificação e de explicitação³. Embora o estudo aqui descrito compartilhe com o trabalho de Paiva a observação de textos médicos com base nos ETC, o foco nas abreviaturas e acrônimos em uma subárea distinta (reumatologia) enriquece a discussão sobre a tradução do texto médico em geral, com destaque para a situação real de uso dos itens objeto desta investigação.

Ademais, acreditamos que a observação, por meio de *corpora*, de um maior número de especialidades amplia a pesquisa na área médica como um todo, possibilitando verificar se determinados padrões de uma especialidade se reproduzem em outras, por exemplo, ou mesmo mostrar particularidades de certas subáreas.

Ainda que haja, como nos lembra Franco Aixelá (2009), certa disposição para manter abreviaturas inalteradas em textos traduzidos do inglês para preservar o caráter de dada disciplina e facilitar a divulgação do conhecimento técnico-científico, como mencionado, a prática nos mostra que esses elementos estão longe de ser estanques e que tampouco há consenso em relação a como traduzi-los, o que acaba por dificultar o trabalho do tradutor. Como grande parte da divulgação do conhecimento técnico-científico é feita em língua inglesa, a tradução de artigos médicos para essa língua vem se tornando imprescindível. Assim sendo, o principal objetivo da pesquisa em andamento é verificar o comportamento tradutório em relação a abreviaturas e acrônimos de termos médicos da área da reumatologia, no par linguístico português-inglês, para que se possa propor um glossário *on-line* como material de consulta.

A pesquisa tem como base teórica a tradução técnico-científica (AZENHA, 1996, 1999; AIXELÁ, 2009) e os Estudos da Tradução baseados em *Corpus* (BAKER, 1993, 1995; OLOHAN, 2004; TYMOCZKO, 1998). Um *corpus* paralelo e dois *corpora* comparáveis, que serão descritos em maiores detalhes na metodologia, serão compilados seguindo alguns princípios e técnicas desenvolvidos pela Linguística de *Corpus* (BAKER, 2013; BERBER SARDINHA, 2002, 2004; BIBER, 1993; LEECH, 1991). A combinação desses *corpora* possibilitará verificar se determinadas abreviaturas e acrônimos são mantidos inalterados ou traduzidos, bem como seus padrões de uso, indicando se determinado padrão está restrito ao texto traduzido ou se é mais frequente nele. Espera-se que tal conhecimento possa auxiliar tradutores e pesquisadores na produção de textos que tenham mais aceitação na comunidade internacional. Para Olohan (2004), a influência da língua de partida (LP) em padrões de uso observados na língua de chegada (LC) seria um dos principais benefícios da combinação entre esses dois

³ Na *explicitação*, o tradutor adiciona, de maneira explícita, no TC, componentes que estariam somente implícitos no TP e na *simplificação*, o tradutor usa uma linguagem mais simples no TC para facilitar a compreensão do leitor.

tipos de *corpora*. Porém, visto que as traduções objeto desta pesquisa são em língua inglesa, e dada sua influência na área aqui pesquisada, espera-se o cenário inverso, ou seja, maior influência da língua inglesa sobre as abreviaturas e acrônimos usados em língua portuguesa, embora resultados preliminares com o *corpus* paralelo já apontem um número estatisticamente relevante de abreviaturas e acrônimos que seguem os padrões das respectivas línguas, como mostramos nos resultados preliminares abaixo.

A escolha da subárea da reumatologia se deu pelo fato de grande parte do trabalho de tradução feito por uma das autoras ter sido nessa área, sobretudo em artrite reumatoide (AR). Segundo o Ministério da Saúde, em 2011, as doenças reumáticas já acometiam 12 milhões de brasileiros⁴. Como o tratamento medicamentoso dessas doenças é garantido pelo Sistema Único de Saúde (SUS), há bastante interesse por parte dos grandes laboratórios médicos em negociar a venda de produtos de última geração ao governo brasileiro, daí o número considerável de projetos de tradução na área. Embora o foco seja nas abreviaturas e acrônimos usados na subárea da reumatologia, não se desprezarão abreviaturas e acrônimos que sejam compartilhados com outras subáreas, como o exemplo do desfecho clínico⁵ *QoL* (*quality of life/qualidade de vida*), ou micro-organismos que impactam doenças tratadas pela reumatologia, como é o caso de HIV (vírus da imunodeficiência humana) e HPV (papilomavírus humano.)

Em relação à escolha dos textos, essa se deu pelo fato de serem esses textos de revistas representativas dessa subárea e também por sua respeitabilidade. Por serem publicações de importantes sociedades médicas, pressupõe-se que estejam em conformidade com os padrões de aceitabilidade de cada língua.

É importante mencionar que, como acontece com o texto acadêmico, a publicação do texto médico em língua inglesa dá ao trabalho maior visibilidade internacional, visto ser o inglês *lingua franca* e ter, assim, abrangência global. Ademais, países como Estados Unidos e Inglaterra são detentores de grande conhecimento nessa área e financiadores de pesquisas robustas. Dessa maneira, são muito mais comuns textos traduzidos para o inglês do que o contrário, por isso o *corpus* paralelo consiste nos originais em língua portuguesa e suas traduções em língua inglesa.

A seguir, discutimos brevemente os pressupostos teórico-metodológicos que nortearão a pesquisa aqui apresentada.

⁴ Disponível em: <<http://www.blog.saude.gov.br/promocao-da-saude/29041-saude-alerta-para-prevencao-as-doencas-reumaticas>>.

