

Linguística de Corpus: Possibilidades para o ensino de Línguas e Tradução

Danielle de Almeida Menezes (UFTM)

Recebido 14 fev. 2011 / Aprovado 7, mar. 2011

Há mais de um século, o pensar sobre o ensino de línguas estrangeiras vem sofrendo mudanças em decorrência dos avanços científico-tecnológicos em diferentes áreas do saber. Na era dos métodos, cada nova teoria oriunda da linguística, da psicologia ou da sociologia, dentre outras, responsabilizava-se por embasar um novo método que, no instante de sua criação, era concebido como revolucionário. Hoje, parece consensual entre professores de línguas a visão de que não há o método mágico para o ensino de idiomas. Ao contrário, na era pós-método¹ (BROWN, 2002), percebe-se que cabe a esses profissionais lançar mão do repertório de teorias existentes e dos incrementos tecnológicos do mundo contemporâneo, em especial aqueles provenientes da informática, a fim de orientar melhor cada aluno ou grupo de alunos, atendendo às suas necessidades específicas de aprendizado linguístico. Nesse sentido, a recente publicação intitulada *Corpora no Ensino de Línguas Estrangeiras*, editada por Vander Viana e Stella Tagnin, e publicada pela HUB Editorial em janeiro deste ano, é um ótimo exemplo de como a Linguística de *Corpus* pode iluminar a prática docente, tornando-se uma aliada do professor de línguas que está constantemente interessado em aprimorar sua prática.

Os nove capítulos da obra são escritos por quinze professores-pesquisadores brasileiros que atuam em diferentes contextos de ensino de línguas. Diferentemente do que o título do volume possa sugerir, cada capítulo não se restringe a apontar possibilidades de utilização de ferramentas computacionais provenientes da Linguística de *Corpus* na sala de aula de línguas. Em outras palavras, *Corpora no Ensino de Línguas Estrangeiras* não deve ser entendido por um manual; ao contrário, devido à experiência acadêmica de seus colaboradores (bachareis, mestres, doutores, professores universitários), os estudos nele contidos buscam estimular o público alvo a realizar suas próprias investigações de ordem empírico-linguística, a elaborar seus próprios materiais didáticos com base em *corpora* e a perceber o discente como potencialmente capaz de nortear seu aprendizado por meio da observação da língua.

¹ Do inglês: post-method era.

Além desse viés instrumentalizador, a obra é a pioneira em língua portuguesa a reunir estudos acerca de diferentes idiomas (alemão, espanhol, francês, inglês e português) sob uma perspectiva educacional. É importante ressaltar aqui que, apesar de a Linguística de *Corpus* estar presente no país desde fins da última década do século XX, pouco da produção brasileira na área tem explorado a “sua interface com a educação” (p.19). No que concerne à organização, em acréscimo aos capítulos, o livro traz um preâmbulo, escrito por Mike Scott, em homenagem a Tim Johns², um dos precursores ao propor a associação entre *corpora* e ensino; um prefácio produzido por Vera Lúcia de Menezes Oliveira e Paiva e a introdução redigida pelos organizadores. O volume ainda contempla dois elementos pós-textuais que têm por objetivo auxiliar os leitores: um glossário e uma lista de *corpora* online, compilados por Stella Tagnin. Ao final, encontra-se uma seção chamada “Sobre os autores”, em que é possível conhecer um pouco do percurso acadêmico dos colaboradores e ter acesso aos seus endereços eletrônicos. Uma breve análise da estrutura da obra já evidencia a preocupação didático-pedagógica que a alicerça.

