

MWEs and the Emotion Lexicon: typological & cross-lingual considerations

Angeliki Fotopoulou and Voula Giouli

Institute for Language and Speech Processing, "Athena" RIC

{afotop;voula}@ilsp.athena-innovation.gr



Ερευνητικό Κέντρο Αθηνών
Athena Research Center

Abstract

We report on work in progress aimed at studying MWEs predicates denoting emotion in Greek and French. After defining a typology of emotion-denoting MWEs in Greek, we attempted their mapping onto their French counterparts. Our aim is to establish correspondences between the lexicons of emotions in Greek and French in view of extracting useful generalizations over the two languages that will be useful not only from a purely linguistic point of view, but also in various NLP applications.

1. Introduction

We present work aimed at studying predicates that pertain to the semantic field of emotions the focus being on MWEs and their syntactic and semantic properties. A core lexicon of nominal and verbal MWEs denoting emotion is extracted from existing lexical resources and corpora have been manually updated and extended; MWE classification has been attempted on the basis of syntactic, selectional and semantic properties. An attempt to map the typology identified onto French constructions has also been performed.

2. Purpose - Rationale

It has been widely accepted that MWE identification plays a crucial role in various NLP applications including Machine Translation, or lexical Sentiment Analysis and opinion mining, etc.

From another perspective, the cross-lingual study of MWEs that pertain to a certain semantic field is expected to reveal a number of generalizations across data of different languages.

Our objective is to create a Lexicon of MWEs denoting Emotion for Modern Greek. Currently, verbal MWEs and Support Verb Constructions (SVCs) have been accounted for; correspondences with French MWEs will be identified at a later stage.

3. Background

The present study builds on previous work aimed at the systematic treatment of nominal and verbal predicates denoting emotion in Modern Greek and on studies on MWEs typologies for the two languages (EL-FR). More precisely:

A typology of French verbal MWEs has been defined in M. Gross (1982) and a typology of MG verbal MWEs with the same principles in (Fotopoulou, 1993, Mini, 2009) (20 classes, 6000 phrases, a combination of fixed and non-fixed constituents in Subject or Object position is permitted).

Moreover, French predicates denoting emotion and their combinatorics have been studied in Leeman (1991), M. Gross (1995), and Tutin et al (2006), (Balibar-Mramti, 1995).

Along the same lines, predicative nouns denoting emotion (Nsent) were identified and their combinatorics with support (or light) verbs were identified in MG (Pantazara et al 2008; Fotopoulou et al 2009).

4. MWEs denoting emotion: the data

The initial list of Greek MWEs denoting emotion was extracted from the afore-mentioned lexical resources and was then manually updated and extended on the basis of corpus evidence.

More precisely, a set of noun (single-word) predicates denoting emotion (Nsent) in Greek has been identified and their combinatorics with specific verbs have been extensively described (Pantazara et al 2008; Fotopoulou et al 2009). Of these, Support Verb Constructions (SVCs), collocations and Nominal MWEs were manually identified and used as the basis for the present study. This material was further supplemented with verbal MWEs that were systematically encoded as paraphrases of verbal emotion predicates (Vsent) (Giouli and Fotopoulou 2014); for these, a number of features (as for example, syntactic structure, selectional restrictions,

control, and their combinatorics) were also defined and encoded formally.

Moreover, the French empirical data were enriched with translations of the Greek MWEs or sentences extracted from on-line sources (google.fr) where needed; these data were aligned to the Greek MWEs.

5. Classification of MWEs denoting emotion-EMOTION_LEXICON

At the next stage, MWEs (all types: fixed -extracted from typologies light verbs constructions...) classification was performed along the following axes: (a) grammatical information (part of speech), (b) syntactic configuration of the expression, (c) selectional restrictions imposed over their Subjects and Objects (+hum, -hum), (d) lexical choice, and (d) semantic information dealing with aspect, intensity of the emotion (low, medium and high), control, manifestation of verbal expression, etc.

MWEs with their properties are grouped under concepts as *θυμός* (anger) .

6. Cross-lingual considerations

In this study, we aimed at defining the typologies of emotion MWEs in Greek and French and at finding cross-lingual analogies, the focus being on verbal light/support verb constructions and their variants (intensity, aspect, expression of emotion). More precisely, the syntactic and lexical properties of these expressions were systematically examined.

For example, the comparative study of Greek and French light/support verb constructions shows that they generally share the same properties, in terms of (a)syntactic structure, namely:

N0hum V Prep Nsent

EL. Ο αδερφός του **κοκκινίζει αν'** το **θυμό** του
FR. Son frère **rougit de colère**
EN. His brother turns red of (his) anger

Nsent V Prep (Npc Nhum-gen)

EL. Το **μίσος** φωλιάζει (στην καρδιά + στην ψυχή + *στα μάτια) του Νίκου
FR. La **haine** niche dans (le cœur + l'âme + les yeux) de Nikos
EN. Hate nests in (the heart + the soul + the eyes) of Nikos.