⁵ “Evento de investigação supostamente causado pelo fator em estudo. Exemplo doença, complicação, efeito terapêutico” (disponível em: <<http://pt.slideshare.net/flavioes/tipos-de-estudo-1814732>>).

2 Tradução técnico-científica

Segundo Aixelá (2004), a tradução técnico-científica não recebe a mesma reverência que a tradução literária dentro dos Estudos da Tradução porque seus textos são vistos como menos criativos e elaborados, requerendo muito pouco além de conhecimento terminológico. Porém, o autor sugere que essa tradução se destaca entre outros gêneros, pertencendo a um campo de pesquisa independente. Para ele, diferentemente de outros tipos de traduções, a técnico-científica precisa respeitar tanto a função referencial da língua quanto as convenções da linguagem técnica no que diz respeito à precisão para melhor se entender o mundo.

Aixelá (2009) trabalha com a noção de que a tradução é regida por uma tensão dupla – a força centrípeta, que a leva em direção às propostas do texto de partida (TP), e a centrífuga, que a leva em direção ao texto de chegada (TC) no que diz respeito às noções do que vem a ser correto e também às expectativas do leitor. Segundo o autor, ao usar exclusivamente a primeira, o tradutor estaria criando um texto repleto do que denomina “interferência”, que pode ser minimizada quando as forças são equilibradas. Porém, em consequência dos limites de entrega bastante curtos com os quais trabalham os tradutores, além de remunerações nem sempre ideais, Aixelá acredita que tradutores somente se desviam do TP se julgam estritamente necessário, pois é mais rápido produzir uma tradução mais próxima a ele.

Em relação a textos técnico-científicos, Aixelá (2009) sugere que pessoas que compartilham uma profissão tendem a criar sua própria terminologia, seja por necessidade (precisão e clareza), ou exclusividade, o que acaba gerando certa opacidade para as pessoas de fora, mas que fortalece o sentimento de pertencimento dos profissionais em questão. Outro aspecto importante é o papel da língua inglesa, como já mencionado, que acaba por fazer com que traduções a partir dessa língua tenham um grau maior de interferência, aumentando, assim, a probabilidade de aceitação, por estarem promovendo a internacionalização da terminologia, facilitando, assim, o fluxo de conhecimento técnico-científico. Como há vários neologismos em inglês nessa área, tradutores tendem a mantê-los para que seus textos ganhem prestígio e exclusividade. Por outro lado, as interferências podem ser rejeitadas porque existe a vigilância de autoridades na cultura de chegada que se preocupam com a pureza linguística.

Outro autor que discorre sobre o texto técnico-científico e sua tradução, Azenha (1999) também nos adverte sobre essa ideia generalizada de que os textos técnicos são estanques e que os problemas de tradução relacionados a eles estão restritos ao plano léxico-terminológico. Para o autor, eles são “formas híbridas expostas à ação de um número elevadíssimo de variáveis e a terminologia, longe de ser estática, é dinâmica e admite uma margem de subjetividade no tratamento de seu objeto” (p. 11). O sistema de pesos e medidas é um bom exemplo dessa variedade de textos técnicos. Ao procurar uma receita de bolo inglês na internet,

por exemplo, o brasileiro espera encontrá-la no sistema métrico, não no imperial, que é o mais usado na Grã-Bretanha. Assim, ao se traduzir essa receita para o português, deve-se levar essa particularidade em consideração e buscar estratégias apropriadas para melhor entendimento por parte do leitor de chegada, visto que a tradução literal por si só não daria conta desse objetivo.

Ainda que textos técnico-científicos estejam sujeitos à variação, Aixelá (2009) observa que os tradutores desse tipo de texto optam pela interferência, ou seja, por manter a terminologia inalterada, em vez de uma tradução mais “pura”. Ele também nos lembra que muitos tradutores em formação se queixam de professores que exigem pureza na LC, quando, de fato, o mercado rejeita traduções que não se enquadram em modelos internacionais. Por outro lado, Azenha (1999) enfatiza que, pela natureza dinâmica desses textos, é de vital importância que os professores conscientizem seus alunos para a variabilidade desse tipo de tradução e também os encorajem a exercer o pensamento crítico, pois somente assim estarão mais bem preparados para as demandas do mercado.

Em relação a essas demandas, embora as pesquisas acadêmicas na área dos Estudos da Tradução tenham tradicionalmente o texto literário como foco, a demanda de mercado pela tradução técnico-científica é bem maior. Assim, ainda que possa existir uma tendência à não tradução de certos elementos dentro do texto técnico-científico, como aponta Aixelá (2009), para se conformar a determinados padrões, é importante que tradutores em formação sejam conscientizados de que existe mais de uma solução para a tradução de terminologia e que o grau de interferência depende do quanto ela é aceita na cultura de chegada. Dessa maneira, futuros tradutores estarão mais aptos a suprir essa demanda.

Como pesquisadoras, interessa-nos observar as variações inerentes ao texto técnico-científico e as soluções encontradas por tradutores na produção de um TC adequado. Porém, independentemente do fato de um tradutor traduzir ou não termos técnicos, engana-se quem pensa que a decisão de transpô-los inalterados para o TC seja uma decisão automática, pois, para tanto, o tradutor deve considerar aspectos que vão além da questão terminológica. É nesse sentido que vemos vantagem na combinação de estudos voltados a textos técnico-científicos e a metodologia de *corpus*, pois, graças aos avanços nas ferramentas de construção de *corpus*, a observação de outras traduções vem se tornando cada vez mais precisa e imediata e vem aumentando consideravelmente a confiabilidade dos resultados por permitir que um número representativo de textos seja incluído em um único *corpus*. Ademais, como essas ferramentas possibilitam a observação do objeto de estudo em contexto, tornou-se mais fácil expandir as discussões para além do plano linguístico. Porém, antes de descrevermos mais detalhadamente o processo de compilação dos *corpora* que estão sendo construídos para a análise aqui proposta, tecemos algumas considerações breves sobre a tradução de abreviaturas e acrônimos e sobre os Estudos da Tradução baseados em *Corpus*.