O primeiro capítulo, intitulado “Linguística de *Corpus*: Conceitos, Técnicas & Análises”, de Vander Viana (*Queen's University Belfast*), oferece uma visão geral da Linguística de *Corpus* a fim de preparar tanto o leitor pouco familiarizado com a área quanto o mais experiente para os capítulos subsequentes e também para possíveis leituras posteriores. Didaticamente, o autor inicia a discussão sobre o assunto a partir de seu conceito mais básico: *corpus*. Ao mostrar que os “*corpora* [são coleções de textos orais e escritos que] representam uma língua ou um recorte dela” (p. 29), Viana fornece o insumo necessário para definir a Linguística de *Corpus* como “uma forma de investigação empírica da linguagem a partir da exploração sistemática de um *corpus*” (p. 34). O autor ainda explica outros conceitos-chave, como, por exemplo, a diferença entre “item” e “forma”, ambos os termos utilizados para se referir a “palavra”, e a distinção entre quatro padrões de uso linguístico: a “colocação”, a “coligação”, a “preferência semântica” e a “prosódia semântica”. A maior parte do texto, contudo, destina-se a explorar as técnicas utilizadas para as análises de *corpora* com o auxílio do programa computacional *WordSmith Tools* (SCOTT, 2009), escolhido por ser empregado nos outros capítulos do livro. Assim, Viana enfoca, em subseções diferentes, os pressupostos teóricos e funcionais de três ferramentas analíticas, a saber: gerador de listas de palavras, extrator de palavras-chave e concordanciador. Cada ferramenta e suas possibilidades são discutidas minuciosamente e ilustradas por meio de referências a estudos variados, o que demonstra conhecimento aprofundado e pesquisa por parte do autor. Ao final dessas subseções, são apresentadas perguntas e respostas direcionadas a professores de línguas estrangeiras.

² Pesquisador que cunhou o termo *data-driven learning* (“aprendizagem direcionada por dados”) mais conhecido por DDL.

Tais perguntas ajudam o leitor a perceber de forma mais concreta como as ferramentas podem contribuir para o trabalho docente, materializando a informação teórica previamente oferecida.

Do segundo ao oitavo capítulo, o foco recai sobre pesquisas realizadas que abordam a relação entre línguas estrangeiras e *corpora*. No capítulo “*Working Closely with Corpus: Análise de Colocações Adverbiais em Inglês para Negócios*”, de Andréa Geroldo dos Santos, da Universidade de São Paulo, a autora busca mostrar que “a aprendizagem de colocações (...) é importante para que um aprendiz de língua estrangeira possa se comunicar na língua alvo” (p. 36). A partir da constatação de que as colocações adverbiais são pouco abordadas em livros didáticos, se comparadas a outras colocações (como as verbais e as nominais), a pesquisadora decidiu compilar seu próprio *corpus* de estudo formado por jornais, revistas e relatórios da área de negócios. Lançando mão de listas de palavras e de linhas de concordância geradas para advérbios que tiveram mais de 55 ocorrências em seu *corpus* especializado, Santos percebeu que há necessidade de outra abordagem no ensino de colocações adverbiais na área de negócios, visto que “os resultados apresentados, com base em dados autênticos apontaram para colocações que podem ser típicas dessa área (...) e que precisariam ser aprendidas e praticadas pelos aprendizes” (p.133).

Seguindo a perspectiva da pesquisa anterior, voltada para o ensino de inglês instrumental, o terceiro capítulo, “*Corpora no Ensino do Inglês Acadêmico: Padrões Léxico-Gramaticais em Abstracts de Pós-Graduandos Brasileiros*”, de Carmen Dayrell, também da Universidade de São Paulo, investiga resumos de artigos científicos em língua inglesa, produzidos por pós-graduandos brasileiros das áreas de Física, Ciências farmacêuticas e computação. Seu objetivo reside em comparar as características lexicais e gramaticais dos referidos textos com as de um *corpus* de referência, compilado a partir de artigos científicos em inglês pertencentes às áreas supracitadas. Restringindo-se à discussão de padrões léxico-gramaticais pertencentes a três itens lexicais, Dayrell conclui que seu estudo “aponta uma série de diferenças relevantes entre os dois *corpora*” (p. 167) e, semelhante à pesquisa realizada por Santos, mostra a importância de chamar a atenção dos aprendizes para aspectos que distinguem sua linguagem daquela de falantes de inglês como língua materna.

