

SALÃO DE
INICIAÇÃO CIENTÍFICA
XXIX SIC
UFRGS
PROPESQ



múltipla 
UNIVERSIDADE
inovadora  inspiradora

Evento	Salão UFRGS 2017: SIC - XXIX SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2017
Local	Campus do Vale
Título	Do corpus ao glossário: iniciação à terminologia
Autor	THIANE CECONI
Orientador	ANNA MARIA BECKER MACIEL

Do corpus ao glossário: iniciação à terminologia

Autoras: Thiane Ceconi e Luísa Feres de Aguiar Rabaldo

Orientadora: Anna Maria Becker Maciel

Universidade Federal do Rio Grande do Sul - Instituto de Letras

O presente trabalho tem o objetivo de relatar como Bolsistas de Iniciação Científica (BICs), graduandas do curso de Bacharelado em Letras, sem conhecimento terminológico prévio, compreendem a metodologia envolvida no processo de construção de um *corpus* especializado multilíngue para a elaboração de um glossário. O tema do glossário em andamento é a terminologia da conservação de documentos em suporte papel tendo como língua de partida o português. A língua estrangeira a ser trabalhada pelas graduandas é o inglês. Com a ajuda de *softwares* de análise textual, AntConc (Windows 3.4.4) e WordSmith (Windows, 6.0), candidatos a termo serão retirados do *corpus* em português. Paralelamente, candidatos a termo, definições, exemplos e outras informações equivalentes deverão ser coletados utilizando as mesmas ferramentas, a partir do *corpus* semelhante em inglês construído pela equipe da língua inglesa com a colaboração das BICs. O uso do *corpus* especializado como fonte de coleta é essencial, pois as informações retiradas dele são de contextos reais de uso dos especialistas. A base do trabalho é a utilização de *corpora* como metodologia terminográfica. Segundo Bevilacqua (2013, p. 12) “tal perspectiva prevê que se identifiquem, analisem e descrevam os termos, considerando as diferentes situações comunicativas em que são utilizados, e que, conseqüentemente, se considerem os aspectos linguísticos, comunicativos e cognitivos”. Nossa tarefa no projeto até agora foi o pré-processamento, organização e catalogação dos textos científicos, apresentações publicadas em anais de eventos, artigos de periódicos acadêmicos, teses e dissertações e manuais escritos por especialistas, compilados pelas orientadoras das páginas da *web*. Em primeiro lugar convertemos todos os textos, independentemente de formato a documentos do Microsoft Word (.doc) para realizar sua limpeza de acordo com Aluizio e Almeida (p. 106; 2006). Dessa forma, removemos todos os elementos não-textuais, ou seja, imagens, tabelas, cabeçalhos, rodapés, gráficos, agradecimentos, referências bibliográficas. Depois, salvamos os textos até agora em formato .doc como texto sem formatação (.txt) para serem nomeados por um código pré-estabelecido, catalogados e armazenados. Essa fase inicial nos proporcionou uma visão panorâmica da área da conservação dos documentos em papel, bem como nos fez antever a complexidade da pesquisa terminológica. Como resultado de nossa iniciação à terminologia, demonstraremos as diferentes etapas da construção do *corpus* de língua inglesa e também nossa compreensão da metodologia da identificação dos termos mais salientes da área.

Palavras-Chave: Conservação de Documentos em Suporte Papel, *Corpus*, Terminologia,

Referências

BEVILACQUA, Cleci. Por que e para que a linguística de corpus na terminologia. In: TAGNIN, S.; BEVILACQUA, C.(Org.). ***Corpora na terminologia***. São Paulo: Hub editorial, 2013. p. 11.

ALUÍSIO, S. M.; ALMEIDA, G. M. de Barcellos. O que é e como se constrói um corpus? ***Calidoscópico***, São Leopoldo. Vol. 4, n. 3, p. 156-178, set/dez 2006. Disponível em: <<http://revistas.unisinos.br/index.php/calidoscopio/article/view/6002>>. Acesso em 16 fev. 2017.