2.1 Tradução de abreviaturas e acrônimos

Trabalhamos com a noção de que a abreviatura se refere à “apresentação de uma palavra por meio de algumas de suas sílabas ou letras”⁶, tendo em conta seus aspectos fônicos, como AAS (Ácido Acetilsalicílico); enquanto o acrônimo “é pronunciado como se fosse uma palavra única, como OTAN e NASA”⁷.

Do ponto de vista morfológico, de um modo geral, as abreviaturas⁸ se classificam em i) simples, formadas pelas iniciais ou qualquer outra letra de um grupo de palavras (e.g. ROM/*Read-Only Memory*); ii) palavras cortadas (*clipped words*), formadas por três ou mais letras contíguas de uma mesma palavras (e.g. del/delete); iii) combinações, que são junções de palavras previamente cortadas ou abreviadas (e.g. *modem /modulator-demodulator*) e iv) complexas, que combinam duas ou mais abreviaturas simples por meio de símbolos tipográficos, como é o caso do hífen, (e.g. CD-ROM). (Cf. BELDA, 2004).

Em relação à tradução dessas reduções de palavras, para Aixelá (2009), por exemplo, há certa disposição em mantê-las inalteradas em textos traduzidos do inglês para manter o caráter da disciplina a qual pertence e facilitar a divulgação do conhecimento técnico-científico:

Assim, se você deseja que sua disciplina seja descrita em seus próprios termos, geralmente é mais fácil importar palavras e estruturas pré-fabricadas do que criar outras; além disso, há a vantagem da exclusividade devido ao fato de que termos importados tendem a ser mais opacos que outros derivados de palavras preexistentes na LC (sendo essa uma das razões pela qual os jargões técnicos modernos tendem a ser mais opacos para o leitor geral quando não estão em inglês). Esse motivo é também reforçado pelo argumento da promoção da internacionalização da terminologia, que facilitaria, assim, o fluxo de conhecimento técnico-científico. *Essa é uma justificativa importante e recorrente dada para a não tradução de abreviaturas, que provavelmente representam o grau máximo de interferência na tradução técnico-científica.* (AIXELÁ, 2009, p.80-81, grifo e tradução nossos⁹)¹⁰

⁶ Dicionário Aurélio.

⁷ “is pronounced as if it were a single word, in the manner of NATO and NASA.” (Farfex Dictionary: <<http://www.thefreedictionary.com/acronym>>).

⁸ Ao apresentar tal classificação, Belda (2004) refere-se somente à abreviatura, porém, seguindo as definições que adotamos, alguns de seus exemplos incluem também acrônimos.

⁹ Todas as traduções feitas das citações em língua inglesa foram feitas por nós.

¹⁰ “Thus, if you want your discipline to be described in its own terms, it is generally easier to import ready-made words and structures than to create new ones, not forgetting the bonus of exclusivity due to the fact that imported terms tend to be more opaque than others derived from pre-existing TL words (one of the reasons why modern technical jargons tend to be more opaque for the general reader when not in English). This motive is also supported by the argument of promoting the internationalisation of your terminology, and thus facilitating the flow of scientific and technical knowledge. This is an important and often quoted reason for the nontranslation of abbreviations, which probably represent the maximum degree of interference in technical and scientific translation”.

Já Belda (2004), que escreve sobre a tradução de abreviaturas na área da tecnologia da informática no par linguístico inglês-espanhol, chama a atenção para um dos principais aspectos que o tradutor deve considerar na tradução de abreviaturas – a ordem das palavras. A língua inglesa é flexível o suficiente para permitir a conversão de substantivos em adjetivos simplesmente colocando-os antes de outros substantivos. Em espanhol, e também em português, tal conversão não é comum, sendo que normalmente usa-se algum conector, sobretudo preposição, como é o caso de *CTS (carpal tunnel syndrome)*, traduzido para o português como *STC (síndrome do túnel do carpo)*¹¹. Como o *corpus* paralelo da pesquisa aqui descrita consiste em textos escritos originalmente em português com suas respectivas traduções para o inglês, tal conversão ocorre no texto traduzido.

O autor, porém, ratifica a ideia de que a prática tradutória segue o padrão de uso que, por sua vez, está atrelado à aceitação ou não por profissionais da LC de termos inalterados. O autor dá o exemplo da tentativa de usar a abreviatura espanhola *MMM (Malla Máxima Mundial)* para a inglesa *WWW (World Wide Web)*, uma proposta feita pelo CVC (Centro Virtual Cervantes), que acabou não se concretizando.

3 Estudos da Tradução baseados em *corpus*

Baker (1995) define *corpus* como sendo uma

coleção de textos armazenados em formato digital que podem ser analisados automaticamente ou semiautomaticamente de várias maneiras [...] O mais importante é que seja construído para um propósito específico seguindo critérios explícitos de desenho a fim de garantir que seja representativo da área ou língua que pretende estudar. (BAKER, 1995, p. 225)¹²

Para Baker (1993), que deu início às pesquisas de Estudos da Tradução baseados em *corpus*, os textos traduzidos “*registram eventos comunicativos genuínos e, como tal, não são inferiores nem superiores a outros eventos comunicativos em qualquer língua. Contudo, são diferentes, e a natureza dessa diferença precisa ser explorada e registrada*”¹³ (p. 234).

