

EmoBase - A platform of multilingual resources for the analysis of emotion lexemes

Sascha Diwersy
Institute of Romance Languages
University of Cologne
sascha.diwery@uni-koeln.de

The present talk reports on EmoBase, a platform of multilingual resources created within the framework of the ANR-DFG funded project EMOLEX, which has been carried out by a group of linguists affiliated to the LIDILEM research lab at the University of Grenoble-Stendhal 3, the Institute of Romance Languages at the University of Cologne and the Institute of English studies at the University of Osnabrück.

Grounded in a multilingual and multidisciplinary approach (intra- and interlingual typology, lexical semantics, syntax, lexicography, foreign language teaching, computational linguistics), EMOLEX has been designed as a corpus-based project pursuing various objectives.

The first aim was to deepen and conjoin the representationist and contextualist approaches to meaning (i. e. meaning \approx concept versus meaning \approx usage).

The second aim was to harness the relevant theories to practical implementation. In order to achieve this we carried out a contrastive analysis of the semantic values, combinatory (lexematic and syntactic) behaviour and discourse roles of the lexemes relating to emotions in five European languages (German, French, English, Russian and Spanish). The research findings aimed at providing a better overview of the lexical field of emotions, thereby making it possible to move beyond existing lexicological and lexicographical studies.

The advocated approach to the lexis of emotion was meant to result in new applications for foreign language teaching, lexicography and computational linguistics. In the context of the EMOLEX project new didactic material has been devised to improve teaching/acquisition of the preferred lexical combinations in the languages dealt with leading to a systematic use of collocations in written and oral text production. At the same time the representation of the combinatory typical to the lexis of emotion has contributed to create a new type of lexicographical microstructures which are able to incorporate sufficient contextual information. Finally, looking at it from a computational linguistic perspective the activities focused on the development of existing applications for the computation and query of frequencies and syntactic structures in large multilingual corpora as well as on the conception of new tools adjusted to the requirements of the project's research questions.

A major achievement of these efforts was the creation of EmoBase as a platform providing access to the compiled resources by means of three main applications:

- (1) EmoConc, a toolbox for lexico-statistical corpus exploration
- (2) EmoLing, a front-end user interface to the database containing the results of lexico-grammatical description
- (3) EmoProf, a framework for the deployment of didactic material related to the lexis of emotion

In our talk, we will outline the features of the EmoBase platform according to various usage scenarios.