# CORPUS DATA AND FRAME SEMANTICS

Rema Rossini
*University of Bologna (Italy)*
*rema.rossini@unibo.it*

In this paper collocation extraction is focused as instrumental to frame description. Repetitiveness in word co-selection and co-occurrence is seen as factual evidence of what is typical and routine in language use and makes it possible to investigate the frames evoked by individuals and shared by their linguistic community.

Under this perspective reference is made to Firthian and Neofirthian approach on the one hand and to Minky's and Fillmore's studies on the other to note the common elements that can be identified in the collocation description and in frame elements identification.

In Firth's contextual meaning theory, the term collocation is systematically introduced to denote recurrent or "habitual" word combinations where words are "mutually expectant and mutually prehended". It is emphasised the role of contextual relations in the constitution of a word meaning.

Minsky is referred to for his description of the notion of frames as associated to stereotyped situations. According to Minsky, the "top levels of a frame are fixed and represent things that are always true about the supposed situation". At the lower levels the structure of frames may be well represented in terms of terminals, fillers and default values. Default values are described as loosely attached to their terminals. They can be replaced by new items that "fit better" the new situation.

In Fillmore's approach the epistemic links that connect a term with a set of beliefs are structured by a general knowledge frame that serves as a set of default assumptions for the term understanding. Frame core elements are seen as normal, "persisting" part of a frame. They coexist with optional, non–persisting elements. Frames are described as frameworks of concepts or terms which are linked together as a system and give structure or coherence on aspects of human experience. They contribute to define the role concepts evoked by linguistic forms to which they are conventionally associated and constitute the underlying, implicit, conceptual structures invoked by individuals.

A word's meaning is understood with reference to a structured background of experience, beliefs or practices which are made accessible through culturally shared cognitive schemata. A common ground can be found in the elements occurring in the description of collocations, on the one hand, and in the characterizing aspects of frame description, on the other. Conventionality, repetitiveness, and shared experience seem to result as common characterizing aspects.

To illustrate these views some case studies are presented considering some concordances in a synchronic corpus of written Italian – CORIS/CODIS, a120-million word reference corpus – and in a diachronic one – DiaCORIS. Both i) words with little or no semantic ambiguity; and ii) polysemous words with a high degree of semantic ambiguity structure will be examined.

Interestingly, the polysemy of the word is disambiguated by the representation of its collocates. Their constellation contributes to distinguish its different meanings and to pair a word with just one of its meanings. Collocates result to pave the way to capture the links connecting lexical units and the frames they evoke. In Fillmore's terms, lexical units are distinguished and identified.