¹¹ Também foram encontradas outras variações do termo em português como, por exemplo, *síndrome to túnel cárpico, síndrome do túnel carpal e síndrome do túnel de carpo*.

¹² “collection of texts held in machine-readable form and capable of being analysed automatically or semi-automatically in a variety of ways [...] What is important is that 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 of language it aims to account for.”

¹³ “record genuine communicative events and as such are neither inferior or superior to other communicative events in any language. They are however different, and the nature of this difference needs to be explored and recorded.”

Essa visão do texto traduzido tem como origem os Estudos Descritivos da Tradução, apresentados por James Holmes em seu artigo seminal, de 1972, *The Name and Nature of Translation Studies*. Segundo o autor, os estudos descritivos “descrevem o fenômeno do ato tradutório e da(s) tradução/traduições como se manifestam no mundo de nossa experiência”^{14,15}.

Os trabalhos de Baker apoiam-se sobretudo nos trabalhos de Even-Zohar e Gideon Toury¹⁶. Na teoria do polissistema de Even-Zohar, na qual um determinado polissistema literário é parte de um polissistema cultural maior, os textos traduzidos são vistos como parte de seus originais, assim, ganhando relevância e sendo considerados dignos de investigação como um sistema em si.

Os ETC emprestam da Linguística de *Corpus* alguns de seus princípios e algumas das técnicas por ela desenvolvidas sobretudo no que diz respeito à metodologia de compilação de *corpus*, mas firmam-se como uma abordagem com forte base descritiva e que dá espaço a investigações bastante abrangentes¹⁷, como atesta Tymoczko (1998, p. 653):

Os Estudos da Tradução com *Corpus* (ETC) surgiram em um momento crítico na disciplina dos Estudos da Tradução. Emergindo da linguística de *corpus* e, portanto, inerentemente fiel às abordagens linguísticas da tradução, os ETC, ao mesmo tempo que marcam uma mudança das abordagens prescritivas para as descritivas, [...] levam em consideração tanto os pormenores do texto escolhido pelo tradutor individual, como padrões culturais mais amplos, sejam internos ou externos ao texto. Embora o material para os *corpora* estejam baseados nos meios linguísticos da tradução, as perguntas feitas a esses *corpora* podem servir para tratar não somente questões de língua e linguística, mas também questões de cultura, ideologia e crítica literária.¹⁸

Os estudos com *corpora* permitem, assim, estudar o fenômeno da tradução de diferentes pontos de vista e a partir de diferentes línguas e culturas, revelando tanto

¹⁴ Cf. Shuttleworth, M; Cowie, M. Dictionary of Translation Studies. Manchester (1997, p. 39).

¹⁵ “describe the phenomena of translating and translation(s) as they manifest themselves in the world of our experience”.

¹⁶ Toury’s *Descriptive Translation Studies- and Beyond* (1995) “has become an indispensable reference point for those working in this area. Toury’s early work came out of the context of polysystem theory, developed by his colleague Itamar Even-Zohar” (MUNDAY, 2009).

¹⁷ Em geral, os ETC não estão restritos à descrição de fenômenos linguísticos. A pesquisa de Baker (2013) sobre a representação da palavra *mulçumano* (*Muslim*) na imprensa britânica, por exemplo, tem por base a Análise Crítica do Discurso.

¹⁸ “*Corpus* translation studies (CTS) has emerged at a critical time in the discipline of Translation Studies. Growing out of *corpus* linguistics and thus inherently having an allegiance to linguistic approaches to translation, CTS at the same time marks a turn away from prescriptive approaches to translation toward descriptive approaches [...] and takes into account the smallest details of the text chosen by the individual translator, as well as the largest cultural patterns both internal and external to the text [A]lthough the materials of *corpora* are based upon the language medium of translations, interrogation of *corpora* can nonetheless serve to address not simply questions of language or linguistics, but also questions of culture, ideology, and literary criticism.”

as similaridades como as diferenças. Tymoczko (1998, p. 253) aponta ainda que ao longo da história da teoria da tradução as diferenças eram vistas negativamente e esperava-se um TC mais fiel ao TP; porém, hoje as diferenças são mais valorizadas.

A presente pesquisa utiliza dois tipos de *corpora*, a saber:

- a) Corpus paralelo – coleção de textos originais na língua A e suas traduções em língua B.
- b) Corpus comparável – i) *corpus* comparável monolíngue: consiste em coleção de textos originais em língua A e textos traduzidos em língua A e ii) *corpus* comparável bilíngue: coleção de textos originais em língua A e textos originais em língua B.

Como mencionado, Olohan (2004) acredita que a combinação desses dois tipos de *corpora* seja bastante vantajosa, pois através dela pode-se observar a influência da LP em padrões de uso observados na LC:

Uma abordagem metodológica que promete emergir em trabalhos futuros é a combinação de resultados de uma análise com *corpora* comparáveis e paralelos [...] Essa abordagem dupla pode ser particularmente frutífera para mensurar a influência do texto de partida sobre padrões de uso observados na língua traduzida. (OLOHAN, 2004, p. 192)¹⁹

3 Metodologia

Berber-Sardinha (2004, p. 18) aponta ser o *corpus* uma coleção de textos naturais grande e criteriosa que deve refletir apropriadamente a variedade escolhida “o mais fielmente possível”. Em caso de variedade específica, sugere o autor, o pesquisador deve ter o cuidado de escolher material que reflita pontualmente tal variedade para evitar vieses e contaminações.

