

# ESTUDOS DE TERMINOLOGIA NO BRASIL: DIÁLOGOS COM PORTUGAL

MARIA JOSÉ BOCORNY FINATTO

## ABSTRACT

This paper contrasts some texts that deal with the history of the terminology research in Brazil, especially the research aimed to extract or recognize terminology in corpora, a widespread practice among us only from the year 2000, with pioneering texts on this topic produced by Portuguese researchers, represented here by Belinda Maia. The intention is to recognize her role as disseminator of the corpus-based methodologies. The paper follows showing how the dialogue between the Terminology studies from Brazil and Portugal is important for the promotion of Portuguese language in the global scenario of scientific and technical communication.

## [1] INTRODUÇÃO

Este texto trata de cotejar algumas publicações que servem como um exemplo de testemunhos da trajetória da pesquisa terminológica no Brasil, especialmente a pesquisa orientada para a extração ou reconhecimento de terminologias a partir de corpora, método de trabalho que só foi nacionalmente disseminado entre nós a partir dos anos 2.000, e textos pioneiros sobre esse modo de pesquisa apresentados por estudiosos portugueses, representados aqui por Belinda Maia. A intenção deste trabalho, assim, é prestar o devido reconhecimento ao papel de Belinda Maia como disseminadora da ideia do trabalho com corpus, quando nossos primeiros estudos sobre Linguística de Corpus recém conseguiam alguma repercussão e reconhecimento no Brasil (Sardinha 2000). Desde então, essa metodologia de trabalho, que tem reunido no Brasil a Linguística de Corpus (LC) e o Processamento de Linguagem Natural (PLN), permanece como algo altamente desafiador, sobretudo entre a comunidade de pesquisadores linguistas que ainda hoje têm pouco contato com técnicas computacionais.

Do lado lusitano, revisamos dois textos de Maia (2003), “Using Corpora for Terminology Extraction: Pedagogical and computational approaches” produzido para um evento de 2001 (PALC) e Maia (2002), “Corpora for terminology extraction – the differing perspectives and objectives of researchers, teachers and language services providers” produzido para um evento de 2002 (LREC). Do lado brasileiro, é tratado um texto de minha autoria (Finatto 2003), publicado em um boletim da Associação Brasileira de Linguística (ABRALIN), no qual, juntamente com ou-

tras colegas da Terminologia e Terminografia brasileiras, apresentava o desafio de tratar manuais de Química Geral como *corpora* em busca de conexões entre suas terminologias e os gêneros textuais e discursivos envolvidos. Por fim, examino, de modo panorâmico, em contraponto com os textos de Belinda Maia e esse meu, o conjunto de ideias postas em um livro brasileiro bastante recente (Tagnin & Bevilacqua 2013) sobre o assunto, cujo título é justamente *Corpora na Terminologia*.

Ao final do texto, mostra-se como o diálogo entre os estudos de Terminologia do Brasil e de Portugal tem sido positivo para a promoção da Lusofonia no cenário global da comunicação científico-técnica. Nesse cenário, os repertórios terminológicos, sejam glossários, dicionários, bases de dados, mapas conceituais ou ontologias de domínio, gerados a partir de *corpora*, têm um destaque especial.

[2] A “NOVIDADE” DO TRABALHO COM CORPUS E DAS FERRAMENTAS COMPUTACIONAIS

Para a construção de um *corpus*, bem sabemos, há todo um conjunto de procedimentos, bastante penosos, mas ao final muito gratificantes, de modo que o acervo, criteriosamente reunido, realmente sirva para representar, com segurança, um dado estado de uso de língua. Maia (1997) já nos apresentava um texto sobre como se poderia enfrentar essa tarefa de um modo relativamente tranquilo, produtivo e colaborativo, reunindo-se esforços de diferentes pessoas que tivessem interesses de pesquisa semelhantes em torno desse trabalho.