Assim como o segundo e o terceiro capítulos, “*Filmes comerciais: uma perspectiva da aplicação de pesquisa em corpus na sala de aula*” reforça a importância da abordagem DDL (aprendizagem direcionada por dados). De autoria de Marcia Veirano Pinto e Renata Condi de Souza, ambas da Pontifícia Universidade Católica de São Paulo, o capítulo possui alta relevância e contribuição clara para o ensino de inglês geral. Menos complexo que os anteriores, o estudo aproxima-se do interesse e da realidade de ensino de línguas em cursos livres e escolas regulares. Pinto

e Souza compilaram um *corpus* formado por 72 roteiros de filmes comerciais, todos disponíveis na rede mundial de computadores, e analisaram todas as linhas de concordância geradas para a palavra “just”. As autoras puderam constatar que o uso de “just” em filmes comerciais apresenta diferenças no tocante ao que materiais de ensino enfocam. O capítulo conta ainda com uma interessante seção em que é proposta uma sequência didática com cinco atividades. Diferentemente dos estudos anteriores, o quinto capítulo, “Corpora e Ensino de Tradução: o Papel do Automonitoramento e da Conscientização Cognitivo-Discursiva no Processo de Aprendizagem de Tradutores Novatos”, de Fábio Alves (Universidade Federal de Minas Gerais) e Stella Tagnin (Universidade de São Paulo), está direcionado a tradutores e a professores e alunos de tradução. Partindo de pesquisas que evidenciam que “tradutores novatos tendem a enfatizar excessivamente aspectos microtextuais do texto, concentrando esforços na solução de problemas lexicais” (p. 189), Alves e Tagnin discutem “formas de modificar as características cognitivo-discursivas que limitam o desempenho desses aprendizes” (p.190). Para tanto, os autores descrevem os procedimentos necessários para a elaboração de glossários para tradução de textos técnicos a partir da compilação de *corpora* especializados para o par português-inglês.

Desviando o foco da língua inglesa e direcionando-o para outras línguas, o sexto capítulo compara, com o auxílio de ferramentas do programa *WordSmith Tools* (Scott, 2009), artigos científicos de cardiologia publicados em português e em alemão. Partindo de um conveniente arcabouço teórico acerca da trajetória do ensino de leitura instrumental no Brasil, Maria José Bocorny Finatto, Leonardo Zilio e Elisandro José Migotto, todos da Universidade Federal do Rio Grande do Sul, contrastam as listas de palavras mais frequentes encontradas nos *subcorpora* de português e alemão do *corpus* bilíngue que compilaram. A fim de validar os resultados, os mesmos foram submetidos à análise estatística. Das constatações apresentadas por “Artigos de Cardiologia em Português e Alemão: Contribuições da Pesquisa em *Corpus* para o Ensino de Leitura Instrumental” é relevante ressaltar que o gênero artigo de cardiologia tende a seguir um mesmo roteiro nas duas línguas e “os colocados adjetivais (...) não são igualmente selecionados (...), de modo que o aprendiz e o professor (...) perceberão que nem sempre o que se esperaria a partir do português se encontra em alemão” (p.230). Apesar de os resultados se limitarem aos idiomas investigados, a descrição dos métodos empregados e o referencial teórico abordado no texto são úteis para todo professor de língua estrangeira instrumental.