Tal variedade acaba sendo um critério sugerido também em relação ao tamanho ideal do *corpus* de estudo, sobre o qual ainda não há consenso. Sinclair (1997), cujo trabalho está voltado para a língua geral para propostos lexicográficos, sugere que o *corpus* deva ser o mais extenso possível. Embora seja da mesma opinião, Berber-Sardinha (2004, p. 29) destaca a impossibilidade de se incluir um idioma inteiro em um *corpus*, por isso a necessidade de “*delimitar ao máximo a variedade (tipo de texto, por exemplo) incluída no corpus*”.

¹⁹ “A methodological approach that is likely to come to the fore in future work is the combination of findings from comparable *corpus* analysis and parallel *corpus* work, [...] This dual approach proves particularly fruitful in measuring the extent of any source-text influence on patterns of usage observed in translated language” (OLOHAN, 2004, p. 192).

Embora o tamanho seja um dos principais critérios a se pensar ao compilar um *corpus*, sua representatividade é considerada ainda mais relevante. Para Leech (1991, p. 27), um *corpus* é representativo “*contanto que os resultados baseados em seus conteúdos possam ser aplicados a um corpus hipotético maior*”²⁰.

Já Biber (1993) defende uma pesquisa teórica criteriosa antes de se desenhar um *corpus* para melhor identificar as características linguísticas que serão analisadas. Berber-Sardinha (2004, p. 29) ressalta que o *corpus* “*deve ser adequado aos interesses do pesquisador, que deve ter uma questão a investigar para a qual necessite de um corpus específico*”.

No caso dos *corpora* compilados para esta pesquisa, acreditamos serem representativos, pois temos claro nosso objetivo de observar o comportamento tradutório para abreviaturas e acrônimos na subárea da reumatologia, assim como o padrão de uso em textos escritos originalmente em língua portuguesa e em língua inglesa para servir a uma determinada população, ou seja, tradutores, revisores e pesquisadores dessa área. Tal representatividade é reforçada pelo fato de que os textos foram extraídos de periódicos bem conceituados e específicos da subárea de estudo, o que faz com que os *corpora*, ainda que não sejam grandes, tenham uma concentração maior de terminologia.

A seguir, descrevemos resumidamente o processo de compilação dos *corpora*.

3.1 Corpora

No presente estágio da pesquisa, estamos compilando um *corpus* paralelo com textos originais em língua portuguesa com suas traduções em língua inglesa da *Revista Brasileira de Reumatologia*. No próximo estágio, compilaremos dois *corpora* comparáveis: i) *corpus* com os textos traduzidos em língua inglesa dessa revista e textos escritos originalmente em inglês da revista *Rheumatology* (*corpus* comparável monolíngue), e ii) *corpus* com textos escritos originalmente em português da *Revista Brasileira de Reumatologia* e textos escritos originalmente em inglês da revista *Rheumatology* (*corpus* comparável bilíngue). A análise dos *corpora* está sendo feita utilizando as ferramentas ParaConc e AntConc, cujas funções são descritas na próxima seção.

Neste primeiro momento, optamos por trabalhar com as edições 2009 e 2010 (volumes nº 49 e 50, respectivamente) da *Revista Brasileira de Reumatologia*, sendo o volume nº 49 a primeira edição a ser traduzida para o inglês. Até o momento, nosso *corpus* paralelo conta com 117 textos em português e inglês, com um total de 284.315 *tokens*²¹ em português e 273.259 em inglês e com 17.117 *types* em português e 12.732 *types* em inglês.

²⁰ “*to the extent that findings based on its contents can be generalized to a larger hypothetical corpus.*”

²¹ Enquanto *tokens* se referem ao número total de palavras em um *corpus*, *types* se referem às palavras diferentes.

3.2 Escolha das ferramentas

Após refletir sobre qual *software* atenderia as necessidades de nossa pesquisa, optamos por utilizar o programa ParaConc 1.0, que é um concordanciador bilíngue usado em análises contrastivas, desenvolvido por Michael Barlow²². O ParaConc possibilita gerar listas de frequência das palavras de texto alinhados, buscar por palavras específicas, além da possibilidade de alinhar até quatro textos paralelos, que podem ser em quatro línguas diferentes, ou um texto original mais três traduções diferentes. Apesar de sua interface pouco amigável e intuitiva, o ParaConc é uma das poucas ferramentas disponíveis que permite salvar o texto alinhado, além de oferecer a opção de salvar o andamento do trabalho (*workspace*). Ademais, devido à falta de atualizações que visem melhorar o programa, passamos por inúmeras dificuldades e a necessidade de refazer boa parte dos alinhamentos, pois quando o arquivo alinhado é exportado e salvo, algumas frases acabam perdendo seu alinhamento no arquivo de texto resultante. Sua utilização requer um alto grau de paciência e atenção redobrada por parte do usuário.

Já os *corpora* comparáveis serão processados com a ferramenta AntConc, uma ferramenta de livre acesso, desenvolvida por Laurence Anthony²³. Essa ferramenta, em particular, permite ao pesquisador: i) gerar listas de palavras, em ordem de frequência, de todas as palavras, de todas as palavras que constam nos arquivos de textos selecionados; ii) gerar lista de clusters/N-grams, que permite observar como determinado termo se combina com outras palavras do *corpus* e iii) gerar lista de palavras-chave (Keyword List), resultante da comparação entre a lista de frequência das palavras do *corpus* de estudo com a do *corpus* de referência, um *corpus* maior que o *corpus* de estudo, como é o caso do BNC (British National Corpus), uma coleção de 100 milhões de palavras em inglês britânico que inclui tanto a língua escrita como a falada. Tal comparação permite que o pesquisador identifique palavras que são estatisticamente mais frequentes no *corpus* de estudo. Quanto maior for a frequência estatística de uma palavra no *corpus* de estudo, maior será sua especificidade.

