A contrastive analysis of prepositional post-copular Multi Word Units in French, Norwegian and Polish

Authors:
Agnieszka Kaliska, Adam Mickiewicz University, Poznan, Poland
Harald Ulland, University of Bergen, Norway

Theoretical background
The typical constituent appearing in the post-copular position of a copular clause is normally an adjective phrase or a noun phrase:
- John is very intelligent
- John is a teacher
But we can also have a prepositional phrase (PP) in this position:
- John is on his horse
A PP occurring in this position will very often be a MWU, having many semantic characteristics in common with a simple adjective or an adjectival phrase.

Following Denton (1985), Dross (1996) and Machonka (1988), we will consider the copula of such clauses as a "support verb", and the PP as a prepositional predicate.

NLP background
- Prepositional predicates have been included in the manually constructed electronic dictionary for French in the Unitex corpus processor.
- They are POS-marked as adjectives (A), and they are also subcategorized according to formal complexity.
- We would like to include MWUs of this kind also in the Norwegian and the Polish dictionaries of Unitex. In doing so, we want to use the same criteria of selection for the three languages.
- Furthermore, we want to investigate the possibility of adding transitional equivalents for each entry in each dictionary, so that an entry in the Polish dictionary, for instance, would have a translation proposal into French and Norwegian, and likewise for the predicational predicates entries in the French and the Norwegian dictionaries.

Prepositional post-copular MWUs
The PREP C sequence is lexicalized and partially or totally formalized, i.e.:
- (a) it has a figurative meaning, separate from the literal meaning, and/or
- (b) it cannot be transformed in a way that the similar non-lexicalized constructions usually can be;

However, the borderline between free combinations and MWU is often difficult to define:
- (a) the PREP C may be considered as being always idiomaticized as far as its locative meaning is figurative;
- (b) a semi-lexicalized PREP C can accept some syntactic manipulations like modifiers insertion if only it corresponds to the internal selection rules;
- (c) an adjectival internal to the PREP C can sometimes be replaced by its antonym (good vs. bad).

Formal characteristics and classification of the BE PREP C MWUs
- The BE PREP C structure contains three basic components:
  - (1) the verb to be (BE)
  - (2) a preposition (PREP)
  - (3) a complement (C)
- The complement may:
  - be determined by an ADJ: e.g.
  - (adjective) + (datum) in French, subject + object in Polish
  - or another prepositional complement, e.g.:
  - (adverb) + (adverb) in Polish
- Each of the components may accept other expansions which make possible a form-based classification of the BE PREP C constructions.

Aspectsual extensions of the copula
Potential co-occurrence with the standard copular verb in each language (être, by, være) is a prerequisite of selection of the items to be included in our lists.
In many cases, the selected PREP C units can also be used with other verbs that could be interpreted as aspectual extensions of the standard copula.
If the standard copular verb is only imperfective (and that is probably in any language), its aspectual extension in Polish usually manifests itself in two ways: perfective and imperfective. Both variants will have a single equivalent in other non-aspectual languages.

Aspectsual extensions of the copula
Inchoatives:
- Fr.: Luç entre dans une colline noire
- No.: komme, gå til, at
- Pl.: znalaz się, wchodzi w namiot, przechodzę przez płot, it al.
Terminatives:
- Fr.: Luç sort du coin
- No.: komme av
- Pl.: (T)
Static:
- Fr.: Luç nage dans le bonheur
- No.: är digit, skilt, etc.
- Pl.: pozostaj w pozostaju, zostaje w pozostaniu.

Dictionaries of MWUs using the DELAF formalism (example)

References