

Integration of Terminology Database and Corpus of Translations

Miran Zeljko¹

Abstract

Bilingual dictionaries, glossaries, terminology databases and corpora are the basic tools for translators.

Dictionaries (as well as glossaries and terminology databases) are usually dealt with separately from corpora. However, a corpus can be regarded as an additional database of possible uses of words and phrases, and in this respect a corpus is a natural add-on to a dictionary. Corpora are used during compilation of dictionaries; however, after the dictionary has been compiled, it usually contains a rather limited number of examples from the corpora. Therefore, I believe that the dictionary and the corpus should be integrated not only during compilation of the dictionary but also during its use.

During translation of EU *acquis communautaire* into the Slovene language the Government Translation Service used terminology database and translation memory software. In order to ensure consistent use of terminology, we had to make those data accessible to other participants in this project (outsourcing translators, language-, expert- and legal revisers), so we had to find a solution to make the data available to all of them without their using expensive software. For this purpose we developed a web-based terminology and corpus software (Evroterm and Evrokorpus: <http://www.gov.si/evroterm/index.php?jezik=angl>).

While these databases have been used it became clear that by integrating them, additional features can be obtained: faster search, more search options, more useful results, user friendliness and faster detection of errors.

The paper will address the problems, theoretical aspects, implementation of the integrated terminology and corpus, as well as benefits of this solution. I am aware that dictionary is a much more complex product than a terminology database, but one has to start with simpler tasks. Provided that suitable equipment is available at the conference, a short presentation of the use of software will be given.

¹ *e-mail*: miran.zeljko@gmail.com