

Introdução

Esta comunicação tem como objetivo apresentar o processo de construção e preparação de um *corpus* textual em francês voltado à conservação de bens culturais móveis em suporte papel. Este trabalho faz parte do projeto *A Linguagem do Patrimônio Cultural Brasileiro: Conservação dos Bens Culturais Móveis – Francês e Italiano*, desenvolvido pelo grupo Termisul.

Contextualização do trabalho

Devido ao crescente aumento da importância dos estudos de Conservação e Restauração no Brasil, este novo projeto do grupo visa à produção de um glossário multilíngue (português, inglês, espanhol, francês, italiano e russo) da área. Futuramente, o glossário estará disponível para consulta em uma base de dados *on-line*. Os *corpora* compilados serão a base da busca dos termos e combinatórias léxicas especializadas que comporão o glossário.

Objetivo da comunicação

- Descrever o processo de construção e preparação do *corpus* textual em francês.

Pressupostos teóricos

- Teoria Comunicativa da Terminologia (CABRÉ, 1999);
- Linguística de *Corpus* (BERBER SARDINHA, 2004);
- Estudos de Tradução (HURTADO ALBIR, 2001);
- Socioterminologia (FAULSTICH, E. 2006).

Corpus

O *corpus* é composto de textos francófonos relativos à conservação do patrimônio em suporte papel. Os textos são acadêmicos (artigos, dissertações, teses), além de técnicos (boletins, relatórios, notas técnicas e documentos de órgãos de referência na área).

Metodologia

1. Coleta de textos especializados;
2. Conversão e limpeza;
3. Catalogação;
4. Armazenamento.

Referências:

- ANTHONY, L. AntConc (3.2.4 v). Tokyo: Waseda University, 2014.
BERBER SARDINHA, T. *Linguística de Corpus*. Barueri, SP: Manole, 2004.
CABRÉ, M. T. *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida, 1993.
FAULSTICH, E. A socioterminologia na comunicação científica e técnica. *Ciência e Cultura*, v.58, n.2, 2006. Disponível em: http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200012&script=sci_arttext.
HURTADO ALBIR, A. *Traducción y traductología*. Madrid: Gredos: 2001.

Coleta de textos especializados

- Seleção de países francófonos: França, Canadá, Suíça e Bélgica;
- Busca de portais de bibliotecas de universidades e seus respectivos repositórios digitais;
- Busca de revistas acadêmicas da área e órgãos de referência (por exemplo, Institut Canadien de Conservation);
- Uso das palavras de busca: *conservation, papier, patrimoine, restauration, preservation*.

Conversão e limpeza

Após a coleta de textos, é necessário convertê-los e limpá-los para que possam ser reconhecidos pelo *software* AntConc. Essa ferramenta será utilizada na etapa seguinte do trabalho, na busca dos equivalentes.

1. Conversão de (.pdf) para (.doc) através do OneNote;
2. Limpeza do texto (retiram-se as informações irrelevantes para a pesquisa, como agradecimentos, sumário, figuras etc.);
3. Conversão do (.doc) limpo para (.txt).

Catalogação

O texto é catalogado de acordo com um padrão estabelecido pela equipe do projeto, a saber: código; autor, título, fonte; endereço da web onde está disponível; data de acesso. O código é estabelecido pela abreviação da nacionalidade do texto (be, ca, fr ou sw), as iniciais PP (projeto Papel) e a numeração do texto.

Exemplo: caPP003 – DE BOGUI, Sarah. Le patrimoine des bibliothèques québécoises. *Documentation et bibliothèques*, Montréal, vol. 54, nº 2, p. 105-110, abr-jun. 2008. Disponível em: <https://www.erudit.org/fr/revues/documentation/2008-v54-n2-documentation01777/1029318ar/> Acesso em: 13 fev. 2017.

Armazenamento

Por fim, o texto original e seu formato limpo e convertido são igualmente armazenados sob o título de seu código de catalogação em nossos arquivos.

Considerações finais

- Obtivemos 94 textos acadêmicos e técnicos em língua francesa da área de Conservação e Restauração, abrangendo a produção intelectual da França, Canadá, Bélgica e Suíça. Embora tenhamos buscado textos em mais países francófonos, como Marrocos e Congo, não encontramos textos passíveis de compor nosso *corpus* devido à falta de pesquisa na área ou de disponibilidade *on-line* de materiais;
- O catálogo desenvolvido será disponibilizado na base de dados do projeto juntamente com os termos obtidos a partir dos textos coletados;
- Na segunda etapa do projeto será realizada a extração de termos em português e serão buscados seus equivalentes nas línguas estrangeiras a partir do *corpus* desenvolvido em cada língua.