

## **PROFICIÊNCIA ESCRITA EM LÍNGUA ADICIONAL: COH-METRIX-PORT E CELPE-BRAS**

Autor: Aline Evers

Orientador: Maria José Bocorny Finatto

Instituição: Universidade Federal do Rio Grande do Sul, Instituto de Letras, Linha de pesquisa: Teorias Linguísticas do Léxico

Data de início da pesquisa: março de 2011

Previsão de data de defesa: março de 2013

É possível contribuir com a mensuração de proficiência escrita em língua adicional (LA) utilizando métricas de avaliação de inteligibilidade textual? Trazemos possíveis respostas a essa pergunta através da discussão de resultados parciais obtidos com o processamento de textos avaliados em um exame de proficiência do português brasileiro. Pesquisas realizadas com a finalidade de mensurar proficiência escrita em LA fazem uso de medidas superficiais, tais como diversidade lexical, repetição de palavras e tamanho de texto (Jarvis, Grant, Bikowski e Ferris, 2003). No entanto, medidas mais sofisticadas, como sentido e intenção, níveis de linguagem e análise conceitual, dificilmente são apreendidos automaticamente por ferramentas computacionais. Essa lacuna vem sendo preenchida por pesquisadores com o uso da ferramenta Coh-Metrix (Graesser et al, 2004; McNamara et al, 2002; Crossley et al, 2007), que tem a finalidade de avaliar a coerência e a dificuldade de um texto (em inglês) explorando três níveis de análise linguística: lexical, sintático e semântico. Tendo em vista os exames de proficiência de larga escala em inglês, muito tem sido feito para compreender indicadores de proficiência escrita em LA nesse idioma a fim de tornar as avaliações mais confiáveis e viáveis. Segundo estudos realizados em/para língua inglesa (Crossley e McNamara, 2012), as métricas calculadas pelo Coh-Metrix foram considerados capazes de prever a proficiência escrita (métricas: diversidade lexical, frequência de palavras, significado atribuído pelo leitor, repetição de aspecto verbal e familiarização com a palavra). Inspirado nas pesquisas mencionadas, este trabalho faz uso de um corpus de 177 textos produzidos para a obtenção do Certificado de Proficiência em Língua Portuguesa para Estrangeiros (CELPE-Bras), um exame de larga escala e de alto impacto que avalia a competência de uso oral e escrito da língua portuguesa em sua variedade brasileira. Os atributos avaliados nesse exame buscam refletir o uso da língua em situações reais de comunicação, em que o candidato é levado a desempenhar tarefas próximas daquelas desenvolvidas cotidianamente (Dell'Isola, Scaramucci, Schlatter, Júdice, 2003). Não se busca aferir conhecimentos sobre a língua, mas sim a capacidade de uso dessa língua, a competência linguística, que é um dos componentes da competência comunicativa. Partimos da mesma hipótese de Crossley e McNamara (2012) de que, embora não esteja sendo observado de maneira pontual pelos avaliadores, o uso de léxico e sintaxe mais complexos é produtivo para a predição de proficiência escrita em português língua adicional (PLA), e está presente, mesmo que de forma não marcada, na apreciação dos avaliadores. Para mensurar esses elementos, utilizamos a ferramenta Coh-Metrix-Port (Scarton e Aluísio, 2010), a adaptação parcial das métricas da ferramenta americana para a língua portuguesa. Processamos os textos na ferramenta e agrupamos os resultados de acordo com os 4 níveis certificados pelo exame (Intermediário, Intermediário Superior, Avançado e Avançado Superior). A partir da observação inicial dos resultados, é possível verificar que textos mais avançados não são mais coesivos, mas são mais sofisticados linguisticamente. Os

resultados contribuem para redimensionar o ensino de escrita em PLA e fornecem subsídios para repensar e melhor descrever os níveis de proficiência certificados pelo exame.

Referências bibliográficas:

CROSSLEY, S. A., LOUWERSE, M. M., MCCARTHY, M. P., McNAMARA, D. S. A linguistic analysis of simplified and authentic texts. **Modern Language Journal**, 91, (2), páginas 15-30, 2007.

\_\_\_\_\_, S. A.; MCNAMARA, D. S., Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication. **Journal of Research in Reading**, 35: 115–135. doi: 10.1111/j.1467-9817.2010.01449.x., 2012.

DELL'ISOLA, R., SCARAMUCCI, M., SCHLATTER, M., JÚDICE, N.. A avaliação de proficiência em português língua estrangeira: o exame CELPE-Bras. **Revista Brasileira de Linguística Aplicada**, 3, jan. 2012.

GRAESSER, A.C., McNAMARA, D.S., LOWERSE, M.M. & Cai, Z. Coh-Metrix: Analysis of text on cohesion and language. **Behavior Research Methods, Instruments, and Computers**, 36, 193–202, 2004.

JARVIS, S., GRANT, L., BIKOWSKI, D. & FERRIS, D. Exploring multiple profiles of highly rated learner compositions. **Journal of Second Language Writing**, 12, 377–403, 2003.

McNAMARA, D. S., LOUWERSE, M. M., GRAESSER, A. C. Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. **Grant proposal**. Disponível em: <http://csep.psyc.memphis.edu/mcnamara/pdf/IESproposal.pdf>, 2002.

SCARTON, C. E. ; ALUÍSIO, S. M. . Análise da Inteligibilidade de textos via ferramentas de Processamento de Língua Natural: adaptando as métricas do Coh-Metrix para o Português. **Linguamática - Revista para o Processamento Automático das Línguas Ibéricas**, v. 2, p. 45-62, 2010.