

INTERFACES NOS ESTUDOS DA TRADUÇÃO: CORPORA E DISCURSO

Célia M. Magalhães
Universidade Federal de Minas Gerais

Em 1993, Mona Baker lançou uma proposta pioneira modelo de estudos da tradução baseados em *corpora*, a partir de duas principais correntes de pensamento nas áreas de investigação da tradução e da lingüística. A primeira, do teórico da Universidade de Tel-Aviv, Gideon Toury, que, segundo ela, melhor elaborou o conceito de estudos descritivos da tradução como “(...) o ramo da disciplina que deve fornecer uma metodologia coerente e procedimentos explícitos de pesquisa de forma a permitir que os resultados de estudos descritivos individuais sejam expressos em termos de generalizações sobre o comportamento tradutório” (Baker, 1993:241). A segunda, do lingüista John Sinclair (apud Baker, 1993), que conseguiu, através da coleção de *corpora* computadorizados e do desenvolvimento de uma metodologia de pesquisa relevante dentro de uma nova área de estudos – a Lingüística de *Corpus*, superar as limitações humanas do pesquisador e minimizar sua dependência da intuição. A inserção e o tratamento de um número substancial de textos traduzidos num banco de dados computadorizado permite à Baker consolidar sua proposta de estabelecimento da tradução como objeto de pesquisa da disciplina estudos tradução, cujo objetivo principal passa a ser a identificação de traços do texto traduzido que levarão ao entendimento do que é e de como funciona a tradução. Tal proposta levou a pesquisadora à criação de um corpus de textos traduzidos, intitulado TEC (Translational English Corpus), armazenado na University of Manchester Institute of Science and Technology (UMIST) e disponibilizado em rede para a comunidade de pesquisadores em tradução.

Uma das preocupações de Baker (1995, 1996a) foi estabelecer as bases de um novo ramo da disciplina, definindo conceitos e propondo perguntas para investigação do objeto de estudo desse ramo. O conceito de *corpus* é usado pela pesquisadora como “(...) qualquer coleção de textos inteiros (...) em formato eletrônico, analisáveis automaticamente ou semi automaticamente (...)” (Baker, 1995:226). Os *softwares* disponíveis no mercado para processamento dos *corpora* são o WordSmith Tools, comercializado pela Oxford University Press, e o Multiconcord, comercializado por David Wools. A ferramenta básica de análise nesses *softwares* é o **KWIC concordance** (**KWIC** é a sigla para **Key Word In Context**), uma lista de todas as ocorrências de uma palavra chave ou expressão especificada no *corpus*, cada uma no meio de uma linha num determinado contexto. As concordâncias de palavras chave podem ser feitas de modos variados, à esquerda ou direita da palavra chave, podendo também ser ampliadas em rede (no nível da frase ou do parágrafo) para revelar mais do contexto, ou dar um perfil da palavra chave listando as combinações mais freqüentes dentro de um espaço determinado. Além das concordâncias, esses softwares oferecem outras ferramentas úteis à investigação da tradução, como a listagem de todas as palavras do corpus, ou de um texto ou grupo de textos específicos do *corpus*, por ordem de freqüência ou alfabética.

Baker (1995:230) designou três tipos principais de *corpora* usados em estudos da tradução, antecipando-se a possíveis confusões de terminologia quando o número de atividades na área for muito grande. O primeiro, os *corpora* paralelos, consistem de textos originais, numa língua de partida A e suas versões traduzidas numa língua B. Eles permitem estabelecer, objetivamente, como os tradutores superam as dificuldades de tradução na prática, e usar essa evidência para fornecer modelos realistas para o treinamento de tradutores; além disso, têm um papel crucial na exploração de normas de tradução em contextos socioculturais

e históricos (Baker, 1995:231). Vale ressaltar os trabalhos recentes de Baker (2000), de estudo dos estilos de tradutores literários e de Kenny (2001), de investigação da criatividade no texto traduzido, em *corpora* paralelos de romances (alemão/inglês).

O segundo, os *corpora* multilíngües, são conjuntos de dois ou mais *corpora* monolíngües em línguas diferentes, criados na mesma instituição ou em instituições diferentes com base nos mesmos critérios. Permitem-nos estudar os itens e traços lingüísticos no ambiente da língua tal como produzida originalmente. Sua utilidade está em dar acesso aos padrões naturais da língua objeto de estudo; portanto, têm papel crucial na preparação de material didático, no treinamento do tradutor e na melhoria do desempenho dos sistemas de tradução automática.

