

A LINGUÍSTICA DE *CORPUS* NA FORMAÇÃO DE TRADUTORES: COMPILAÇÃO E ANÁLISE DE UM *CORPUS* DE APRENDIZES DE TRADUÇÃO

Autor: Joacyr Oliveira

Orientador: Professora Doutora Stella Esther Ortweiler Tagnin

Universidade de São Paulo – USP

Linguística de Corpus

Início da pesquisa: 1º semestre de 2012 (Porém iniciei informalmente sob orientação da professora Stella no 2º semestre de 2011.)

Previsão de defesa: Início de 2014

Com pouquíssimas exceções, os cursos superiores de formação de tradutores e intérpretes no Brasil são relativamente novos e, em muitos centros universitários da rede privada, são classificados como uma ramificação de licenciatura em Letras. Além disso, pouquíssimos são os estudos sobre o ensino e a aprendizagem da tradução em faculdades no Brasil.

Tal cenário serve de estímulo para nosso projeto de pesquisa, cujo passo inicial está na construção de um corpus de aprendizes de tradução formado por textos oriundos dos alunos de duas faculdades particulares na cidade de São Paulo, onde atuamos como professor de tradução há 10 anos. Nosso objetivo é, por meio do mapeamento, da identificação e da análise das preferências e tendências nas escolhas lexicais e morfossintáticas dos tradutores-aprendizes, compreender, a partir da perspectiva do aluno, como se dá a aprendizagem das técnicas de tradução.

Esta apresentação tem por finalidade expor e colocar para discussão a primeira questão encontrada no desenvolvimento de nossa pesquisa: como alinhar um texto original com cerca de 50 ou 60 traduções. Iniciamos, claro, com as ferramentas tradicionalmente usadas para alinhamento como *ParaConc*, *Aligner* do WordSmith Tools, *EAlign*, *CBW*, assim como algumas on-line como *CasualConc*(<http://sites.google.com/site/casualconc/utility-programs/casualmultipconc/casualmultipconc-text-aligner>) e *ABBYY Aligner Online* (<http://aligner.abbyyonline.com/en>) entre outras; porém nos deparamos com o mesmo problema: o número elevado de traduções referentes a um mesmo original. Além dos alinhamentos, também precisávamos buscar uma solução para a etiquetagem das linhas e inserção de cabeçalho em cada tradução.

A ferramenta que nos ofereceu as condições necessárias para os alinhamentos, etiquetagem das linhas e inserção de cabeçalhos em todos os textos foi o *Microsoft Excel* do Pacote Microsoft Office 2010.

Cerca de 400 alunos cederam-nos de 10 a 15 de suas traduções para a compilação do corpus. Os textos originais foram segmentados por períodos sintáticos, para, em seguida, serem traduzidos. O trabalho final dos alunos foi apresentado em unidades de tradução, ou seja, segmento original seguido paralelamente por sua tradução. Por fim, o arquivo, já digitalizado, foi enviado por e-mail. Ao recebermos os arquivos, manualmente transferimos apenas a tradução para uma planilha do *Excel* previamente formatada para receber um original em uma coluna e todas as traduções do referido texto nas colunas à direita. Tanto o original como cada uma de suas traduções ocupam três colunas da planilha: a coluna 1 contém as etiquetas de abertura, por exemplo <ua_id=en00101>, onde os três primeiros dígitos identificam o texto e os dois últimos, o número da linha do texto original em inglês; as traduções também seguem o mesmo código, sendo os dígitos são antecidos por “pt”; a coluna 2 traz o texto

(original ou tradução); e a coluna 3 apresenta a etiqueta que fecha cada linha, </ua>. Todas as traduções contêm um cabeçalho devidamente etiquetado e com informações como idade, conhecimento de inglês, experiência com tradutor etc., provenientes de um questionário respondido por cada aluno participante. Finalmente cada grupo de três colunas com etiquetas e texto é selecionado e transferido para um arquivo TXT, que fará parte do corpus e que será analisado por meio da ferramenta *WordSmith Tools*.

Referências

BERNARDINI, S. (2002). “Educating translators for the challenges of the new millennium: The potential of parallel bi-directional corpora”, *in*: HALLER, J.; MAIA, B. e ULRICH, M. (eds.) *Training the language services provider for the new millennium*. Porto: Universidade do Porto. 173-186.

LEVY, J.(2000). “Translation as a Decision Process”, *in*: *The Translation Studies Reader*. London: Routledge, 148-59.

MARTINEZ, R.; JOSEBA, A. e ARANTIZA, C. *Bitext Correspondences through Rich Mark-up*. UKPMC Labs. UK PubMed Central. Disponível em: <http://ukpmc.ac.uk/abstract/CIT/578243/reload=0;jsessionid=pCe6rIpJ8E3NUFwFkbpM.34>.

SCOTT, M. (2012). *WordSmith Tools*. Liverpool: Lexical Analysis Software.