

DIFICULDADES INERENTES À COMPILAÇÃO DE CORPUS DE APRENDIZES

Orientador: Prof. Dr. Tufi Neder Neto

Autores: Leonardo de Oliveira, Sheila Maria Resende e Carolina Alves Carvalho

Data de início: 2011

Data aproximada de término: 2015

Este trabalho originou-se durante a etapa de transcrição de entrevistas gravadas de brasileiros aprendizes de inglês, feita pelos estudantes participantes do grupo de estudos em Língua e Literatura Estrangeira, da Universidade Federal de Lavras. Este grupo está vinculado ao COBAI, ramo do projeto internacional LINDSEI, da universidade de Louvain, para constituição de um corpus oral brasileiro de aprendizes de língua inglesa. Atualmente, prioriza-se trabalhos com corpora de falantes nativos no ensino/aprendizado de L2. Tono (2003) observa que a maioria dos projetos com corpus de aprendizes surgiu na última década e os resultados das pesquisas baseadas neles ainda são relativamente limitados. “Profissionais da linguagem tomam como verdadeiro que os modelos apropriados de uso de uma língua venham de seus falantes nativos” (Cook, 1999: p. 185), todavia, falantes nativos e aprendizes de L2, segundo Cook, diferem nos conhecimentos de seus respectivos primeiro e segundo idiomas, bem como em alguns de seus processos cognitivos. Assim, falantes de L2 devem ser considerados como tal e não como a aproximação de um falante nativo. O objetivo deste trabalho é identificar e discutir as principais dificuldades encontradas pelos participantes do projeto quanto à adequação das transcrições às convenções pré-estabelecidas pelo LINDSEI para indicar ocorrências exclusivas da fala. Para tanto, cada membro do projeto faz a sua própria transcrição. Depois, elas são comparadas entre si e seus aspectos divergentes são discutidos. As dificuldades mais comuns, em geral decorrentes diferentes interpretações das regras e do nível individual de proficiência na língua inglesa, são aquelas relacionadas às pausas preenchidas, aos retornos dados pelo receptor, às palavras truncadas e à sobreposição de discurso. Nos dois primeiros casos, o principal entrave se deve a precisão exigida na identificação dos sons produzidos por entrevistador e entrevistado, sendo corriqueiras interpretações distintas, variando de acordo com a capacidade de distinguir e interpretar individualmente cada aspecto. O mesmo se aplica à transcrição das palavras truncadas. Quanto aos turnos de fala sobrepostos, estes se devem especialmente às entrevistas nas quais ocorrem muitas sobreposições em um único turno, o que torna difícil a identificação de cada palavra e de seu respectivo falante e, ainda, como apontar esta sobreposição. Espera-se que, no decorrer do trabalho, a prática elimine a dificuldade em empregar as convenções. A correta aplicação destas proporciona precisão aos corpora de aprendizes e confere a eles um efetivo potencial de benefício a áreas da lingüística de corpus como a da análise de erros (Tono, 2003), os estudos comparativos entre L1 e L2, bem como o estudo das questões referentes à inter-linguagem (Cook, 1999).

BÉLGICA. Sítio da Universidade de Louvain. Disponível em: <http://www.uclouvain.be/em-cecl-lindsei.html>. Acesso em 12-03-2012.

COOK, V. Going Beyond the Native Speaker in Language Teaching. TESOL Quarterly, v. 33, n. 2, p. 185-209, 1999.

TONO, Y. Learner corpora - design development n applications - Proceedings of the 2003 Corpus Linguistics Lancaster UK 28-31 march.