Os *corpora* comparáveis consistem de duas coleções separadas de textos na mesma língua: um *corpus* de textos originais na língua em questão e o outro de textos traduzidos para essa língua, a partir de uma ou mais línguas determinadas. Os dois *corpora* devem cobrir um domínio, uma variedade de língua e um tempo cronológico semelhantes e ser de tamanho semelhante. O *corpus* de textos traduzidos deve ser representativo em termos do número de autores originais e de tradutores. O papel desses *corpora* na disciplina de estudos da tradução é o de identificar padrões específicos dos textos traduzidos, sejam quais forem as línguas de partida ou de chegada.

Em 1999, partindo da proposta de Baker (1993), um grupo de pesquisadores da Faculdade de Letras da UFMG, do qual faço parte, decidiu elaborar um projeto de ensino baseado em *corpora*, atendendo a edital interno da instituição, de projetos para melhoria das condições de ensino da graduação. Logo de início, houve o entendimento do grupo de que tal projeto deveria contar, no mínimo, com equipamentos avançados de computação, além de *softwares* específicos e *CD-ROMS* de *corpora* já disponíveis no mercado. Assim sendo, foi elaborado um projeto de melhoria de ensino na graduação que otimizou um dos laboratórios de computação já existentes na faculdade com a aquisição de *kits* multimídia, *softwares* de análise e *CD-ROMs*, e cujo objetivo era a criação de *corpora* para uso no ensino de leitura, produção e tradução de textos. Esta se constituiu na primeira etapa do trabalho, subdividida em duas fases distintas: os primeiros seis meses para aquisição, instalação e exploração dos materiais, os segundos para o início da seleção de textos para os *corpora*, bem como a transformação de alguns deles em formato eletrônico. Vale observar que o objetivo era criar os três tipos de *corpora* definidos por Baker (1995): os *corpora* paralelos e multilíngües, visando a aplicação imediata nas atividades de ensino bem como o uso em pesquisa em estudos da tradução, e os *corpora* comparáveis, para investigação de traços dos textos traduzidos para o português do Brasil.

No ano seguinte, três dos membros do primeiro grupo elaboraram um projeto de pesquisa para obter subsídios financeiros que lhes permitissem agregar estudantes da graduação ao trabalho, como iniciantes à pesquisa e com bolsa equivalente à de iniciação científica. Esse projeto integrou seis estudantes que exploraram um dos *softwares* de análise adquiridos, o *WordSmith Tools*, além de iniciar a montagem de pequenos *corpora* para as primeiras análises. O objeto e objetivos de pesquisa de cada membro do grupo foram delineados de modo a atender os interesses e áreas específicas de cada um. No início de 2001, a Pró-Reitoria de Graduação concedeu, novamente, recursos equivalentes a seis bolsas para estudantes da graduação atuarem por dez meses no projeto, desta vez ampliado para contemplar também a abordagem de estudos cognitivos da tradução.

No momento, o corpus já desenhado, intitulado CORDIALL (Corpus Discursivo para Análises Lingüísticas e Literárias), conta com *corpora* paralelos e comparáveis, abrigando quatro pesquisadores orientadores de sub-projetos de pesquisa, dentre eles três da área de tradução e uma da área de ensino de português para estrangeiros, duas mestrandas da área de tradução e

uma da área de análise do discurso, além de dois bolsistas de PIBIC/UFMG. O CORDIAL tem possibilitado o desenvolvimento dos seguintes projetos: a análise de corpora paralelos de romances policiais das décadas de trinta e quarenta, visando à investigação de mudanças no fazer tradutório de acordo com o eixo histórico; os estudos baseados em *corpora* paralelos de traduções de textos da literatura de expressão inglesa, visando à investigação das representações da questão cultural de raça, de acordo com mudanças ideológicas na prática social da tradução e, finalmente, estudos baseados em corpora comparáveis, visando à análise de elementos coesivos e dos verbos de relato em textos traduzidos e originalmente escritos em português brasileiro.

