

Abs-72

Ana Rita Remígio (Universidade de Aveiro)

Corpus-based terminography: popularizing discourse and the role of communicative contexts in specialised corpora building

The lack of, or reduced, understanding of the consumer, as far as food items with health claims that are available on the market are concerned has led to the proposal of a terminological database targeted to non-experts, with the aim of making scientifically valid and accurate information on the so called functional food systematically available. The resource was conceived, built and populated within the framework of a corpus-based approach for terminography, i.e., the applied field of Terminology, which besides the textual, also comprises the conceptual and communicative dimensions.

The singular nature of this project lies in the fact that texts on functional food targeted to the consumer are not only written by researchers, professors or science communicators, but also by actors from the food industry and journalists, who produce texts according to given communicative intentions: if the target public is heterogeneous, so are the text producers. Consequently and upon familiarization with the special subject field, the following question was raised: how to select and organise texts on functional foods produced by different actors, in the corpus, in order to be able to identify and extract term candidates and contexts rich in conceptual information, as a basis for definition writing, to populate the resource with information that suits the needs of the consumer?

The answer to that question was based on the hypothesis that the identification of communicative contexts, i.e the circumstances of discourse production, considering text producers, communicative intention and target public, in which popularising discourse on functional food is produced could be intimately related to text selection criteria and the design of the specialised corpus.

Therefore, texts from different genres were selected, with the aim of obtaining a representative sample of popularizing discourse on functional foods, and organized within the corpus according to the communicative context in which they were produced. Those texts were then separately analysed, using WordSmith Tools, so that terminological information identified and extracted could then be compared: quantitatively and according to relevance. This comparison aimed to evaluate the adequacy of inclusion of texts produced in three different communicative contexts in the corpus, of corpus design and, as a consequence, to assess the need to redesign the corpus, in order to obtain more and better results. A reference corpus of texts targeted at subject field specialists, and thus representative of scientific discourse on the field, was also built, again for comparison purposes and to verify if information not present in the corpus of study and relevant for the resource was identifiable.

Results of this study are going to be disclosed and concluding remarks on the process of designing a specialised corpus, representative of popularizing discourse according to communicative contexts for terminographical purposes, will be discussed.

BIBER, Douglas – Representativeness in corpus design. *Literary and Linguistic Computing*. ISSN 1477-4615. 8:4 (1993). p. 243-257.

COSTA, Rute; SILVA, Raquel – De la typologie à l'ontologie des textes. In *Terminologies & Ontologies: théories et applications*. Actes de la 2ème Conférence – Toth Annecy. Annecy : Institut Porphyre, 2008. Savoie et Connaissance.

COSTA, Rute – Terminology, corpus linguistics and ontologies. In *Contrastive studies and valency: studies in honour of Hans Ulrich Boas*. Frankfurt: Peter Lang. 2006. p. 107-118. ISBN 3-631-54935-0.

JACOBI, Daniel – La communication scientifique: discours, figures, modèles. Saint-Martin-d'Hères: PUG, 1999. 277p. ISBN 2-7061-0822-3 (Communication, médias et sociétés).

JACOBI, Daniel – Sémiotique du discours de vulgarisation scientifique. Semen, De Saussure aux média [Em linha] 02 (1985). [Consultado a 17/06/08] Disponível na WWW: <http://semen.revues.org/document4291.html>.

L'HOMME, Marie-Claude – La terminologie: principes et techniques. Québec: PUM, 2004. 202 p. ISBN 2-7606-1949-4. (Paramètres).

MEYER, Ingrid – Extracting knowledge-rich contexts for terminography: a conceptual and methodological framework. In BOURIGAULT, Didier; JACQUEMIN, Christian; L'HOMME, Maria-Claude, eds. – Recent advances in Computational Terminology. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2001. p. 279-302. (Coll. NLP) ISBN 1588110168.

MEYER, Ingrid; MACKINTOSH, Kristen – The corpus from a terminographer's viewpoint. International Journal of Corpus Linguistics. ISSN: 1384-6655 1:2 (1996), 257-285.

PEARSON, Jennifer – Terms in context. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1998. 242p. ISBN 1 55619 342 4.

REMÍGIO, Ana Rita – Processo terminográfico: vertentes conceptual, comunicativa e textual. Proposta de uma base de dados terminológica para o consumidor. Aveiro: Universidade de Aveiro, 2010. 606 p. Tese de Doutoramento.

REMÍGIO, Ana Rita – Do processo tradutivo à actividade terminológica: contributo para uma analogia de abordagens. 3Ts – Revista de Tradução, Terminologia e Tecnologias. ISSN 1646-849X n.º 0 (2008) p. 37-51.

REMÍGIO, Ana Rita; ROBERTO, Maria Teresa; COSTA, Rute – A divulgação científica no sector alimentar: os contextos comunicativos e a informação terminológica veiculada. In XI Simpósio Ibero-Americano de Terminologia (RITerm 2008): "A terminologia no terceiro milénio: pela adopção de boas práticas terminológicas" [Em linha] Lima, 2008. Disponível na WWW: http://www.riterm.net/actes/11simposio/Remigio_Ana-Roberto_MT.htm

REMÍGIO, Ana Rita; ROBERTO, Teresa; COSTA, Rute – Organização conceptual e diversidade de contextos de comunicação na construção de um corpus: o caso das Ciências da Nutrição. In Jornada REALITER sobre Metodologia para a Recolha e Sistematização de Corpora para fins Dicionarísticos. [Em linha] Rio de Janeiro, 2006. Disponível na WWW: <http://www.realiter.net/spip.php?article552>.

SCOTT, Mike – Oxford WordSmith Tools version 4.0. [Em linha] Oxford: Oxford University Press, 2006a. 259 p. [Consultado a 11/10/07] Disponível na WWW: <http://www.lexically.net/wordsmith/>.

SCOTT, Mike – Word-lists: approaching texts. In SCOTT, Mike; TRIBBLE, Christopher – Textual patterns: key words and corpus analysis in language education. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2006b. p. 11-32. ISBN 90 272 2294 0.

SINCLAIR, John – Introduction. In HOEY, Michael [et al.] – Text, discourse and corpora: theory and analysis. London/New York: Continuum, 2007. p. 1-5. ISBN 978-08264-9171-8. (Studies in Corpus and Discourse).

SINCLAIR, John – Current issues in Corpus Linguistics. In SINCLAIR, John – Trust the text: language corpus and discourse. London: Routledge, 2004. p. 185-193 ISBN-10 0415317681.