Mais tarde, em 2004, conforme apontavam Maia et al. (2004, pg. 45), em um texto que tratava justamente da cooperação entre brasileiros e portugueses em torno de *corpora* — para ensino, ensino de tradução, tradução e pesquisa de terminologias, uma vez construído um *corpus*, definido, *grosso modo*, como uma coleção de textos em formato digital, sendo ele etiquetado ou em ‘cru’, havia toda uma parte de ferramentas para observar e analisar o uso da língua nesse conjunto de textos. Essas ferramentas, como víamos àquela época no Brasil, pareciam uma “mágica”. Afinal, elas permitiam a observação de muitos dados ao mesmo tempo, em vez do “antigo” — mas familiar — trabalho de se ler uma mesma ocorrência de palavra ou de uma dada expressão linha a linha ao longo de um texto ou de vários textos disponíveis apenas em formato impresso. Eu mesma, em 1998, ainda examinava os textos das mais de centenas de leis brasileiras sobre o meio ambiente desse modo, com lápis, caneta sinalizadora colorida e papel, para a produção de um dicionário da sua terminologia.

Pois aquelas ferramentas computacionais “mágicas” já ofereciam informação de natureza estatística, que poderia, posteriormente, ser analisada para fins específicos. E, conforme ensinavam Maia et al. (2004), os grandes *corpora* monolíngues traziam a possibilidade de se estudar a língua no nível lexical e sintático, o que tenderia a auxiliar imensamente quem se interessasse, por exemplo, por identificar terminologias em grandes acervos de textos científicos ou técnicos.

Infelizmente, ainda hoje no Brasil, em 2015, muitos colegas linguistas desconhecem quaisquer metodologias do trabalho com *corpus*, embora estejamos em uma época de grande informatização, quando nem mesmo se precisa mais comprar algum *software* para realizar o papel das ferramentas. Afinal, não faltam opções gratuitas e ferramentas prontas para uso *on-line*. Conforme já afirmamos em trabalho recente (Novodvorski & Finatto 2014), a LC no Brasil — e por extensão os trabalhos com *corpora* e com ferramentas para sua exploração — associou-se a diferentes aventuras de investigação e praticamente nada rejeitou em termos de parcerias de trabalho. O diálogo tem sido uma marca constante, mesmo com quem conceba a LC “apenas” como um *modus operandi* computacional e quantitativo. A despeito dessa impressão, deve ter ficado claro, pelo menos nesses primeiros 10 anos de percurso do trabalho com *corpora* no Brasil — comemorados em 2015 com uma década da publicação do artigo de Sardinha (2000) —, que fomos muito além de “apenas contar palavras”.

### [3] RECONHECIMENTO DE TERMINOLOGIAS EM CORPORA

Conforme já assinalava Maia (2002), para além do trabalho da padronização de termos — que estabelece os termos “corretos” de uma dada especialidade, levado a cabo por especialistas de um domínio, geralmente sem a colaboração de linguistas, há todo um outro trabalho terminológico, que inclui uma parte de recuperação de informação e de observação de descrição de usos além dos usos padronizados. Esses inventários de termos padronizados geralmente ficam registrados em dicionários e glossários “oficiais”, publicados como terminologia padronizada e preconizada por associações de classe, visando uniformizar denominações e conceituações que lhes pareçam mais adequadas. Desse modo, trata-se de normatizar usos.

Assim, num outro cenário, descritivo e não prescritivo, são estudados os diferentes modos e circunstâncias de emprego de termos — o foco é o de descrever usos (quaisquer usos), independentemente de serem mais ou menos normatizados. O trabalho descritivo, feito dessa forma, é geralmente levado a cabo por linguistas, tradutores, professores que lidam com ensino de línguas para propósitos específicos. Esse seria o caso, por exemplo, do ensino de escrita científica em inglês para médicos, ou a situação de profissionais que lidam com recuperação de informação e representação de conhecimento — como cientistas da informação e informatas que trabalham com Processamento de Linguagem Natural (PLN). Nessa situação, são identificadas todas as realizações encontradas em diferentes *corpora*: os termos corretos, incorretos, desviantes, aceitáveis, toleráveis, “adaptados” (entre outros), desde que empregados em textos produzidos por pesquisadores de diferentes perfis. São observados também e especialmente os termos presentes em diferentes cenários comunicativos, incluindo a comunicação científica, técnica ou tecnológica dirigida para leigos e semileigos — entrando aqui o sempre pungente tema da popularização ou vulgarização de temas científicos.

