

## Lexical combinations in specialized resources

Marie-Claude L'Homme  
Observatoire de linguistique Sens-Texte (OLST)  
Université de Montréal  
[mc.lhomme@umontreal.ca](mailto:mc.lhomme@umontreal.ca)

### Topic 1: Understanding specialized lexical combinations

Terminologists are concerned with the description of single-word and multi-word units that label concepts precisely within a specialized subject field (such as *bus*, *data*, *operating system* in computing, or *carbon dioxide*, *contaminant*, *greenhouse effect* in the field of the environment). Standard descriptions of terms comprise data categories that help understand concepts (a definition, usage notes, relations between concepts) but very little linguistic information. However, users of terminological resources often need to access information about how to use terms in specialized texts: for instance, how to combine them with other terms or lexical units (*absorb a contaminant*, *discharge a contaminant*, *toxic contaminant*; *enter data*, *store data*, *critical data*). For this reason, terminologists have started taking into account combinations in which terms appear and devised methods to collect them from text and encode them in specialized resources. We will examine some of this work and address the following questions:

- Which combinations are relevant for specialized resources?
- Are specialized combinations different from “general language” combinations?
- What do combinations tell us about terms?
- How can specialized combinations be added to specialized resources such as term banks?
- What are the theoretical and methodological consequences of taking into consideration combinations for terminology?

### Topic 2: Advanced encoding of specialized combinations in resources

The first attempts to add combinations to specialized resources (terms banks and printed dictionaries) were made in the 1980s. Since then, different proposals were made in order to provide an efficient access to specialized combinations based in different parameters: the part of speech of collocates, meanings of collocates, translations of combinations, etc. We will examine a method for encoding specialized combinations that takes into account their linguistic features. The method is based on lexical functions, a system devised by Mel'čuk and his collaborators (1988; 2012) for collocations in general language, but that can be adapted to specialized combinations in which terms appear. We will show how lexical functions were used to encode combinations in two resources: the first one contains terms linked to the field of the environment; the second one comprises terms in the domains of computing and the Internet. We will then show that this encoding can improve the access to combinations since it allows to classify them semantically and can be used to retrieve translations of combinations.