

A systematic study on the fixedness degree of verbal multiword expressions: application to Modern Greek and French

Mathieu Constant & Angeliki Fotopoulou

Université Paris Est-LIGM & Alpage INRIA France - ILSP/R.C “Athena” Greece

Mathieu.constant@u.pem.fr

afotop@ilsp-athena.innovation.gr

WG1

Abstract

Multword expressions are complex linguistic phenomena (Sag et al. 2002) as they have multidimensional properties and varying degree of compositionality (Nunberg et al. 1994). This poster aims at presenting a preliminary work on the identification of the fixedness degree of multiword expressions (MWE).

1. Introduction

We specifically focus on Modern Greek and French as lexical MWE database exist with a larger coverage of encoded syntactic semantic properties of verbal MWE that are organized in the form of a lexicon-grammar (M. Gross 1986). Our study is built on the work of M. Gross (1984) for French and the work of Fotopoulou (1993) and Mini (2009) for Modern Greek.

2. Linguistic data – Method- Criteria

For our study, we constructed a small set of 65 verbal multiword expressions for both languages. The extracted expressions have different syntactic structures in order to test the fixedness degree with respect to a series of tests. These tests are those which, in principle, enable to define a fixed expression like more and more studies show such as in Vincze (2011), Saittureer and Wintner (2014), Stone (2015):

- Lexical criteria:** Fixedness can be identified by testing whether there exists a paradigmatic break. These criteria enables to evaluate the exclusive co-occurrence of expression components.

Max a cassé (les pieds +*le jouet+*le verre) à Marie [FR]
Max broke the (feet +*toy+*glass) to Marie
Max gets on Marie's nerves'

- Morphosyntactic criteria:** non-regular restrictions apply on the determiner distribution or on the morphological variants (e.g. number), as well as on some transformations like passivation or pronominalization.

Μου κόπηκαν τα ήπατα [GR]
me.PRO-GEN cut.V-3-PL-PASS the.ART-DEF liver.N-PL-ACC
was very frightened

* Μου κόπηκε το ήπαρ
me.PRO-GEN cut.V-3-SG-PASS the.ART-DEF liver.N-SG-ACC

- Semantic criteria:** the meaning of the expression is non compositional, i.e. it is not predictable from the meaning of its components. Similarly to Mini et al. (2011), we examined independently the verb and the nominal arguments of the expressions with tests like

- the element keeps its literal meaning
- the element has a metaphorical meaning, or is an extension of the literal meaning
- the element has nothing to do with the literal meaning

Luc nage dans le bonheur [FR]
Luc swims in the happiness
Luc is happy
[nage (swims) -> case (iii) ; bonheur (happiness) -> case (i)]

μου ράγισε την καρδιά [GR]

me.PRO-GEN broke.V-3SG the.ART-DEF heart.N-ACC

he broke my heart

[ράγισε (broke) -> case (iii) ; καρδιά (heart) -> case (ii)]

All expressions are encoded in the form of a table: a row corresponds to a lexical entry (a multiword expression), a column corresponds to a lexical, syntactic or semantic property like in M. Gross (1986).

3. Discussion - Results

Results show a large variety of behaviors. In general, we can observe a correlation between the three types of criteria in the extreme cases of the continuum between entirely non-compositional and almost free expressions. We also observe a grey zone with some unexpected behaviors. For instance, in French, *casser du sucre sur le dos* (break some sugar on the back, meaning talk about someone behind her/his back) accepts passivation.

Let's now examine our results with the perspective of Mini et al. (2011). Grammatical and semantic tests examine :

- exclusive cooccurrence between the verb and its complements (nominal elements) ;
- the meaning of the verb and its complements (literal meaning, metaphorical meaning), and the global meaning of the expression.

The combinatorics of these criteria led to two groups : **compositional** and **non-compositional expressions**.

Compositional expressions can be divided in two classes : **quasi-typical** and **conventional expressions**. In quasi-typical expressions, the verb keeps its meaning combining with a metaphorical element or the verb has a metaphorical meaning by extension from its literal or concrete meaning such as in η καρδιά μου ["my heart cracked"], the verb ράγισε retains the core meaning of "cracking without being cut into pieces" that it has in the sentence ράγισε το ποτήρι ["the glass cracked"] by a semantic extension to a non-tangible/abstract level. On the other hand, the noun καρδιά ("heart") is a semantically autonomous constituent of the phrase since it refers to the inner emotional world. η καρδιά μου ["my heart cracked"]. Conventionalized phrases are those whose verb maintains its fundamental meaning whereas the noun has a parallel to the basic-literal meaning which is a conventionalized meaning. A conventionalized meaning is a meaning widely accepted by a linguistic community and used with all possible connotations acquired in the course of time. For example, in the phrase τον τρώει το σαράκι της ζήλειας [literally "the woodworm of jealousy eats him"] meaning "jealousy wears him out", the verb eat retains the core meaning of wear. What differentiates this phrase from the phrase το σαράκι τρώει το ξύλινο τραπέζι [literally "the woodworm eats the wooden table"] is that what the phrase τον τρώει το σαράκι της ζήλειας describes is realized on an abstract level, whereas its consequences are perceived on a tangible level. In other words, one can probably recognize somebody who is jealous, but one can hardly see him "eaten alive" by jealousy.

This is a first coarse-grained classification with few elements yet. We need to extend the coverage by adding new entries for both languages.

V	PREP	DET	Adj	N	LEXICAL		MORPHOSYNT				SEM			GLOSS	TRANSLATION
					V	N	flex. det	V ell.	N ell.	pass.	verb	noun	global		
trembler	de			peur	+	+	-	+	+	/	0	1	0	shake of fear	be terrified
remuer		l'		âme	+	+	+	+	+	+	0	1	0	stir the soul	touch
casser		les		prix	-	-	+	-	-	+	-1	1	0	break the prices	cut prices
briser		le		coeur	-	+	+	+	+	+	-1	0	0	break the heart	break one's heart
monter	à	la		tête	-	+	+	-	-	/	-1	0	0	climb to the head	go to one's head
monter	sur	POSS	grands	chevaux	-	-	POSS	-	-	/	-1	-1	-1	climb on his big horses	get on POSS high horses
casser		les		pieds	-	-	-	-	-	-	-1	-1	-1	break the feet	get on one's nerves

V	PREP	DET	Adj	N	LEXICAL		MORPHOSYNT				SEM			GLOSS	TRANSLATION
					V	N	flex. det	V ell.	N ell.	pass.	verb	noun	global		
γελάει και		τα		μουστάκια του Νh	-	-	-	-	-	/	1	0	0	Even his moustache laughs	He is delighted
ανοίγω				δρόμο (για)	-	-	+	-	-	-/erg	1	0	0	The road is open for ...	There are no obstacles left for...
χάνω		τα		λόγια Poss-0	+	-	-	-	-*	+	0	1	0	N loses his words	Be speechless (with emotion)
πάγωσε		το		αίμα Poss-0	-	-	+	-	-	-	0	0	0	my blood froze	I was terrified
πλέω	σε			πελάγη ευτυχίας	-	-	-	-	-	/	-1	0	0	sail in seas of happiness	to be very happy
φόρτωσα τα	σε	τον		κόκορα	-	-	-	-	-	/	-1	-1	-1	I loaded them on the rooster	I did not act at all, as I was feeling lazy
πέφτει δεν				καρφίτσα	-	-	-	-	-	/	-1	-1	-1	a pin doesn't drop	there are a lot of people crowded'