Desse modo, já que a variação terminológica é um fenômeno natural da língua em uso, coloca-se a necessidade de se organizar *corpora* de diferentes perfis para o reconhecimento de terminologias, pois há muita coisa envolvida, conforme Belinda, magnificamente, já nos ensinava e resumia (Maia 2002, pg. 2, tradução minha<sup>1</sup>):

*[O trabalho da] Terminologia não é a simples acumulação de palavras, seus equivalentes em outras línguas, definições e uma certa quantidade de informação gramatical. Também não é a simples correspondência de um termo para um conceito. A pessoa tem que lidar com todos os problemas usuais da linguagem - sociais, geográficos, históricos, políticos, e outros aspectos de estilo e de registro. Ao nível da padronização, a pessoa pode até mesmo se envolver em verdadeiras batalhas entre acadêmicos ou sociedades comerciais que querem ver prevalecer as palavras que eles usam para descrever suas teorias ou produtos específicos.*

Esse trabalho descritivo, como normalmente é visto e produzido por linguistas, abastece recursos voltados para o uso de tradutores. Mas, claro, há um número sem fim de finalidades para ele. Além disso, ainda que se possa argumentar que a variação ou variabilidade de formas e de significados dos termos seja um entrave para um ideal de comunicação científica unívoca, sem ambiguidades, há que se descrever a situação, reconhecendo o que ocorre para que se possa, inclusive, melhor administrar uma situação de padronização.

Em um outro trabalho, que corresponde a uma comunicação apresentada em 2001, Maia (2003) discutia sobre o uso de *corpora* menores para a extração de terminologias, trazendo-nos uma abordagem pedagógica e computacional. Na parte pedagógica, estava a formação de tradutores que se envolviam, sob a supervisão de seus professores, com o processo de compilar e explorar os seus *corpora*; e, na parte computacional, aquilo que ela denominou de “terminologia computacional”, o que já havia mencionado no seu texto de 2002 antes referido.

O relato da parte pedagógica mostrava que Belinda então se concentrava em produzir e usar pequenos *corpora* para estudar assuntos especializados com mais profundidade e para a extração de termos (cf. *ibid.*, pg. 1). Dessa parte, é preciso salientar a sua seguinte frase aqui reproduzida no original: “*One of the principal objectives of our work is to convince the humanities educated translator that the serious study of specialised language – from the text to the term – is essential training for future translators.*” (pg. 2, grifo nosso). O grifo, conforme entendo, reflete uma visão

[1] No original: “Terminology is not the simple accumulation of words, their equivalents in other languages, definitions and a certain amount of grammatical information. Nor is it the simple matching of term to concept. One has to deal with all the usual problems of language - social, geographical, historical, political, and other aspects of style and register. At the level of standardisation, one can even become involved in authentic battles between academics or commercial companies who want to see the words they use to describe their particular theories or products prevail.”

pioneira da perspectiva textual do estudo e do trabalho com as terminologias, com a qual nos temos alinhado há bastante tempo, o que se confirma nesses trechos a seguir apresentados de um trabalho nosso produzido em 2001 (Finatto 2003). Nele já apresentávamos uma proposta diferenciada — pelo menos no âmbito dos estudos de Terminologia do Brasil — para o enfoque linguístico-terminológico de manuais acadêmicos de Química (Finatto 2003, pg. 186):

