Derivation and multi-word expressions

Verginica Barbu Mititelu, Svetlozara Leseva, Borislav Rizov

The focus of this chapter is the morphological flexibility of multi-word expressions (MWE), manifested as their capacity to create derivationally related items. When subject to derivation, MWEs produce either other MWEs (Bg. *(neshto) spira dăha mi* "(something) takes (one's) breath away" – *spirasht dăha* "breathtaking") or words (Ro. *a face un lucru muşama* (literally: "to make an object oilcloth") "to hide something" – *a muşamaliza* "to cover something up"). In the former case, one content word (usually the syntactic head) of the original MWE is subject to affixation, while the rest of the MWE may or may not be syntactically reorganized, depending on the morphosyntactic characteristics of the original and the derived MWEs: the syntactic reorganization is illustrated by: Ro. *a mustra cugetul (pe cineva)* (literally: "the conscience chides (someone)") – *mustrare de cuget* (literally: "the chiding by the conscience (*of someone)"), while the lack of syntactic reorganization can be supported by: Ro. *minte de îngheaţă apele* (literally: "He lies so much that waters freeze.") – *mincinos de îngheaţă apele* (literally: "[He is such a] liar that waters freeze."). In the latter case, one content word of the MWE is subject to affixation and the obtained word has the semantic content of the original MWE; thus, the other words of the MWE simply do not occur in the result of the derivation.

These types of derivation seem to affect collocations (Bg. *pera pari* "launder (money)" – *prane na pari* "money laundering" – *perach (na pari)* "money launderer" – *perachnitsa (na pari)* "(money) laundry"), terms (Ro. *instalator de gaze* "gas installer" – *gazist*) and idioms (see above) alike. Linguists have described the phenomenon as very frequent (Groza 2011:100) and systematic (idem: 118). However, no quantitative support has been offered for these claims.

MWEs derivation has not been studied extensively, especially with respect to the languages in question, Romanian and Bulgarian. As far as Romanian is concerned, the research on phraseological derivation is scarce: the phenomenon has been recorded and discussed since the 1950s (Iordan 1950, Dimitrescu 1958, Hristea 1984, Căpăţînă 2007, Groza 2011), but no exhaustive or quantitative analysis of phraseological derivation has been made. In the Bulgarian traditional linguistics the derivation involving MWEs has not been studied as such; it has been mentioned in relation to the capacity of some components of phraseological units to acquire a lexical meaning, and derive other words (Blagoeva, 2011). Recently there have been publications

on compounds derivation in Bulgarian (Osenova and Simov, 2014); however, no full account of the subject and the related phenomena has been proposed so far.

The relevance of the proposed cross-lingual study is also substantiated by the fact that Romanian and Bulgarian belong to different language groups of the Indo-European family (Romance and Slavic, respectively), but, being languages in contact, they have developed common linguistic features. In our previous work on derivation (Tarpomanova et al. 2014), we found parallels between the two languages, and it will be interesting to explore the extent of their validity with respect to MWE derivation.

We are interested in identifying and classifying the possibilities of derivation between the morphological classes of the (syntactic) head of the MWEs. We will also propose a typology of the (productive) patterns of syntactic reorganization involved in the derivation, both monolingually and cross-lingually. As a further step, we will focus on studying the semantics of derivation and analyse the types of semantic and derivational patterns.

Our aim will be the identification of similarities and differences in the behaviour of MWEs in the languages under study, Romanian and Bulgarian, with respect to derivation and its types. We will also compare our results with the ones reported for Polish (Piela 2007) and Turkish (Oflazer et al. 2004) and will be open to collaboration with researchers who can report on other languages.

The foreseen working methodology for this chapter consists of the following steps: (1) a comprehensive list of MWEs will be extracted from existing lexicographic resources (big explanatory monolingual dictionaries or dictionaries of idioms); (2) a shallow syntactic analysis of these MWEs will be done; (3) using either our resources of derivationally related words or automatic derivation based on the derivation models that can be extracted from our resources, we will be able to: (i) identify the derivationally related MWEs extracted at step (1); (ii) automatically create possible MWEs whose syntactic head is derivationally related to the syntactic head of the target MWE. The real existence or use of the derivationally created MWEs will be checked against our big national corpora and, in case no results can be obtained from them, against the Internet; (4) cross-lingually classify the pairs of base MWE – derived MWE according to their morphological classes; (5) identify and (if possible, cross-lingually) classify the types of syntactic reorganization of the MWE implied by the derivation; (6) analyse and classify the semantics of the derivational patterns.