OMBI-Dutch-Arabic

(OMBI bilingual resources in XML-LMF format)

1. Introduction

The OMBI bilingual resources were developed under the auspices of the Committee for Lexicographical Translation Resources (in Dutch, *Commissie Lexicografische Vertaalvoorzieningen*, CLVV) under the chairmanship of Prof. Willy Martin. The Committee was set up in 1993 by the Ministry of Education of both Flanders and the Netherlands to improve and stimulate the production of bilingual dictionaries and lexical databases with Dutch as the source or target language.

The CLVV has launched about twenty bilingual dictionary projects, most of which have already been completed. It not only supported these projects financially, but also supplied a lexicographical infrastructure. The Dutch Reference File (*Referentiebestand Nederlands*, RBN) and the Belgian-Dutch Reference File (*Referentiebestand Belgisch-Nederlands*, RBBN) were developed as a basis for the Dutch lexical database. Moreover, the Committee gave the Software Engineering Research Centre (SERC) the assignment to develop an editing tool that would ease the lexicographical process - SERC developed the OMBI dictionary editor (*Omkeerbare Bilinguale Bestanden*, Reversible Bilingual Lexical Databases).

In 2004 the CLVV was replaced by the Advisory Committee for Interlingual Lexicographical Resources (*Adviescommissie Lexicografische Vertaalvoorzieningen*, ALVV) of the Dutch Language Union (*Nederlandse Taalunie*, NTU), which has as its first and foremost goal to supervise the completion of ongoing CLVV projects.

2. Realization of the OMBI-Dutch-Arabic lexical file

The OMBI-Dutch-Arabic lexical file was compiled within the framework of the "*Woordenboek Nederlands-Arabisch, Arabisch-Nederlands, Nijmegen*" dictionary project, conducted between 1998 and 2002 by the Department of Arab and Islamic Studies at Radboud University Nijmegen, under the supervision of Jan Hoogland, Kees Versteegh and Manfred Woidich. The project resulted in a lexical database, used as the basis of a set of bilingual dictionaries published in 2002 by Bulaaq, Amsterdam.

3. Format

The OMBI-Dutch-Arabic lexical file was compiled using the OMBI dictionary editor. In 2007, the Dutch-Flemish HLT Agency (*Centrale voor Taal- en Spraaktechnologie*, TST-Centrale) converted the data to XML in accordance with the standard representation for lexica: Lexical Markup Framework. More information about this model can be found on the LMF website; for information about the conversion see Maks et al. (2008).

4. Relation to other files

4.1. Relation to the Dutch Reference File (*Referentiebestand Nederlands*, RBN)

The RBN is used as a basis for the Dutch component of the OMBI files. Most of the data from RBN is incorporated unchanged into OMBI. In some cases the values are adjusted or left out.

4.2.   Relation to other OMBI bilingual files

The OMBI-Dutch-Arabic, OMBI-Dutch-Danish and OMBI-Dutch-Indonesian bilingual files were developed within the framework of the Committee for Interlingual Lexicographical Resources (CLVV) bilingual dictionary projects and are compiled using the OMBI dictionary editor.

5.   IPR

The property rights to the bilingual files belong to the Dutch Language Union (NTU).

6.   Conditions of use

The files are suitable for use as a component of automatic translation, computer assisted language learning, cross-lingual information retrieval, parallel corpora etc. The data may not be copied for use, for example, in the production of a bilingual dictionary.

7.   Costs

See the website for current price information.