*O texto especializado (também chamado texto técnico-científico) é fruto da ação perceptiva e transformadora de um sujeito enunciador, individual e múltiplo, sobre um conjunto de conhecimentos e textos com os quais se relaciona. Essa ação, que pode ser vista como um redizer algo ou o recontar a estruturação de um conhecimento tornando-o acessível ao outro, está materializada e sobremodalizada (ou particularizada) no amplo conjunto de enunciados que estabelece o texto e também envolve relações de intertextualidade. Isto é, o texto do Manual de Química, nosso objeto de observação particular, é construído pelo enunciador e é, igualmente, uma síntese de diversos outros textos: os que o precedem e o acompanham, compartilhando com ele a constituição sócio-histórica de uma área de conhecimentos, de um continuum de conhecimentos que é a Química.*

Nesse ponto, ainda vale destacar nossa outra afirmação, colocada naquele trabalho de 2001 (cf. *ibid*, pg. 186) e que sustentamos como válida até hoje, de que não

*há como desvincular termos e textos, mesmo que se escolha privilegiar os aspectos lexicais mais pontuais de uma terminologia. São, assim, esses referenciais elementos muito importantes e que certamente podem contribuir para os novos desenvolvimentos de uma teoria Terminológica de base comunicativa voltada à compreensão do texto técnico-científico em seus diferentes tipos.*

O artigo de Maia (2003) trazia já, além dessa “postura textual”, uma reflexão sobre o português de Portugal e o do Brasil no âmbito da comunicação científica posta em textos. Trazia ela um relato, cuja leitura é indispensável para professores de Tradução ou de Terminologia do Brasil, sobre como conduzia, com alunos de um curso de Tradução em Portugal, as relações travadas com os especialistas do domínio em foco e a formação desses futuros profissionais frente ao “atravesamento” do Português do Brasil. Afinal, ao buscarem seus *corpora* em português, os alunos via de regra deparavam-se com fontes brasileiras e lusitanas.

Conforme explicava Maia (2003, pg. 10), muitas vezes textos em português simplesmente não existiam ou não eram encontráveis para compor os *corpora* dos estudantes. Mas, em geral, quando eram encontrados, estavam em Português do Brasil, algo a que as pessoas de Portugal muitas vezes reagem negativamente

— especialmente os especialistas do domínio e também os estudantes, conforme pude supor. A estrutura do vocabulário e texto do Português do Brasil — particularmente em domínios científicos e técnicos — conforme relatava — sofria de uma influência considerável do inglês norte-americano o que não ocorria do mesmo modo no português de Portugal. No material brasileiro, termos em inglês eram muitas vezes adotados e a estrutura de sentença em tais textos tendia a seguir a ordem SVO dominante de Inglês em detrimento do estilo vernáculo.

Anos mais tarde, em 2010, experimentaríamos, *in vivo*, a situação relatada por Belinda, mas pela via inversa, em relação a *corpora* de um domínio disponíveis apenas em português de Portugal. Com estudantes de Tradução da nossa Universidade, alunos de graduação e de pós-graduação, acompanhados por colegas médicos associados à nossa Universidade, colaboramos em um trabalho sobre o tema das pneumopatias ocupacionais (Brasil)/profissionais (Portugal) integrado por uma equipe brasileira e uma equipe portuguesa, o que registramos nos trabalhos [Marcolin et al. \(2010\)](#) e em [Finatto & Goldnadel \(2013\)](#).

Mais uma vez, ter tido acesso ao relato de experiências de Belinda Maia permitiu que nos entendêssemos e nos situássemos em um enquadramento de uma situação geográfica e temporalmente diferente, mas muito semelhante. Havia, enfim, o desafio de uma lusofonia a enfrentar e a promover com nossos alunos e com os especialistas do domínio. Esse percurso, ainda que parcialmente registrado, pode ser conferido ao acessar-se o nosso *Glossário Experimental de Pneumopatias Ocupacionais*.<sup>2</sup> Esse é um protótipo de glossário brasileiro, fruto do trabalho comum com os colegas de Portugal, que estamos construindo ano a ano com nossos estudantes da disciplina de Terminologia do curso de Tradução. Sua proposta é apenas a de ser um material didático para nossos alunos de Letras.