(b) **lexical choice** of the (support) verb. In terms of the underlying semantics, in both languages, the verb selected, ultimately, functions as an intensifier of the emotion denoted:

N0 hum Vsup Nsent

EL. Η Άννα (νιώθει + αισθάνεται) **χαρά**
FR. Anna (éprouve + ressent) de la **joie**
EN. The Anna feels happiness (=Lena is happy).

N0 hum V Prep Nsent

EL. Η Άννα πετάει από (χαρά + τη χαρά) της
FR. Anna rayonne de **joie**
EN. Anna jumps out of (joy - her joy).

In this respect, the focus is placed on the aspects of fixedness and the effect of lexical choice over the feature intensity of the emotion denoted, elaborating further on cross-lingual considerations. The examination of nominal structures shows that:

□ fixedness and intensity have an impact on our expressions that present a structural and lexical analogy (Moustaki et al, 2008)

□ figurative expressions (which are particularly interesting because they are very close to collocations), often present fixedness (both in Greek

and in French); i.e., there is no possibility to modify any of the elements of the sentence.

□ This fixedness is irrelevant to semantic transparency.

For example :

N0hum V Prep C Nsent

EL. **Πλέω** σε **πελάγη** (ευτυχίας + *αισιοδοξίας + *χαράς + *έρωτα)
FR. Je nage dans (le bonheur + optimisme + joie + amour)
EN. I'm sailing on seas of (happiness + optimism + joy + love)

N0hum V Prep Nsent

EL. ***Πλέω** (στην ευτυχία + στην αισιοδοξία + στη χαρά + στον έρωτα)
FR. Je navigue dans (le bonheur + l'optimisme + la joie + l'amour-éros)
EN. I'm sailing on (happiness + optimism + joy + love)

N0hum V Prep C Nsent

EL. ?***Κολυμπάω** σε πελάγη (ευτυχίας + αισιοδοξίας + χαράς + έρωτα)
FR. Je nage dans les mers (de bonheur + d'optimisme + de joie + d'amour-éros)
EN. I'm swimming in seas of (happiness + optimism + joy + love)

N0hum V Prep Nsent

EL. ?***Κολυμπάω** (στην ευτυχία + στην αισιοδοξία + στη χαρά + στον έρωτα)
FR. Je nage dans (le bonheur + l'optimisme + la joie + l'amour-éros)
EN. I'm swimming in (happiness + optimism + joy + love)

7. Conclusions

In the paper we will further elaborate on the classification of Greek and French MWEs denoting emotion on the basis of pre-defined criteria (syntactic and semantic properties, degree of fixedness and lexical choice), and on the generalizations observed during the comparative study performed.

References

- Sag, I. A., T. Baldwin, F. Bond, A. Copestake and D. Flickinger. 2002. Multiword Expressions: A Pain in the Neck for NLP. LinGO Working Paper No. 2001-03. In Alexander Gelbukh, ed., (2002) *Proceedings of COLING-2002*. Springer.
- Balibar-Mrabet, A. 1995. Une étude de la combinatoire des noms de sentiment dans une grammaire locale. *Langue Française* 105, Paris : Larousse, 88-97.
- Fotopoulou, A., Mini, M., Pantazara, M., Moustaki, A. 2009. La combinatoire lexicale des noms de sentiments en grec moderne. Le lexique des émotions, Iva Novacova & Agnes Tutin (eds), ELLUG, Grenoble.
- Fotopoulou A. (1993). Une classification des phrases à compléments figés en grec moderne - étude morphosyntaxique des phrases figées. *Thèse de Doctorat*. Paris: Université Paris VIII - St. Denis.
- Gross, G. 1996. Les expressions figées en français. Noms composés et autres locutions. Paris/Gap: Ophrys.
- GROSS M. (1982). Une classification des phrases "figées" du français. *Revue Québécoise de Linguistique*, 11/2, 151-185.
- Gross, M. 1995. Une grammaire locale de l'expression des sentiments. *Langue Française* 105. Paris: Larousse. 70-87.
- Mathieu, Y.Y. 1999. Les prédicats de sentiment, *Langages* 136, Paris : Larousse.
- Mini M. 2009. *Linguistic and Psycholinguistic Study of Fixed Verbal Expressions with Fixed Subject in Modern Greek: A Morphosyntactic Analysis, Lexicosemantic Gradation and Processing by Elementary School Children*. Unpublished doctoral dissertation. University of Patras
- Moustaki, A., Pantazara, M., Fotopoulou, A., Mini, A. 2008. Comment traduire les noms d'émotion: étude contrastive entre le grec moderne et le français. DISCOURS
- Leeman, D. 1991. Hurler de rage, rayonner de bonheur: remarques sur une construction en de Langue française 91. Paris : Larousse, 80-102.
- Pantazara, M., Fotopoulou, A., Moustaki, A., Mini, M. 2008. La description des noms de sentiments du grec moderne. In *Linguisticae Investigationes* Volume 31:2(2008), 323-331, John Benjamins, Paris, France.
- Wehrl, E. 1998. *Translating Idioms*", COLING-98, Montréal, 1388-1392.