3.3 Limpeza dos arquivos, catalogação e primeiras extrações

Uma vez que selecionamos os volumes da *Revista Brasileira de Reumatologia* que disponibilizam os artigos em língua portuguesa e sua tradução em língua inglesa, esses foram salvos em formato arquivo de texto (*.txt). Após diversos testes com o ParaConc, utilizamos a codificação ANSI para os arquivos do *corpus* paralelo, posto que a ferramenta funciona melhor com essa codificação (já para o

²² Disponível em: <<http://www.paraconc.com/>>.

²³ Disponível em: <<http://www.laurenceanthony.net/software/antconc/>>.

corpus comparável, será utilizada a codificação UTF-8. O AntConc aceita diversas codificações, a UTF8 é a mais apropriada para línguas com muitos diacríticos, como é o caso do português. O primeiro passo foi remover dos artigos todos os elementos extratextuais, como tabelas, gráficos, imagens, agradecimentos e *links* externos. Elementos como sinais de porcentagem e abreviações latinas (e.g. et al. e apud), que entraram em conflito com a ferramenta ParaConc ou que geraram ruído nos resultados, também tiveram que ser removidos.

Em caso de necessidade de recuperação de informação, embora tenhamos optado por não usar o cabeçalho, estabelecemos o sistema de catalogação, exemplo do qual mostramos na Tabela 1, abaixo:

Tabela 1 – Catalogação de arquivos

| Código | Autor, Título, Fonte | Disponível em |
|---------------|--|---|
| pt2010-01-01P | Elisabeth Gonzaga Canova Fernandes; Vanessa Ramos Guissa; Cynthia Saviolli; José Tadeu Tesseroli Siqueira; Marcelo Valente; Clovis Artur Almeida da Silva. Osteonecrose de mandíbula em pacientes com lúpus eritematoso sistêmico juvenil observada em exame de imagem. <i>Rev. Bras. Reumatol. vol.50 no.1 São Paulo jan./ fev. 2010</i> | http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0482-50042010000100002&lng=pt&nrm=iso&tlng=pt (acesso em 20 Mar. 2017) |
| en2010-01-01P | Elisabeth Gonzaga Canova Fernandes; Vanessa Ramos Guissa; Cynthia Saviolli; José Tadeu Tesseroli Siqueira; Marcelo Valente; Clovis Artur Almeida da Silva. Osteonecrosis of the jaw on imaging exams of patients with juvenile systemic lupus erythematosus. <i>Rev. Bras. Reumatol. vol.50 no.1 São Paulo jan./ fev. 2010</i> | http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0482-50042010000100002&lng=en&nrm=iso&tlng=en (acesso em 20 Mar. 2017) |

Fonte: Elaborada pelas autoras

Em relação à busca por abreviaturas e acrônimos, basta que se carreguem todos os arquivos alinhados na ferramenta ParaConc para se gerar a lista de frequência das palavras, por ordem alfabética ou por frequência. O ParaConc ainda oferece opções para refinar a lista gerada, como a opção que permite ignorar as palavras maiúsculas. Para esta pesquisa, porém, optamos por não ignorá-las simplesmente porque as abreviaturas e acrônimos aparecem notadamente em letras maiúsculas.

A lista de frequência gerada nos permite procurar candidatos a abreviaturas e acrônimos. Com esses candidatos em mãos, pode-se gerar uma lista específica (por meio da opção “search” da ferramenta), para observar esses elementos em seus contextos de uso nos pares linguísticos trabalhados. A partir dessa lista, pode-se também verificar as formas por extenso das abreviaturas e acrônimos, como é possível observar na Figura 1 o exemplo da abreviatura LES (Lúpus Eritematoso Sistêmico).

No caso das abreviaturas e acrônimos, interessa-nos extrair as formas por extenso, para que também possamos identificar como são traduzidas (*corpus paralelo*), e seu padrão de uso (*corpora comparáveis*). O conhecimento dessa relação entre abreviatura/acrônimo e sua forma por extenso poderá auxiliar tradutores e pesquisadores a escolherem colocações já consolidadas no campo.