Quanto ao modo como lidar com essa lusofonia em meio à pedagogia de terminologias para futuros tradutores, nossa opção tem sido a de separar as variantes de um modo bastante claro. Recomendamos agrupar os dados em blocos distintos, ainda que esses blocos estejam compartilhados em um mesmo trabalho, de modo a salvaguardar as diferenças importantes entre usos terminológicos e práticas textuais entre Brasil e Portugal. Incluir essas diferenças ou separar as informações, em um *corpus* ou em um glossário, pode incomodar algumas pessoas; mas o incômodo também demonstra a necessidade de se pensar em um “português internacional”.

De nossa parte, alertamos que é um perigo, para o tradutor brasileiro que produz um texto em português para clientes brasileiros, utilizar uma fonte portuguesa sem dar-se conta de sua procedência e das potenciais diferenças terminológicas. Vale mencionar alguns exemplos bastante ilustrativos de diferenças terminológicas entre Brasil (BR) e Portugal (PT): **AIDS** (BR) e **SIDA** (PT), **ácidos**

[2] Disponível gratuitamente em: <http://www.ufrgs.br/textecc/pneumopatias/index.php>.

**graxos** (BR) e **ácidos gordos** (PT), **disbarismo** (BR) e **embolia gasosa** (PT), entre outros casos concretos que se poderia conferir, por exemplo, no *Glossário Panlatino de Pneumopatas Ocupacionais/Profissionais*.<sup>3</sup>

Em síntese, recomendamos que *corpora* portuguesas sejam tratados em separado dos *corpora* brasileiros, sempre muito bem identificados, especialmente se o uso dos dados extraídos servir para abastecer produtos para a tradução. Há muitos pontos coincidentes, naturalmente, mas as diferenças não se pode ignorar, tampouco essas diferenças, devidamente repertoriadas, devem servir para que se possa pensar na inviabilidade de se escrever também em português o conhecimento científico e técnico. Em um trabalho que reúna as fontes e as terminologias dos dois países, as denominações comuns ficariam marcando, assim, um português internacional, ao passo que se assinalam, sempre, os usos diferentes de Portugal (PT) ou do Brasil (BR).

Uma tal postura, é o que discute Santos (2014) quando trata da questão de diferenças linguísticas entre Portugal e Brasil em seu trabalho intitulado *Como estudar variantes do português e, ao mesmo tempo, construir um português internacional?* A autora entende que é importante termos *corpora* em português sem uma separação de variantes considerando a ideia de um português internacional. No contexto dessa pergunta, pelo menos no âmbito da Terminologia, entendo que importa descrever essas variantes e pô-las em contato, em conjunto, ainda que individualizadas, de modo que todos saibamos uns dos outros e de seus usos específicos.

Sob uma outra ótica, igualmente interessante para uma reflexão que abarca e extrapola o trabalho de Santos (2014), temos o estudo de Coulthard (2005). Nesse trabalho, o autor já aponta, com base em um extenso estudo em *corpus*, uma influência do estilo redacional do artigo em inglês sobre a escrita original de artigos em português por parte de pesquisadores de Pediatria no Brasil. Assim, o estilo em português brasileiro, pelo menos em artigos científicos de Pediatria, já aparece “anglofonizado”, talvez até para facilitar a tradução do texto para um inglês *lingua franca*. O *corpus* paralelo de Pediatria reunido por Coulthard (2005), por nós expandido, encontra-se disponível para consulta, em diferentes formatos e recursos.<sup>4</sup>

[4] CORPORA NA TERMINOLOGIA E NAS TERMINOLOGIAS

Em 2013, organizado por Tagnin & Bevilacqua (2013), foi publicada no Brasil uma coletânea de artigos que servem, em tese, para atestar a boa junção e o sucesso do trabalho terminológico baseado ou guiado por *corpora*. O objetivo da obra é, na verdade, reiterar, para nós do Brasil, que já há uma interface produtiva e pro-

[3] Disponível gratuitamente em [http://www.oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/panlatin\\_pneumopathies20130124.pdf](http://www.oqlf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/panlatin_pneumopathies20130124.pdf).