O referencial metodológico dos estudos de *corpora* possibilita a necessária articulação de categorias macrodiscursivas de análise, próprias das tradições analítico-interpretativas de áreas pós-disciplinares como a literatura comparada e os estudos culturais (cf. Magalhães 1997), com aspectos microdiscursivos dos textos. O impacto que a proposta atual pode causar (cf. Magalhães 2001a), tanto para os estudos da tradução quanto para os estudos de discurso, é a explicação do fenômeno da linguagem de forma mais abrangente, conjugando aspectos macro e micro discursivos a partir de uma abordagem original, até então anunciada ou sistematizada por apenas três teóricos, dois da tradução e um de estudos culturais. Baker (1996b) previu a possibilidade de união dos paradigmas de estudos culturais e lingüísticos através da abordagem de *corpora*; Harvey (2000) investiga efetivamente textos traduzidos a partir de um referencial teórico de *gay studies* e de lingüística crítica e Barker e Galasinski (2001) promovem um diálogo entre os estudos culturais e da linguagem, investigando representações de identidades usando o referencial teórico da análise crítica do discurso.

No caso do presente projeto, de interfaces nos estudos da tradução entre as abordagens de análise do discurso e de *corpora*, o espaço pós-disciplinar proposto como capaz de fazer confluir abordagens da tradução até então competitivas, como a dos estudos culturais e a dos estudos lingüísticos (cf. Baker, 1996a), é o de uma abordagem discursiva crítica (Fairclough 1992, 1999; Stubbs, 1996; Magalhães, 2001b), tendo o texto como ponto de partida para uma análise que articula categorias de análise sociocultural com categorias lingüístico-discursivas, através da metodologia da Lingüística de *Corpus*.

Bibliografia

- BAKER, Mona. "Corpus Linguistics and Translation Studies: Implications and Applications". In: BAKER, Mona, FRANCIS, Gill e TOGNINI-BONELLI, Elena. *Text and Technology: In honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins Publishing Co., 1993, p. 233-250.
- BAKER, Mona. "Corpora in Translation Studies: An Overview and Some Suggestions for Future Research". *Target*, Vol. 7:2, 1995, p. 223-243.
- BAKER, Mona. "Corpus-based Translation Studies: The Challenges that Lie ahead". In: SOMERS, Harold (ed.). *Terminology, LSP and Translation Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins Publishing Co., 1996a, p. 177-186.
- BAKER, Mona. "Linguistics & Cultural Studies: Complementary or Competing Paradigms in Translation Studies?". In: LAUER, Angelika; GERZYMISCH-ARBOGAST, Heidrun; HALLER, Johann e STEINER, Erich. *Übersetzungswissenschaft im Umbruch: Festschrift für Wolfram Wilss zum 70. Geburtstag*. Tübingen: Gunter Narr Verlag, 1996b, p. 9-19.
- BAKER, Mona. "Towards a methodology for investigating the style of a literary translator". *Target*, 12:2, 2000, p. 241-266.
- BARKER, Chris e GALASÍNSKI, Dariusz. *Cultural Studies and Discourse Analysis: a Dialogue on Language and Identity*. London, Thousand Oaks, New Delhi: Sage Publications, 2001.

- FAIRCLOUGH, N. *Discourse and Social Change*. Cambridge: Polity Press, 1992.
- FAIRCLOUGH, N. *Discourse in Late Modernity*. Edinburgh: Edinburgh University Press, 1999.
- HARVEY, Keith. "Translating Camp Talk: Gay Identities and Cultural Transfer". IN: VENUTI, Lawrence. *The Translation Studies Reader*. London & New York: Routledge, 2000, P. 446-467.
- MAGALHÃES, Célia M. Os monstros e a questão racial na literatura pós-colonial brasileira. Belo Horizonte: UFMG, 1997. (tese de doutorado inédita).
- MAGALHÃES, Célia M. "Pesquisas textuais/discursivas em tradução: o uso de corpora". In PAGANO, Adriana S. *Metodologias de pesquisa em tradução*. Série Estudos Lingüísticos, Vol. 3. Belo Horizonte: FALE/UFMG, 2001a, p. 93-116.
- MAGALHÃES, Célia M. *Reflexões sobre a análise crítica do discurso*. Série Estudos Lingüísticos, Vol. 2. Belo Horizonte: FALE:UFMG, 2001b.
- STUBBS, Michael. *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture*. Oxford, UK e Cambridge, MA: Blackwell Publishers, 1996.