

Corpus TecEM – Construção de um Corpus de textos escritos por alunos de cursos técnicos integrados ao ensino médio

Vitor Angelo dos Santos Gouvêa¹, Maitê Moraes Gil^{1*}
*Orientadora

¹Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) -
Campus Osório. Osório, RS

O projeto "Corpus TecEM – Construção de um Corpus de textos escritos por alunos de cursos técnicos integrados ao ensino médio" tem como objetivo central organizar um corpus de aprendizes de língua portuguesa, mais especificamente, com textos de alunos de cursos técnicos integrados ao ensino médio, oriundos de Institutos Federais localizados no Rio Grande do Sul. Ainda que tenha grande centralidade no currículo escolar, a produção textual ainda possui fragilidades que requerem medidas especializadas. Não há um corpus que organize e disponibilize textos de aprendizes de língua portuguesa como língua materna, o que dificulta pesquisas e desenvolvimento de conhecimentos focados nessa questão no Brasil. O potencial acadêmico e a elaboração de futuras práticas docentes são possíveis resultados que evidenciam a importância da construção de um corpus desse alcance. Os princípios da Linguística de Corpus foram essenciais para a elaboração da metodologia deste projeto. Neste contexto, quatro pré-requisitos para a formação de um corpus computadorizado precisam ser respeitados: sua autenticidade, produção dos textos por falantes nativos, uma seleção criteriosa de textos e a preocupação com a representatividade. Para que houvesse uma representatividade igualitária dos Institutos Federais localizados no estado, foi utilizado o agrupamento dos COREDEs, fóruns regionais de discussão sobre estratégias, políticas e ações que visam ao desenvolvimento regional, em nove regiões funcionais do estado do Rio Grande do Sul. Das etapas previstas, o site do Corpus está em fase final de construção, já foram definidos os critérios da coleta de textos, os professores dos campi selecionados foram contatados e a coleta de textos foi iniciada. O potencial resultado deste projeto, que tem a autenticidade como norteadora, portanto, é disponibilização de um conjunto organizado de textos que possam servir como base para a formulação de materiais didáticos vinculados com as reais necessidades desse grupo de alunos, assim como se tornar um espaço de investigação sobre as produções textuais que fazem parte do corpus.

Palavras-chave: Corpus. Linguística Cognitiva.