[4] Consultar em <http://www.ufrgs.br/textecc/textped/Dicionarios/DicPed/>.

missora entre Linguística de Corpus (LC) e Terminologia. Seus textos podem ser divididos em três grupos de artigos que tratam: a) da trajetória histórica dessa inter-relação; b) de situar essa junção no cenário da formação de tradutores; e, c) de experiências de elaboração de produtos terminográficos para tradutores e redatores.

Entre vários textos que poderia citar, um trabalho digno de referência aqui, visando ilustrar a riqueza de perspectivas da obra, é o de [Maciel \(2013\)](#). Isso porque a autora frisa que, embora a LC goze de prestígio no cenário da pesquisa terminológica brasileira atual, há muitos trabalhos que, embora lidem com *corpora*, não seguem seus princípios teórico-metodológicos. Mais, ainda, ela nos alerta que ([op. cit.](#), pg. 40):

*O uso de corpora em formato digital abre múltiplos caminhos e aqueles que se dedicam aos estudos terminológicos não conseguem prescindir deles. No entanto, não raras vezes, o pesquisador é levado a acreditar em falácias e a sonhar com o impossível. Engana-se com o poder da máquina, busca a mágica de um aplicativo automático, ilude-se com a velocidade operacional prometida, entrega-se a cálculos matemáticos e, em clima de entusiasmo pela evidência empírica, passa ao largo da reflexão, do estudo e dos fundamentos teóricos do seu trabalho.*

A partir disso, é importante não perder de vista que a informatização, mediada pela LC, mostra para a pesquisa em Terminologia que a melhor observação da linguagem é, sim, extensiva. É evidente que é preciso contrastar padrões da linguagem científica ou técnica em diferentes *corpora*, pois, pela lente de aumento dos *corpora*, evidenciam-se as características dos textos e dos discursos ([Finatto 2007](#)). Ainda assim, pelo menos no Brasil, conforme percebo pessoalmente, a pesquisa em Processamento da Linguagem Natural (PLN) não tem dialogado tão intensivamente quanto poderia com a pesquisa em LC, visto que a primeira é território de informatas e a segunda, de linguistas. Apesar dessa percepção negativa, uma boa integração entre essas duas áreas temos visto em alguns grupos do Brasil e de Portugal, como o Núcleo Interinstitucional de Linguística Computacional (NILC) de São Carlos (São Paulo) e a Liguatca (citados em [Maia et al. \(2004\)](#)), que reúnem linguistas e informatas. Mas, pelo menos no âmbito acadêmico brasileiro, há ainda muito a se avançar rumo a um diálogo produtivo e igualmente positivo para as partes nele envolvidas.

#### [5] CONSIDERAÇÕES FINAIS

É claro, hoje, entre nós linguistas brasileiros que lidamos com Terminologia e Terminografia, que um termo especializado é, antes de tudo, um valor ativado no discurso (o termo *discurso*, para mim, não é exatamente um sinônimo de texto, mas não cabe aqui essa discussão). Essa concepção devemos principalmente à Teoria



Comunicativa da Terminologia (TCT), desenvolvida por Maria Teresa Cabré e seus colaboradores na Espanha.

Em função desse entendimento de um valor terminológico ativado no discurso/texto, identificar uma terminologia passa a ser uma tarefa que envolve também reconhecer um texto, um modo de dizer, uma “família de textos”, elementos que se concretizam em diferentes usos de uma língua. Além dos textos, e por causa deles, temos que acessar todo um conjunto de itens de conhecimento inter-relacionados e apresentados social, cultural e historicamente em um dado espaço-tempo. Portanto, identificar os termos não representa “apenas” o reconhecimento de um “conjunto de palavras” de um domínio determinado, tampouco apenas listar as palavras “técnicas” que se “extraem” de um conjunto de textos.

