

Roberta Facchinetti
Faculty of Education
University of Verona, Italy
roberta.facchinetti@univr.it

Exploiting Specialized Bilingual Dictionaries as Corpora: The Case of *DITELI*

The aim of the present paper is to discuss the role of dictionaries in corpus linguistics when it comes to viewing them as sources of information; more specifically, would it be too far-fetched to hypothesize the exploitation of dictionaries – and particularly of specialized bilingual dictionaries – as finite corpora? Can a dictionary be considered a full-fledged corpus? If so, exactly what type of corpus is the specialized bilingual dictionary?

In discussing these issues, I will refer mostly to *DITELI* (*DI*zionario *TE*matico *LE*gale Italiano-Inglese / Thematic Italian-English Legal Dictionary), which has been compiled at the University of Verona. Its Italian texts and their English translations are currently being aligned sentence by sentence by means of *ParaConc*, a tool developed by Michael Barlow for parallel corpus compilation and analysis.

In the present study I will highlight the most important aspects that have come to light from the linguistic analysis of the texts. Specifically, on the one hand, the examples provided by *DITELI* will testify to the fact that indeed this dictionary – and possibly any specialized bilingual dictionary – can yield interesting and valuable lexical and semantic information when exploited and browsed as a corpus. On the other hand, having been compiled as a dictionary first, and not as a corpus of everyday speech, its peculiarities cannot be overlooked when analyzing and drawing conclusions from its text samples.

While discussing the above issues, I will also illustrate some of the aspects that have been encountered so far in aligning and visualizing the data, particularly with reference to the difficulty of aligning words and phrases when translating a whole concept with no direct corresponding term in English.

References:

- Alcaraz, Enrique and Hughes Brian. 2002. *Legal Translation Explained*. Manchester: St. Jerome Publishing.
- Bowker, Lynne. 2000. "Towards a methodology for exploiting specialized target language corpora as translation resources". *International Journal of Corpus Linguistics*, 5(1), 17-52.
- Foley, Joseph A. 1996. *J.M. Sinclair on Lexis and Lexicography*. Singapore: UniPress, The Centre for the Arts, National University of Singapore.
- Kraif, Olivier. 2003. "From translational data to contrastive knowledge: Using bi-text for bilingual lexicons extraction". *International Journal of Corpus Linguistics* 8, 1, 1-29.
- Ooi, Vincent B.Y. 1998. *Computer Corpus Lexicography*. Edinburgh: Edinburgh University Press.
- Sinclair John M. 2001. "Data-derived multilingual lexicons". *International Journal of Corpus Linguistics* 6 (special issue), 79-94.
- Teubert, Wolfgang. 2001. "Corpus linguistics and lexicography". *International Journal of Corpus Linguistics* 6 (special issue), 125-153.