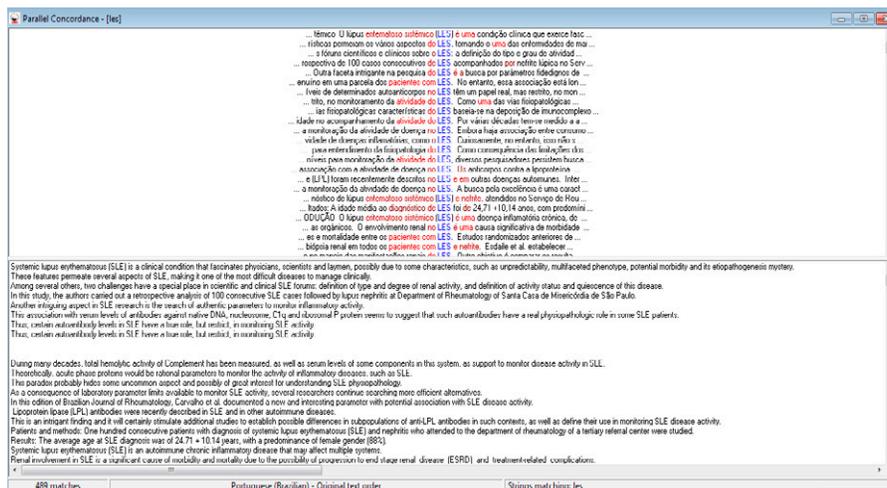


Figura 1 – Opção “search” do programa ParaConc

Fonte: Elaborada pelas autoras

4 Resultados preliminares

Como a pesquisa se encontra em estágio de compilação dos *corpora*, este trabalho não visa a apresentar resultados conclusivos. Contudo, de uma lista de frequência das cinquenta primeiras ocorrências de abreviaturas e acrônimos (a Tabela 2 mostra as cinco primeiras mais frequentes), já foi possível observar que há um número expressivo de abreviaturas que seguem os padrões das línguas em questão (54%), o que refuta em parte a sugestão de Aixelá (2009) de que haja certa disposição em manter abreviaturas inalteradas em textos traduzidos do inglês para preservar o caráter de determinada disciplina e também para facilitar a divulgação do conhecimento técnico-científico, que, como já mencionado, é feita sobretudo em língua inglesa. Em nosso *corpus paralelo*, por exemplo, a abreviatura LES/SLE (Lúpus eritematoso sistêmico / *Systemic lupus erythematosus*) aparece 386 vezes em português e 378 vezes em inglês, sendo a abreviatura mais frequente em nosso *corpus*. Já a segunda abreviatura mais frequente, AR/ RA (*Artrite Reumatoide/ Rheumatoid arthritis*), aparece 119 vezes em português e 128 em inglês. Em seguida, temos ES/SSc (*Esclerose Sistêmica / Systemic Sclerosis*), PCR/CRP (Proteína

C-reativa / *C-reactive protein*) e LESJ/JSLE (Lúpus Eritematoso Sistêmico/Juvenil / *Juvenile systemic lupus erythematosus*), que aparecem, respectivamente, 117 vezes em português e 95 em inglês; 101 vezes em português e 72 em inglês e 97 vezes em português e 103 em inglês .

Tabela 2 – As 5 abreviaturas/acrônimos traduzidos mais frequentes

| | Freq | PT | Forma por Extenso | Tradução | Forma por Extenso | Freq |
|---|------|------|-------------------------------------|----------|---------------------------------------|------|
| 1 | 386 | LES | Lúpus eritematoso sistêmico | SLE | Systemic lupus erythematosus | 378 |
| 2 | 119 | AR | Artrite Reumatoide | RA | Rheumatoid arthritis | 128 |
| 3 | 117 | ES | Esclerose Sistêmica | SSc | Systemic Sclerosis | 95 |
| 4 | 101 | PCR | Proteína C-reativa | CRP | C-reactive protein | 72 |
| 5 | 97 | LESJ | Lúpus Eritematoso Sistêmico Juvenil | JSLE | Juvenile systemic lupus erythematosus | 103 |

Fonte: Elaborada pelas autoras

Em relação às abreviaturas e acrônimos que são iguais nas duas línguas (46%), é interessante notar que, apesar de esses elementos terem se mantido iguais, as formas por extenso de três das cinco ocorrências mais frequentes (Tabela 3) seguem suas respectivas línguas, sendo que as outras duas, NK (*Natural Killer Cell / célula Exterminadora Natural*) e IgM (*Immunoglobulin M / Imunoglobulina M*) aparecem no *corpus* sem as formas por extenso (s/n).

Tabela 3 – As 5 abreviaturas/acrônimos inalterados mais frequentes

| | | | | | | |
|----|----|------------|-------------------------------|------------|---|----|
| 9 | 78 | PRL | Proclatina | PRL | Proclatin | 76 |
| 10 | 71 | NK | s/n | NK | s/n | 72 |
| 12 | 60 | IgM | s/n | IgM | s/n | 61 |
| 14 | 54 | (anti-)LPL | Lipoproteína lipase | (anti-)LPL | Lipoprotein lipase | 47 |
| 15 | 52 | SLEDAI | Índice de atividade da doença | SLEDAI | Systemic Lupus Erythematosus Disease Activity Index | 48 |

Fonte: Elaborada pelas autoras

Em uma análise parcial dessas primeiras 50 ocorrências, parece-nos que há uma tendência em usar as abreviaturas e acrônimos nas respectivas línguas quando esses se referem a denominações de doenças (e.g. LES, AR, ES e LESJ), enquanto permanecem inalterados quando se referem a elementos como hormônios (e.g. PRL), proteínas (e.g. IgM, LPL), escores mensuradores de atividade de doença (e.g. SLEDAI), vírus, como é o caso de HIV (vírus da imunodeficiência humana) e HPV (papilomavírus humano), respectivamente nas posições 22^o e 43^o do *corpus*, entre outros. Contudo, tal conjectura poderá somente ser confirmada com a expansão do *corpus*, que pretendemos ampliar para a marca de 500 mil palavras.

5 Considerações finais

Embora possamos observar um aumento gradual no número de pesquisas acadêmicas com textos médicos na área dos Estudos da Tradução, ainda é inexpressiva a quantidade de pesquisas que tenham como foco a tradução de abreviaturas e acrônimos. A partir da prática tradutória, tivemos a oportunidade de identificar esses elementos como um problema de tradução, visto não haver padronização na maneira de traduzi-los. Como o conhecimento científico é normalmente divulgado em língua inglesa, vem se tornando cada vez mais necessário que pesquisadores falantes de outras línguas, incluindo a portuguesa, submetam artigos nesse idioma para aumentar as chances de publicação. Com base nesse cenário, deu-se início à pesquisa aqui descrita, a fim de verificar padrões tradutórios e de uso desses elementos através da compilação de um *corpus* paralelo e dois *corpora* comparáveis.