Um termo, como sabemos, corresponde a um conceito, que é uma porção de conhecimento. E, tal como as palavras comuns, também os conhecimentos morrem, nascem ou caem em desuso. Afora isso, a sua utilização peculiar pode identificar grupos sociais, perspectivas nacionais ou internacionais de uma ciência e correntes de pensamento “locais”. Afinal, os produtores do pensamento de uma ciência, de uma tecnologia ou de uma especialidade expressam-se com escolhas lexicais próprias do seu discurso.

Por outro lado, no cenário da pesquisa sobre terminologias partindo de *corpora*, destacam-se hoje os *softwares* especialmente desenhados ou adaptados para a localização semi-automática de terminologias, de construções, de fraseologias, de sintagmas e de palavras ou de expressões que seriam potencialmente termos. Isso sem contarmos os novos recursos de PLN, hoje capazes de gerar automaticamente ontologias a partir de *corpora*, tal como, por exemplo, já faz a ferramenta *ExATOLp – Extrator Automático de Termos para Ontologias em Língua Portuguesa*. (Lopes 2012).

Neste momento atual, marcado pela intensa informatização do trabalho, considerando o sempre necessário e produtivo diálogo entre os estudos de Terminologia do Brasil e de Portugal, é importante refletirmos sobre o quanto *corpora* – gigantes ou pequenos – e ferramentas computacionais não são panaceias. Afinal, ainda há muito trabalho subjetivo envolvido. O apoio computacional e estatístico dá um suporte importantíssimo ao terminólogo/terminógrafo que segue lidando com acervos de conhecimento escritos em português. Embora a tarefa esteja hoje muito menos árdua com tal suporte tecnológico, o pesquisador ainda precisa trabalhar intensamente, sem perder de vista o cenário, algumas vezes belicoso, da prescrição *versus* descrição.

E, em meio a isso, permanecem os desafios para a promoção da lusofonia no cenário global da comunicação científico-técnica, de modo que possamos apreciar a nossa própria diversidade, dando-nos conta de que a ideia de um português internacional ou supranacional implica o reconhecimento de um patrimônio co-

num nosso (conforme bem assinala Santos 2014, pg. 21. Nesse cenário global, temos um português pan-nacional — como uma língua que abarca e subsume as suas diferentes variantes — frente a um inglês *lingua franca* ou frente a outras línguas que também têm suas variantes nacionais — como o francês da França ou do Canadá. Esse inglês *lingua franca*, entretanto, parece a tudo e a todos engolir, pois o que não estiver publicado em inglês, na prática, em termos de textos científicos ou técnicos, tende a não ser citado ou reconhecido, ignorado, mesmo que a tradução automática partindo do português para o inglês esteja à disposição de todos. Com o apoio da tradução automática, apesar dos seus erros mais comuns, poderia-se ao menos tentar entender os tópicos principais de um trabalho originalmente escrito em português. O problema é que, em geral, por vários motivos, sequer se tenta.

Será a influência maior ou menor do inglês uma medida de distanciamento entre as terminologias do Brasil e de Portugal e seus respectivos fazeres científicos? Eis aqui um bom tema para pesquisas com base em *corpora*. Mas, se isso for verdade, qual seria, então, a medida de aproximação entre elas? Como conduzir um trabalho terminológico que acolha e promova a lusofonia em uma época como a atual em que o Brasil se abre, intensivamente, à internacionalização e que parece dominar o cenário em muitas frentes? Creio que as respostas “certas” para essas perguntas não estejam dadas prontas em trabalhos de colegas portugueses como os de Belinda Maia aqui citados. Apesar disso, qualquer traço de resposta disponível, legitimamente construído em um percurso de reflexão, em meio à profusão de perguntas e de desafios a equacionar, já pode nos ajudar a encontrar as nossas — quem sabe — novas melhores perguntas. O diálogo renovado entre Brasil e Portugal, pelo menos no âmbito dos estudos de Terminologia, tende a nos ajudar nessa direção.