O representante da língua francesa no volume intitula-se “Aplicação do Conceito de Transcategorialidade ao Ensino de Francês como Língua Estrangeira: o Caso da Marca ‘Mas’”, por Adriana Zavaglia e Marion Celli, da Universidade de São Paulo. Partindo

do pressuposto de que o ensino de gramática nas aulas de francês ainda se pauta por “concepções normativas ou tradicionais” (p. 235), Zavaglia e Celli apresentam o histórico das metodologias de ensino de francês como língua estrangeira e definem o conceito de transcategorialidade e sua importância para o ensino. A partir disso, mostram, por meio da comparação de dois *corpora online* (um para cada idioma), que palavras gramaticais podem apresentar outros valores além daquele que lhes é prototípico. Para tanto, as autoras analisam as conjunções ‘mas’ em português e ‘mais’ em francês e constataam que, em francês, há 34 opções de tradução para ‘mas’ em português, além de ‘mais’, vocábulo frequentemente escolhido por brasileiros devido à semelhança com a língua portuguesa. O capítulo apresenta ainda uma série de exercícios elaborados a partir dos achados.

O penúltimo capítulo, “Gramaticalização da Dor em Português e Espanhol: uma Abordagem Comparada com Subsídios da Linguística de Corpus e da Linguística Sistêmico-Funcional”, por Adriana Silvina Pagano e Giacomo Patrocínio Figueredo (Universidade Federal de Minas Gerais), analisa como usuários de português e espanhol abordam suas dores físicas em entrevistas, textos informativos e fóruns de discussão de livre circulação na internet. Com o respaldo teórico da Linguística Sistêmico-Funcional, que fornece bases para a compreensão de como a experiência é representada por meio da linguagem, os resultados apontam semelhanças entre as duas línguas investigadas, apesar de a ‘dor’ ser representada de maneira mais diversificada em espanhol. Como em quase todos os capítulos, Pagano e Figueredo apontam possibilidades de aplicações pedagógicas de seu estudo. Tony Berber Sardinha, da Pontifícia Universidade Católica de São Paulo assina o nono e último capítulo do livro. Seu texto, “Como Utilizar a Linguística de Corpus no Ensino de Língua Estrangeira. Por uma Linguística de *Corpus* Educacional Brasileira”, difere-se quanto à organização dos demais capítulos e se divide em duas partes: uma prática e a outra de cunho mais teórico. Na primeira e mais extensa, Sardinha oferece ao professor de línguas ideias e exemplos de como elaborar seus materiais de ensino com base em *corpus*, ampliando as possibilidades pedagógicas apresentadas em capítulos anteriores. O autor trata de atividades centradas na concordância, de atividades centradas no texto e de atividades multimídia/multigênero. A segunda parte destina-se a discutir a integração entre a Linguística de *Corpus* e o ensino de línguas no Brasil. Por meio de levantamento bibliográfico, o pesquisador mostra que esta união, cada vez mais desejável, é ainda modesta e recente no país. Contudo, a intensificação, nos últimos anos, do número de pesquisas que abordam aspectos variados da interface *corpus*-ensino demonstra que este interesse encontra-se em franca expansão.

Apesar da clara relevância da obra, algumas críticas talvez lhe caibam. Em primeiro lugar, alguns estudos tornam-se repetitivos ao apontarem justificativas e propostas pedagógicas muito semelhantes. Além disso, apesar de a obra se destinar a professores com diferentes níveis de conhecimento de Linguística de Corpus, alguns capítulos demandam grande esforço cognitivo por parte de leitores pouco familiarizados com a área e, principalmente, com modesto ou nenhum conhecimento sobre análises estatísticas. Contudo, essas questões menores não diminuem a qualidade do livro e nem de longe ofuscam um dos maiores méritos da obra, que é mostrar ao professor que ele independe dos materiais didáticos disponíveis no mercado e que não só pode como deve estimular a autonomia de seus alunos no tocante a investigações linguísticas.

REFERÊNCIAS

- BROWN, H. D. "English Language Teaching in the 'Post-Method' Era: Toward better diagnosis, treatment, and assessment". In: RICHARDS, J.C.; RENANDYA, W.A. *Methodology in Language Teaching*. Cambridge UK: Cambridge University Press, 2002.
- VIANA, V; TAGNIN, S. (Org.). *Corpora no Ensino de Línguas Estrangeiras*. São Paulo: HUB Editorial, 2011.