Ainda que haja certa disposição em manter abreviaturas nas formas usadas e consolidadas em língua inglesa (AIXELÁ, 2009), observamos que um número considerável desses elementos são usados em suas respectivas línguas. Ainda que estejamos nas primeiras etapas do estudo, esses resultados já nos encorajam a nos perguntar se a tradução ou manutenção dessas abreviaturas e acrônimos estaria relacionada a algum elemento em particular, uma vez que observamos que há um número expressivo de abreviaturas e acrônimos traduzidos que se referem a nomes de doenças, por exemplo. Porém, estaremos em melhor posição para responder a essa e a outras perguntas que possam surgir ao longo do estudo quando completarmos a coleta e análise dos textos.

Uma vez concluído o processamento dos *corpora*, nosso objetivo é criar um glossário *on-line* nos pares linguísticos português-inglês e inglês-português para que seja disponibilizado a todos os interessados na tradução de abreviaturas e acrônimos em reumatologia. Não se objetiva criar um glossário definitório, visto que o problema detectado está sobretudo na questão de saber se esses elementos são ou não traduzidos, assim, ele será composto das abreviaturas e acrônimos seguidos das formas por extenso. Uma vantagem adicional de se criar um glossário desse tipo é ter esses elementos convenientemente em um único *site*, uma vez que a procura em diferentes fontes consome considerável tempo do tradutor. Finalmente, uma vez que todo o processo esteja concluído, nosso intuito é que os *corpora* da subárea da reumatologia acompanhem os *corpora* existentes na plataforma do Termisul²⁴ da Universidade Federal do Rio Grande do Sul. Os *corpora* de outras subáreas da saúde (cardiologia, pediatria, enfermagem) e de química do Termisul já se encontram disponíveis para consulta.

²⁴ Disponível em: <<http://www.ufrgs.br/termisul/>>.

Referências

- AZENHA, J. Jr. *Tradução Técnica e Condicionantes Culturais*. São Paulo: Humanitas, 1999.
- _____. Tradução técnica, condicionantes culturais e os limites da responsabilidade do tradutor. *Cadernos de Tradução*, Florianópolis, v. 1, n. 1, p. 137-148, 1996.
- BAKER, M. *Corpora in Translation Studies*. An overview and suggestions for future research. *Target*, 7(2), p. 223-243, 1995.
- _____. *Corpus Linguistics and Translation Studies: implications and applications*. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (Org.). *Text and Technology: in honour of John Sinclair*. Amsterdam: John Benjamins, 1993, p. 233-250.
- BAKER, P. Sketching Muslims: a *corpus* driven analysis of representations around the word 'Muslim' in the British press 1998–2009. *Applied Linguistics*, v. 34, n. 3, p. 255-278, 2013.
- BELDA, J. R. M. Translating Computer Abbreviations from English into Spanish: main types and problems. *Meta: Translators' Journal*, v. 49, n. 4, p. 920-929, 2004.
- BERBER SARDINHA, A. P. *Linguística de Corpus*. Barueri: Manola, 2004.
- _____. *Corpora Eletrônicos na Pesquisa em Tradução*. Florianópolis: *Cadernos de Tradução*, v. 9, n. 1, p.15-59, 2002.
- BIBER, D. Representativeness in *corpus* design. *Literary and Linguistic Computing*, v. 5, n. 4, p.243-257, 1993.
- COLLET, Thaís. *Procedimentos tradutórios na legendagem de house: análise da terminologia médica referente a exames e aparelhos*. Dissertação (mestrado em Estudos da Tradução). Programa de Pós-Graduação em Estudos da Tradução, Universidade Federal de Santa Catarina, Florianópolis, 2012.
- COULTHARD, R. J. *The application of corpus methodology to translation: the JPED parallel corpus and the pediatrics comparable corpus*. Dissertação (mestrado). Programa de Pós-Graduação em Estudos da Tradução, Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2005.
- FRANCO AIXELÁ, J. An overview of interference in scientific and technical translation. *The Journal of Specialised Translation*, n. 11, p. 75-88, 2009.
- _____. The Study of Technical and Scientific Translation: An Examination of its Historical Development. *JoSTrans - The Journal of Specialised Translation*, v. 1, n. 1, 2004.
- MUNDAY, J. (Org.). *The Routledge Companion to Translation Studies*. Revised Edition. Oxon: Routledge, 2009.
- OLOHAN, M. *Introducing Corpora in Translation Studies*. London: Routledge, 2004.
- PAIVA, P. T. P. *Uma investigação de traduções de textos da área médica sob a luz dos estudos da tradução baseados em corpus*. Tese (doutorado). Instituto de Biociências, Letras e Ciências Exatas, UNESP, São José do Rio Preto, 2009.
- PASQUALI, A. B.; PINTO, P. T. A tradução de resumos médicos como meio de aprendizagem do processo tradutório e da terminologia especializada. *Caminhos em Linguística Aplicada*, v. 9, n. 2, p. 25-49, 2013.
- SINCLAIR, J. *Corpus evidence in language description*. In: WICHMANN, A. S.; FLIGELSTONE, S.; MCENERY, T.; KNOWLES, G. (Ed.). *Teaching and language corpora*. London: Longman, 1997, p. 27-39.
- TYMOCZKO, M. Computerized *Corpora* and the Future of Translation Studies. *Meta*, v. 43, n. 4, p. 652-660, 1998.