#### AGRADECIMENTOS

Agradeço a Diana Santos pela oportunidade de participar desta publicação e também à FAPERGS, à CAPES, no âmbito do Programa Stic-AmSud (projeto 047/2013), à FAPERGS e ao CNPq, instituições de apoio à pesquisa no Brasil, pelo apoio às minhas iniciativas de estudo de pesquisa.

#### REFERÊNCIAS

- Coulthard, Robert James. 2005. *The application of corpus methodology to translation: the jped parallel corpus and the pediatrics comparable corpus*: Universidade Federal de Santa Catarina. Tese de Mestrado.
- Finatto, Maria José Bocorny. 2003. Sobre o enfoque lingüístico-terminológico de manuais acadêmicos de Química Geral. Em Associação Brasileira de Linguística ABRALIN (ed.), *II congresso internacional da ABRALIN, 2001*, 184–186.

- Finatto, Maria José Bocorny. 2007. Exploração terminológica com apoio informatizado: perspectivas, desafios e limites. Em Aparecida Negri Isquerdo & Ieda Maria Alves (eds.), *As Ciências do Léxico. Lexicologia, Lexicografia, Terminologia. Volume III*, 447–458. Editora da UFMS/Humanitas.
- Finatto, Maria José Bocorny & Marcos Goldnadel. 2013. Formação de terminólogos: experiência com /corpus/ em uma graduação em tradução. Em Stella Tagnin & Cleci Bevilacqua (eds.), *Corpora na Terminologia*, 87–112. HUB Editorial.
- Lopes, Lucelene. 2012. *Extração automática de conceitos a partir de textos em língua portuguesa*: Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Tese de Doutorado.
- Maciel, Anna Maria Becker. 2013. Terminologia e corpus. Em Stella Tagnin & Cleci Regina Babilacqua (eds.), *Corpora na terminologia*, 29–45. HUB Editorial.
- Maia, Belinda. 1997. Do it yourself corpora... with a little bit of help from your friends! Em B. Lewandowska-Tomaszczyk & P. J. Melia (eds.), *Practical applications in language corpora*, 403–410. Lodz: Lodz University Press.
- Maia, Belinda. 2002. Do-it-yourself, disposable, specialised mini corpora - where next? Reflections on teaching translation and terminology through corpora. *Cadernos de Tradução* 1(9). 221–235.
- Maia, Belinda. 2003. Using Corpora for Terminology Extraction: Pedagogical and computational approaches. Em Barbara Lewandowska-Tomaszczyk (ed.), *PALC 2001: practical applications in language corpora*, 56–68. P. Lang.
- Maia, Belinda, Luís Sarmento, Stella E. O. Tagnin & Sandra Maria Aluísio. 2004. Idéias que cruzam o Oceano. *CROP - Revista da Área de Língua e Literatura Inglesa e Norte-Americana* 10. 43–64.
- Marcolin, Paula, Aline Evers, Maria José Bocorny Finatto & Marcos Goldnadel. 2010. Pneumopatologias: formação em terminologia em curso de tradução no Brasil. Em *Actas da RiTerm 2010*, 254–278.
- Novodvorski, Ariel & Maria José Bocorny Finatto. 2014. Linguística de Corpus no Brasil: uma aventura mais do que adequada. *Letras & Letras - UFU* 30(2). 7–16.
- Santos, Diana. 2014. Como estudar variantes do português e, ao mesmo tempo, construir um português internacional? <http://www.linguateca.pt/Diana/download/VariantesPIGSCP.pdf>.
- Sardinha, Tony Berber. 2000. Linguística de Corpus: histórico e problemática. *DELTA* 16(2). 323–367.

Tagnin, Stella Esther Ortweiler & Cleci Regina Bevilacqua. 2013. *Corpora na terminologia*. São Paulo: HUB Editorial.

CONTACTOS

Maria José Bocorny Finatto  
Instituto de Letras, Universidade Federal do Rio Grande do Sul, Brasil  
[maria.finatto@gmail.com](mailto:maria.finatto@gmail.com)