

Working group 1: Lexicon/Grammar Interface

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

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

The aim of the present poster presentation is to provide a classification of the idiomatic predicates based on the *BE PREP C* structure, in French, Norwegian and Polish. 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 adjective (*dobry*[ADJ] + *data*[N] in *być pod dobrą data* in Polish) or another prepositional complement (*takt*[N] + *med*[Prep] *tiden*[N] in *være i takt med tiden* in Norwegian). Each of the components may accept other expansions which make possible a form-based classification of the *BE PREP C* constructions:

Class	Explication	Example FR	Example PL	Example NRV
EPC	BE PREP C (without determiner)	<i>être à cran</i> [be edgy]	<i>być w modzie</i> [be in fashion]	<i>være på vakt</i> [be carefull]
EPDETC	BE PREP DET C (with determiner)	<i>être à l'abandon</i> [be abandoned]	<i>być po paru głębszych</i> [be drunk]	<i>være uten en tråd</i> [be without a stitch]
EPAC	BE PREP ADJ C (prepositive adjective)	<i>être en plein boum</i> [be booming]	<i>być w dobrym tonie</i> [be convenient]	<i>være i full gang</i> [be in full swing]
EPCA	BE PREP C ADJ (postpositive adjective)	<i>être devant un choix difficile</i> [be faiced with a difficult choice]	<i>być na porządku dziennym</i> [be on the day's order]	-
EP CDC	BE PREP C GEN C (genitive complement)	<i>être à l'abri du vent</i> [be alee]	<i>być na ustach wszystkich</i> [be on everyone's lips]	<i>være på alles lepper</i> [be on everyone's lips]
EPCPC	BE PREP C PREP C (additional PREP C)	<i>être en avance sur son âge</i> [be advanced for one's age]	-	<i>være i takt med tiden</i> [be up to date]
ENEGPC	(BE PREP C) NEG (negation as part of the idiom)	<i>ne pas être dans son assiette</i> [not be feeling well]	<i>być nie w humorze</i> [not be feeling well]	<i>ikke være av veien</i> [not be out of the way]
EPACDC	BE PREP ADJ C GEN C (adjective and genitive complement)	-	<i>być po drugiej stronie barykady</i> [be on the other side]	-
EP CONJ C	BE PREP C CONJ C	<i>être entre le</i>	<i>być między młotem a</i>	<i>være mellom</i>

	(conjunction)	<i>marteau et l'éclume</i> [be caught between a rock and a hard place]	<i>kowadlem</i> [be caught between a rock and a hard place]	<i>barken og veden</i> [be caught between a rock and a hard place]
--	---------------	---	--	---

In each of the cases, the whole sequence should be idiomatized, 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-idiomatized constructions usually can be. The former feature usually comes with the latter one but the latter one doesn't need to come with the former one. Free combinations of (BE) *PREP C* will not be included in our classification, only sequences of words that can truly be considered as multi-word units. However, the borderline between free combinations and multi-words units is often difficult to define, and will thus be a matter of discussion in our poster presentation.

Potential co-occurrence with the standard copular verb in each language (*être, być, 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, usually a verb that could be interpreted as an aspectual extension of the standard copula (e.g. French inchoative *tomber en panne* coexisting with the aspectual neutral *être en panne*). The post-copular *PREP C* part could be considered as the predicate of the sentence, having one or two arguments. In the first stage of our project we will concentrate on *PREP C* predicates with a single argument (which will normally be the syntactic subject of the sentence). In a second stage, we will also include *PREP C* constructions with two arguments, like Norwegian *være på utkikk etter*).

The classification which has been proposed for French by Maurice Gross 1996, and which is used in the corpus processor Unitex, will be the starting point of our classification. However, both the number and the characteristics of the distinguished classes had to be modified to fit the Norwegian and Polish language systems, which haven't been so far analyzed in the proposed way. As for other languages, studies have been made for English (Machonis 1988), Portuguese (Baptista 2005) and Spanish (Mogorrón Huerta 2000).

We believe that an important characteristic of this project, as well as its advantage, is that the three analyzed languages – French, Norwegian and Polish – are genealogically and typologically different.

Possible applications of our project could be: (1) The creation of electronic dictionaries containing MWUs of this kind, using the DELAF formalism described in the Unitex manual, to facilitate automatic recognition in corpora, (2) Trilingual MWU dictionaries with entries including translational equivalents like

de retour, .A+EPC + TransNorw *tilbake* + TransPol *z powrotem*
på plass, .A+EPC + TransFr *en place* + TransPol *w toku*
w modzie, .A+EPC + TransFr *à la mode* + TransNorw *på moten*

which could be useful for computer assisted translation.

References:

- Baptista, 2005. *Sintaxe dos Nomes Predicativos com verbo-suporte SER DE*. Lisboa: Fundação para a Ciência e a Tecnologia/Fundação Calouste Gulbenkian.
- Gross, 1996. Les formes *Etre Prép X* du français. *Linguisticae Investigationes* 20:2, Amsterdam-Philadelphie, John Benjamins.
- Insalata, Sabrina. 1989. Constructions *essere Prep X* en italien, Mémoires du CERIL 5, Université Paris 7, 138-190.
- Machonis, 1988. Support Verbs: An analysis of *be + Prep + X* idioms, *The SECOL Review* 12:2
- Marques-Ranchhod, Elisabete. 1988. *Construções nominais com verbo-suporte Estar : nominalizações e nomes autônomos*, Thèse de doctorat, Universidade de Lisboa.
- Mogorrón Huerta, 2000. Étude comparée et traduction des expressions figées des verbes *ser/estar* suivies de *Prép X* et de *être Prép X*. In María Luz Casal Silva et al. (eds.), *La lingüística francesa en España camino del siglo XXI*.
- Moustaki, Argyro. 1990. Les expressions figées *Nó εἶμαι Prép X W* en grec moderne, *Mémoires du CERIL* 7, Université Paris 7, 203-271.
- Unitex corpus processor: <http://www-igm.univ-mlv.fr/~unitex/>
- Vietri, Simonetta (1996) The Syntax of the Italian Verb *essere Prep*, *Linguisticae Investigationes* XX : 2, 287-363.