Frequency and typological patterning: morphosyntactic properties of multiword expressions in a loanword database

Matthias Heinz

The paper will present analyses based on data extracted from the OIM (Osservatorio degli Italianismi nel mondo) database, which draws on the lexicographical corpus of the Dizionario di Italianismi in francese, inglese, tedesco (DIFIT), a dictionary of language contact describing the diachronic and synchronic impact of loanwords from the donor language Italian in three recipient languages (French, English, German). Its enlargement towards several other contact languages is currently in progress.

The new corpus resource supports lexicological research of various kinds, such as the study of the distribution of lexical categories and morphological patterns across the borrowed units. In the data extant, approximately 11% of the lemmata can be classified as multiword expressions. In many cases the different formation types within this category display morphosyntactic properties and a quantitative distribution that are marked in the target languages (e.g. with regard to adjective position). Interesting lexicological questions are related to tendencies of borrowed multiword expressions to result in certain types of loans. A corpus analysis of the morphology of the borrowings yields a preference for certain types of internal structure. Accordingly, the paper investigates differences in the frequency of the principal morphological patterns between the donor language and the recipient languages, highlighting their respective typological characteristics.

References: