Chapter 3

MWEs and the Emotion Lexicon: Typological and cross-lingual considerations

Aggeliki Fotopoulou

Institute for Language and Speech Processing, Athena RIC, Greece

Voula Giouli

Institute for Language and Speech Processing, Athena RIC, Greece

The work presented in this paper is aimed at studying predicates that pertain to the semantic field of emotions, the focus being on Modern Greek verbal multiword expressions (verbal MWEs) and their counterparts in French. A core lexicon of verbal MWEs denoting emotion was extracted from existing Modern Greek lexical resources; the initial list was further extended and revised manually in view of corpus evidence. A classification of MWEs is proposed based on syntactic, selectional and semantic properties; an attempt to map the expressions identified onto their French counterparts was also made. The cross-linguistic study reveals similarities and discrepancies in the two languages, and highlights the interaction between MWEs structure and their underlying semantics, in that the intensity of the emotion denoted and the degree of fixedness of the relevant expressions seem to be highly correlated in both languages.

1 Introduction

The availability of user-generated content over the web and the increasing need to make the most out of it has brought about a shift of interest from factual information to the identification of subjective information (as opposed to facts) expressed by people or groups of people with respect to a specific topic. To this end, the task of determining the so-called *private states* (that is, beliefs, feelings,

and speculations) expressed in running text and the entities involved has been the focus of attention in the field of Natural Language Processing (NLP). Therefore, identification of expressions denoting emotion or emotional state in textual data and their classification is of paramount importance. In this respect, MWEs can hardly be overlooked since they constitute a significant proportion of the emotion lexicon.

We hereby present work aimed at treating verbal multi-word predicates that pertain to the *semantic field of emotions* from a cross-lingual perspective and systematising their lexical, syntactic and semantic properties. In this context, verbal MWEs in Modern Greek denoting emotion or emotional state were selected from existing language resources. Their lexico-semantic properties were also retrieved from these resources and new entries were encoded following the same principles. All MWEs were further assigned *semantic features* inherent to the semantic field. At the next stage, their mapping onto their counterparts in French was performed. The comparative study of Greek and French MWEs resulted in the identification of cross-lingual similarities and discrepancies. Moreover, correlations between lexical features and the underlying semantics of MWEs were also revealed. Our working hypothesis was that despite idiosyncrasies, MWEs that belong to a given semantic class share features that are characteristic for this class; moreover, these field-specific features are attested cross-linguistically. One step further, the (cross-lingual) treatment of MWEs might be useful not only from a purely linguistic point of view but also for NLP applications.

The paper is outlined as follows. An overview of background work on the study of the emotion lexicon and of MWEs is presented in §2, §3 outlines the methodological framework adopted, whereas the selection process of the lexical data is described in §4. The lexicon of emotion MWEs and the syntactic, selectional and semantic properties encoded are presented in §5; we discuss our findings in §6 and elaborate further on cross-lingual considerations in §7. Finally, our conclusions and prospects for future research are outlined in §8.

2 Background work

The seminal work at the syntax-semantics interface by Levin (1993) involves large-scale classification of English verbal predicates on the basis of shared meaning and syntactic properties. In this work, more than 3000 verbs were grouped into semantically coherent verb classes, each depicting a syntactic configuration that reflects verb meaning. A more fine-grained semantic classification of French verb and noun predicates denoting feeling, emotion and psychological

states has also been performed (Mathieu 1999; 2005), aimed at a wide range of NLP applications. French nominal and verbal predicates denoting emotion and their lexicalised word combinations have been studied (Leeman 1991; Gross 1995; Balibar-Mrabti 1995; Tutin et al. 2006) from a different point of view. Finally, a comparative analysis of English and French single-word verbal predicates denoting emotion (Mathieu & Fellbaum 2010) reports on properties shared among the two languages on the grounds of syntax and semantics, unveiling at the same time the idiosyncrasies of each language.

As far as MWEs are concerned, a systematic treatment of French fixed expressions has been carried out (Gross 1982). In this work, the classification and the analysis of c. 20000 French verbal MWEs consists of the formal representation of their syntactic properties, selectional restrictions and the distinction between fixed and non-fixed constituents. Along the same lines, the classification of Greek fixed expressions (c. 6000 entries) has been performed based on the same formal principles and criteria (Fotopoulou 1993b; Mini 2009).

The present study is part of a larger effort aimed at developing lexical resources that encompass the Greek emotion lexicon, i.e., words and phrases that refer to emotional states and emotion-related mental events. Previous work involves treatment of nouns and verbs. In this context, 130 Greek noun predicates denoting emotion (*Nsent*) were identified and classified on the basis of the verbs' syntactic, semantic and distributional properties (Pantazara et al. 2008; Fotopoulou et al. 2009). In this context, support verbs (*Vsup*) and other verbs expressing diverse modalities (aspect, intensity, control, etc.) were identified and encoded as properties; these properties reveal the restrictions nouns impose on the lexical choice of verbs. Similarly, 339 Greek verbal predicates denoting emotion (*Vsent*) were classified into homogenous syntactico-semantic classes based on their syntactic, lexical and semantic properties (Giouli & Fotopoulou 2012); a number of syntactic features (i.e., argument structure, alternations), selectional restrictions imposed on the verbs' subject and object complements, emotion type, polarity and intensity were also defined and encoded formally.

In this respect, this work is further aimed at enriching the set of lexical resources pertaining to the semantic field of emotions with a lexicon that comprises verbal MWEs denoting emotion or emotional state. Moreover, the Greek MWEs were mapped onto their French counterparts. The ultimate goal was not only to develop a bi-lingual lexical resource, but also to test the hypothesis that, despite the idiosyncrasies that are inherent to MWEs in general, a certain degree of regularity (in terms of inherent properties) can be observed within a semantic class. To this end, we opted for reusing and extending existing lexical resources that encompass verb MWEs in Greek and French.

3 Methodological framework

The resources that form the basis of the present study have been developed using the Lexicon-Grammar (LG) methodological framework (Gross 1975). Being a model of syntax limited to the elementary sentences of a natural language, the theory argues that the unit of meaning is not located at the level of the word, but at the level of sentence of the form *Subject – Verb – Object*. Therefore, the elementary sentence is transformed to its predicate-argument structure, and the main complements (subject, one or more objects) are separated from other complements (adjuncts) on the basis of formal criteria. Distributional properties associated with words, i.e., types of prepositions, semantic features inherent to nouns in subject and object positions, etc. are also taken into account, resulting in a more fine-grained classification and in the creation of homogeneous word classes. Finally, transformation rules, construed as equivalence relations between sentences, generate additional equivalent structures. All this information (argument structure, distributional properties and permitted transformational rules) is formally encoded in the so-called *LG tables*.

Each table is defined by a set of distinct properties (syntactic, distributional, and semantic) and includes all the lexical items sharing these properties. Predicates with more than one usage or meaning are treated as separate lexical items possibly represented in different tables, and the syntactic and semantic properties are assigned to each entry as appropriate. In this sense, entries in one table are considered to form a homogeneous class. In an LG table, the set of properties that describe the entries are encoded as headers of the columns, whereas entries are listed at separate rows. At the intersection of a row corresponding to a lexical item (entry) and a column corresponding to a property, the cell is set to '+' if the property is valid for the given entry or '−' if it is not.

Similarly, MWEs are also treated as *elementary sentences* for which all possible fixed and non-fixed (or variable) arguments (if any) are consistently and uniformly encoded. The formalism provides the mechanism for encoding properties that are appropriate for the identification and processing of MWEs. More precisely, the MWE structure is represented as a Part-of-Speech sequence. According to the LG notation, *N* denotes a non-fixed nominal, whereas, *C* signifies a fixed one; numbers are used to represent the syntactic function of fixed or non-fixed constituents. In this sense, *N0* is used to represent a non-fixed noun in subject position whereas, *C0* denotes a fixed subject. Similarly, *N1*, *N2*, *N3*, etc., along with *C1*, *C2*, *C3* etc. denote complements in object position (or complements of prepositional phrases), marked also for fixedness. It should be noted, however,

that the internal structure of the noun phrase is not represented explicitly in general; patterns depict the elementary sentence or structure characterising each MWE class, whereas information regarding modifiers, determiners, etc. allowed for by certain expressions is provided in the form of *features* or *properties*. *Selectional restrictions* over the non-fixed or variable elements of MWEs as well as syntactic phenomena (e.g., passive alternation, etc.) – if any – are also encoded formally. Finally, other grammatical phenomena such as agreement features are accounted for.

For example, the MWE in (1) below comprises two fixed (or lexicalised) elements, a verb and a noun in subject position, and two variable elements, namely a nominal phrase in accusative and a possessive pronoun (*Poss*) that modifies the fixed nominal constituent. The variant nominal phrase is most often realised as a weak personal pronoun in pre-verbal position (*Ppv*); agreement in number and person between the two variable elements is mandatory:

(1) my devils catch me 'to become very angry'
    με πιάνουν τα διαόλια   *μου / *σου / *του Γιάννη*
    me pianun ta δiaolia   mu / *su / *tu Γiani
    me catch.3PL the devils.NOM.PL.POSS my / your / the John.GEN
    'to become very angry'

In this case, a generic syntactic pattern like the one depicted in (2) below is used to describe a class in a LG table.

(2)   a. Ppv V C0 Poss
        b. Ppv-1 V C0 Poss-1

The agreement attested between variable elements is then depicted via co-indexing as shown in (2b).

An example of MWE representation within the LG framework is illustrated in Table 1; the table comprises verbal MWEs with the underlying structure *N0 V Prep C1* (Fotopoulou 1993b).

It becomes evident, therefore, that the LG framework together with the requirement of substantial coverage leads to a *uniform* and *consistent* description of elementary sentences and the formal encoding of properties across languages in a comparable manner. In this respect, one of the main advantages of LG is that it allows comparisons between languages and facilitates the construction of *cross-language* resources.

Table 1: LG table of verbal MWEs (sample).

N0 =: +Hum	N0 =: -Hum			⟨ E ⟩			N1 =: Npc	N0V	N0 Vcmt N1 Prep C2	PhraseVsup
-	+	ακτινοβολώ	από	E	ευτυχία	E	-	+	-	-
+	-	αφρίζω	από	(E+τη)	λύσσα	(E+Poss-0)	-	-	-	-
+	-	βράζω	σε	το	ζουμί	Poss-0	-	-	-	-
+	-	γελάω	με	την	καρδιά	Poss-0	+	+	-	-
+	-	έρχομαι	σε	τα	λογικά	Poss-0	-	-	+	+
+	-	έρχομαι	σε	τα	συγκαλά	Poss-0	-	-	+	-
+	-	κάθομαι	σε	τα	αβγά	Poss-0	-	-	-	-
+	-	κάθομαι	σε	τα	αγκάθια	E	-	-	-	-
+	-	κιτρινίζω	από	τον	φόβο	Poss-0	-	+	-	-
+	-	λύνομαι	σε	τα	γέλια	E	-	-	-	-

4 Data selection

The initial list of Greek and French MWEs that pertain to the semantic field of emotions was manually compiled from data listed in existing LG tables for Greek (Fotopoulou 1993b; Mini 2009) and French (Gross 1982). The selection of the Greek MWEs was performed as a two-stage procedure: (a) manual identification of candidate MWEs that pertain to the semantic field emotion, and (b) validation of these candidate MWEs for inclusion or deletion on the basis of formal criteria besides intuitive judgments. The initial list of MWEs was further updated and extended, drawing on corpus evidence. More precisely, Greek MWEs were selected manually from a suite of specialised corpora (Giouli & Fotopoulou 2014) that were developed and annotated in view of guiding sentiment analysis. In this sense, our work is corpus-based and thus empirical rather than purely intuitive.

Since the scope of the current work is limited to clear instances of emotion denoting predicates (i.e., verbal MWEs), a formal distinction between direct and indirect affective expressions that correspond to emotion concepts was in order. For this reason, a set of lexical semantic tests (lexical substitution, paraphrasing, etc.) was adopted as a formal device guiding the selection of Greek verbal

emotion predicates. Therefore, a candidate MWE is selected for inclusion in the lexicon if at least one of the following criteria is met:

Criterion 1: A candidate Emotion MWE is selected if it can be replaced by a sequence that comprises one of the verbs *feel* or *cause* and a *noun that denotes emotion* (*Nsent*), that is, if there exists an *Nsent* that is related with the concept EMOTION via the IS-A relation, and the relation *MWE is semantically equivalent to "feel/cause Nsent"* is true. For example, the expression in (3) is semantically equivalent to an expression of the form *to feel* EMOTION, where EMOTION is *panic*:

(3) *με   πιάνει   πανικός*
     me piani   panikos
     me catches panic.NOM
     'to panic'

Criterion 2: A candidate Emotion MWE is selected if it can be replaced by a verb predicate that denotes emotion (*Vsent*), that is, if there exists a *Vsent* defined as a conceptualization of a FEEL-EMOTION or CAUSE-FEEL-EMOTION event and the relation *MWE is semantically equivalent to "Vsent"* is *true*. For example, the expression in (4) is semantically equivalent with the *Vsent φοβάμαι* (fovame) 'to be frightened':

(4) *πάγωσε το αίμα   μου*
     payose to ema   mu
     froze   the blood.NOM my
     'I was terrified'

Criterion 3: A candidate Emotion MWE is selected if it can be replaced by the verb *to be* and an adjective that denotes emotion (*Asent*), that is, if there exists an *Asent* defined as conceptualizing an EXPERIENCER-EMOTION or TRIGGER-EMOTION entity, and the relation *MWE is semantically equivalent to "to be Asent"* is true. In the example (5) below, the expression is semantically equivalent to an expression of the form *to be Asent* – *είμαι έκπληκτος* (ime ekpliktos) 'to be surprised':

(5) *μένω με   το στόμα ανοικτό*
     meno me   to stoma anikto
     stay   with the mouth open
     'to be aghast'

Finally, the selection of French MWEs denoting emotion and their mapping onto their Greek counterparts was performed manually. First, translations or

translational equivalents of the Greek MWEs were either provided by human translators or extracted from standard mono- and bilingual lexicographic resources, such as the *Trésor de la Langue Française Informatisé*[1] and *WordReference.com*. In certain cases, translations were obtained using English as a pivot language. These translations were checked against entries in existing LG tables that define the typologies of French MWEs (Gross 1982). Once an expression was spotted, it was selected and aligned to its Greek counterpart(s).

The afore-mentioned process resulted in the identification of 607 Greek and 520 French MWEs that constitute the linguistic data of the current study. As one might expect, the numbers show that there is no 1:1 correspondence between Greek and French MWEs denoting emotion. In fact, the process of translating the list of Greek MWEs to the target language proved that the transition from one language to the other was not always straightforward. The outcome of this procedure can be summed as follows (see also §6.2):

- a Greek MWE is mapped onto a French MWE;

- more than one Greek MWEs are mapped onto a single French MWE;

- a single Greek MWE corresponds to more than one French MWEs;

- one or more Greek MWEs correspond to a single-word French verb rather than an MWE.

5 Description of the MWEs Emotion Lexicon

Data encoding was performed after data selection. The challenge of representing MWEs in lexical resources is to ensure that the variability along with extra features required by the different types of MWEs can be captured efficiently (Calzolari et al. 2002; Copestake et al. 2002). To this end, features and properties that are appropriate for the robust computational treatment of MWEs were retained from existing LG tables where applicable. MWEs extracted from corpora were encoded from scratch. Syntactic information includes the argument structure of the elementary sentence (by also depicting fixed and variable elements), modification information (if permitted), syntactic alternations, and selectional restrictions imposed over the variable elements of the MWE (often in subject and object(s) position). Additionally, all MWEs were coupled with information about

---

[1]The resources are available online (http://atilf.atilf.fr/tlf.htm; http://www.wordreference.com/).

their type in terms of compositionality, syntactic rigidity idiosyncrasies, and lexical choice. Moreover, semantic features that are relevant to the semantic field to which each of these predicates adheres are also encoded, namely: emotion type, polarity and intensity. In this way, the typologies of emotion MWEs in Greek and French were consolidated and cross-lingual analogies or discrepancies were identified. In the remainder, we will elaborate further on the encoding of verb MWEs. As we have already mentioned above, linguistic information is encoded formally in both the Greek and French tables, and this common representation facilitates the extraction of shared patterns – if any.

5.1 Emotion MWEs: fixed expressions – SVCs

In this section, we present the classification of verbal MWEs included in the emotion lexicon. Entries were assigned a value corresponding to the type they belong to, namely (a) fixed (or idiomatic) expressions and (b) support (or light) verb constructions (SVCs).

The identification of fixed expressions involves lexical, morphosyntactic and semantic criteria (Gross 1982; 1998b; Lamiroy 2003), to be taken into account, namely: *non-compositionality*,[2] i.e., the meaning of the expression cannot be computed from the meanings of its constituents*; non-substitutability*, i.e., at least one of the expression constituents does not enter in alternations at the paradigmatic axis; and *non-modifiability*, in that they enter in syntactically rigid structures, posing further constraints over modification, transformations, etc. To this end, linguistic tests were applied to all MWEs. The examples that follow conform to the criteria mentioned and are classified as *fixed expressions*:

(6) *δαγκώνω τη λαμαρίνα*
δagono   ti   lamarina
bite   the panel.ACC

'to be in love'

(7) *serrer   les dents*
to.clench the teeth

'to grit one's teeth/to be stressed or angry'

---

[2]We distinguish between *composability/decomposability* (Nunberg et al. 1994: 496) and *compositionality/non-compositionality*. Composability concerns the property of phrase elements to "[c]arry identifiable parts of the idiomatic meaning".

On the other hand, identification of SVCs for inclusion in the emotion lexicon is based on the following criteria:

SVCs Criterion 1: SVCs comprise a *support verb* (*Vsup*) and a *predicative noun denoting emotion (Nsent)*; support or light verbs of this type bear no meaning and are simply carriers of tense and person;

SVCs Criterion 2: SVCs comprise specific (modal) verbs expressing *diverse modalities* (aspect, intensity, control, etc.) and an *Nsent*. These verbs are considered as *Vsup variants.*

In this respect, SVCs are – to some extent – characterised by semantic transparency due to the fact that the predicative noun, which carries the predicative function within the SVC, is used in one of its literal senses. Basic support verbs are *έχω* (eχo)/*avoir* 'to have', *είμαι Prep* (ime Prep)/*être Prep* 'be Prep', *κάνω* (kano)/*faire* 'to make', the operator verb *δίνω* (δino)/*donner* 'to give', and the causative verbs *προκαλώ* (prokalo)/*défier, provoquer* 'to cause', *προξενώ* (prokseno)/ *provoquer* 'to cause', *αφήνω* (afino)/*laisser* 'to leave', which have an effect on structures with the basic *Vsup*. In practice, however, SVCs are highly *idiosyncratic* and for this reason, it is quite difficult to predict which *Vsup* combines with a noun (Abeille 1988). In the case of emotion MWEs, a close inspection of the data, showed that domain-specific verbs assume the function of a basic *Vsup*. Greek SVCs in this semantic field usually select for the verbs *νιώθω* (nioθo) 'to feel' or *αισθάνομαι* (esθanome) 'to feel' (see (8)); similarly, their French counterparts select for the verbs *éprouver* 'to feel' and *ressentir* 'to feel', as shown in the example (9) below. These constructions are semantically equivalent with single-word verb predicates denoting emotion.

(8) *νιώθω χαρά*
     nioθo χara
     feel   joy.ACC
     'to feel joy'

(9) *ressentir de la   joie*
     to.feel   of the joy
     'to feel joy'

Additionally, certain verbs selected by the *Nsent* predicates that function as *Vsup* variants may further denote the degree or intensity of the emotion. From a cross-linguistic perspective, these *Vsup* variants usually form a pair of translational equivalents in Greek and French as shown in the examples (10) and (11) respectively:

(10) *πετάω από χαρά / τη   χαρά μου*
     petao apo χara / ti   χara mu
     fly   from joy   / the joy   my
     'to be very happy'

(11) *sauter   de joie*
     to.jump of joy
     'to be very happy'

Classification of MWEs as fixed expressions or SVCs is not always straightforward or clear-cut, as shown in §5.2.2 and §6.1. In fact, some expressions seem to comprise an intermediate class placed in between fixed expressions and SVCs. In other words, there seems to be a continuum between fixed expressions and SVCs (or between fixed and free expressions in other cases). These expressions may be considered (under syntactic and semantic conditions) as *semi-fixed*. A study of these expressions related to the *degree of fixedness* is currently in progress (Constant & Fotopoulou 2016).

5.2 Syntactic properties

Syntactic (and semantic) information is extracted from the LG tables for those MWEs that were accounted for in the past; new MWEs selected for the purposes of the current study were encoded as appropriate. Syntactic information in the LG tables comprises the *argument structure of each MWE*, *the syntactic alternations* defined for the particular MWE, and *selectional restrictions* imposed over the variable elements of the expressions. The encoding of modifiability specifically concerns the fixed modifiers of SVCs. In the next sections, we elaborate on these aspects.

5.2.1 Argument Structure

Verbal MWE expressions (fixed non-compositional and SVCs) that denote an emotion bear no syntactic idiomaticity, since they generally conform to the argument structure of the main verb and there is nothing exceptional in their syntactic behavior. This information is only implicitly encoded in the LG tables. In this respect, naming conventions of the initial tables correspond to specific configurations cross-linguistically, and this information can be easily and effectively retained in the current lexical resource. Information with respect to the underlying structure and the syntactic function of the (fixed and variable) constituent(s)

further shows that verbal MWE predicates conform to the following patterns: (i) fixed subject MWEs, (ii) fixed complement MWEs, and (iii) any combination of the above. These types are presented in detail in the following paragraphs.

Fixed Subject MWEs comprise a verb and an *NP* in subject position; these are both lexicalised. Complements (if any) are represented as variant elements. According the LG notation, the generic syntactic pattern that describes MWEs of this type is *C0 V Ω*. The symbol *Ω*[3] is used to denote one or more complements a predicate subcategorises for, without further specifying their form. In the LG tables, however, the form and function of variable elements are further encoded. For example, the patterns *C0 V N1* and *C0 V Prep C1 N2gen*, used to describe Greek and French expressions in (12) and (13) below, further license a variable nominal phrase in object position or as the complement of a *PP* modifier respectively:

(12)   cold sweat bathes me 'I am terrified'
       Κρύος ιδρώτας   έλουσε την Άννα.
       krios iδrotas   eluse tin Ana
       cold   sweat.NOM.SBJ bathed the Anna.ACC.OBJ
       'Anna was terrified.'

(13)   *La haine   niche dans le   coeur de Anna.*
       the hate.SBJ nests in   the heart of Anna
       'Anna hates.'

It should be noted, however, that the variable complement is usually employed in its cliticised form as shown in (14); this property is also encoded in the LG tables.

(14)   cold sweat baths me 'I am terrified'
       Την   έλουσε κρύος ιδρώτας.
       tin   eluse krios iδrotas
       her.OBJ bathed cold   sweat.SBJ
       'She was terrified.'

Similarly, Greek SVCs may comprise an aspectual variant of a *Vsup* and a predicative noun denoting an emotion in subject position:

---

[3] We will not discuss the possible forms assumed by *Ω* in detail.

(15) *με πιάνει πανικός*
    me piani   panikos
    me catches panic.NOM.SBJ
    'to panic'

Fixed Complement MWEs. Verbal MWEs of this type comprise a verb and one lexicalised complement. Most often, this lexicalised complement is an *NP* in direct object position. The subject is represented as a variable argument of the elementary sentence; the generic syntactic pattern that describes fixed verbal MWEs of this type is *N0 V C1,* whereas the syntactic pattern of SVCs is *N0 Vsup Nsent*:

(16) *δαγκώνω τη λαμαρίνα*
    δagono ti lamarina
    bite   the panel.OBJ
    'to be in love'

(17) *avoir   du chagrin*
    to.have of grief
    'to be sad'

Fixed PP Complement MWEs comprise a verb and a lexicalised prepositional phrase (*PP*) complement. The variable *NP* in subject position along with other non-fixed elements (if any) is also represented as appropriate. The generic pattern that describes this class is of the form *N0 V Prep C1*. In (18), the Greek MWE consists of the verb *κάθομαι* (kaθome) 'to sit' and the lexicalised *PP στα καρφιά* (sta karfia) 'on the nails'. Similarly, the French MWE in (19) consists of the verb *rire* 'to laugh' and the PP *aux larmes* 'to tears':

(18) *κάθομαι στα   καρφιά*
    kaθome sta   karfia
    sit   to.the nails
    'to be anxious', 'to be on tenterhooks'

(19) *rire   aux larmes*
    to.laugh to.the tears
    'to roar with laughter'

Fixed Adjunct MWEs comprise a verb plus an adjunct (often adverb) that are both lexicalised. Other variable complements are depicted in the structure of the relative elementary sentence:

(20) *φέρω βαρέως*
      fero   vareos
      carry heavily
      'to be very sad'

(21) *Ils   s'   aiment comme deux tourtereaux*
      they REFL love   like   two lovebirds
      'They are in love.'

Finally, a number of verbal MWEs have a *syntactic* structure that is a combination of the configurations presented. These structures are exhaustively represented in the resource:

(22) *μου   ανεβαίνει το αίμα   στο   κεφάλι*
      mu   aneveni to ema   sto   kiefali
      me.GEN raises   the blood.NOM to.the head
      'to become very angry'

(23) *la   moutarde monte au   nez*
      the mustard   raises to.the nose
      'to become very angry'

(24) *avoir   froid dans le   dos*
      to.have cold in   the back
      'to be terrified'

5.2.2 Modification

Fixed non-compositional verbal expressions do not allow for any modification over the fixed constituents. On the contrary, SVCs are considered as syntactically more flexible constructions, and adjectival modification is allowed over the *Nsent*. However, constructions with a *Vsup* do not conform to a uniform pattern of modification (Moustaki et al. 2008). Adjectival modification within the MWE is found to be *free*, *semi-fixed* or even *fixed*. Modification in both languages involves *intensifiers* or – more generally – *grade indicators* like *μεγάλος* (meγalos)/*grand* 'big', *λίγος* (liγos)/*petit* 'few', *φοβερός* (foveros)/*intense* 'awful', *άκρατος* (akratos)/*intense* 'awful', etc.:

(25) *O   Γιάννης νιώθει ένα παθολογικό   / υπαρξιακό / αόριστο / *δυνατό άγχος.*
    o   Γιανις nioθi ena παθολογiko / iparksiako / aoristo / δinato anχos
    the John   feels a   pathological / existential / vague / strong anxiety

    'John feels a pathological / existential / vague / *strong anxiety.'

(26) *Jean éprouve une angoisse pathologique / vague / sourde / mortelle / de*
    John feels   an anxiety pathological / vague / silent / deadly / of
    mort / existentielle.
    death / existential
    'John feels a pathological / vague / silent / deadly / existential anxiety.'

(27) *Με έπιασε μαύρη απελπισία   / *λύπη.*
    me epiase mavri apelpisia   / lipi
    me cought black dispair.NOM / sorrow.NOM
    'I was in total despair.'

(28) *J' ai   eu une peur bleue / *tristesse bleue.*
    I have had a   fear blue / sadness blue
    'I was terrified.'

The *fixed* modifiers, i.e., modifiers that seem to be idiosyncratic to a given *Nsent* cannot be employed productively. We note that in example (27) , the adjective μαύρη, mavri, 'black' is only used as a modifier of the nominal predicate απελπισία, apelpisia, 'despair', which cannot be described literally as being of black colour. Similarly, the French adjective *bleu* 'blue' in (28) is only used with the nominal predicates *peur* 'fear'. These expressions are also encoded as fixed in the LG tables. Actually, this is evidence of the existence of grey zones between SVCs and fixed expressions (cf. §5.1).

To conclude, Greek and French *Nsent* predicates in a SVC select from a variety of modifiers in an idiosyncratic manner. Moreover, the respective Greek and French expressions seem to present a variable degree of fixedness depending on the *Nsent* and the modifier selected. Free and semi-fixed modifiers are not encoded in the lexicon so far. On the contrary, fixed modifiers of the predicative noun are encoded as fixed elements of the expression.

5.2.3 Syntactic alternations

Information relative to syntactic alternations encoded in the LG tables was also kept in the lexical resource. The causative-inchoative alternation is a syntactic

property that involves verbs (or pairs of verbs) which have an intransitive and a transitive usage. The inchoative form (intransitive) denotes a *change of state*, and the causative form (transitive) denotes a *bringing about of a change of state*. A number of emotive MWEs were found to enter this alternation. The following cases have been attested in the LG tables:

First case: a pair of two MWEs each one comprising a distinct verb, whereas all the other fixed elements are identical. The two verbs (which are often predicates denoting movement) normally enter (or signal) the transitive-intransitive alternation:

(29)   to take one out of one's clothes 'to make someone angry'
    *o Γιάννης την **βγάζει** τη Μαρία από τα ρούχα της.* (CAUS)
    o Γianis tin vγazi ti Maria apo ta ruχa tis
    the John.sʙᴊ her.oʙᴊ takes.out the Maria.oʙᴊ from the clothes hers

    'John makes Maria very angry.'

(30)   to get out of one's clothes 'to be made angry'
    *η Μαρία **βγήκε** από τα ρούχα της.* (INCHO)
    i Maria vγikie apo ta ruχa tis
    the Maria.sʙᴊ went.out from the clothes hers

    'Maria was made very angry.'

(31)   to send someone to the seventh sky 'to make someone happy'
    *Eric **envoie** Léa au septième ciel.* (CAUS)
    Eric sends Lea to.the seventh sky

    'Eric makes Lea very happy.'

(32)   to go up to the seventh sky 'to be happy'
    *Léa **monte** au septième ciel.* (INCHO)
    Lea goes-up to.the seventh sky

    'Lea is in the seventh heaven.'

Second case: MWEs that comprise a verb that enters the transitive-intransitive alternation (ergativity):

(33)   to turn someone's lights on 'to make someone angry'
    *o Γιάννης μου **άναψε** τα λαμπάκια.* (CAUS)
    o Γianis mu anapse ta labakia
    the John.sʙᴊ I.ɢᴇɴ turned.on the lights.oʙᴊ

    'John made me very angry.'

(34) my lights turn on 'I get angry'
      *Μου **άναψαν** τα λαμπάκια*   (INCHO)
      mu   anapsan   ta   labakia
      I.ɢᴇɴ turned.on the lights.ѕʙᴊ
      'I got very angry.'

Similarly, other syntactic properties were encoded in the LG tables where applicable (i.e., passivisation, genitive-dative alternation, etc.).

5.2.4 Selectional restrictions

A number of *selectional restrictions* that are imposed on the *variable* elements of the MWEs (in subject and object(s) position) were encoded as properties in the LG tables. Like their single word counterparts, verbal MWEs denoting emotion select a nominal element that is obligatorily [*+human*]. Being at the heart of the syntax-semantics interface, this information relates to the participants of the emotion event. An emotion event generally involves an Exᴘᴇʀɪᴇɴᴄᴇʀ (that is, the individual experiencing the psychological state) and a Tʜᴇᴍᴇ (that is, the content or object of the psychological state) or – occasionally – a Cᴀᴜѕᴇ. These participants, however, are not realised in a uniform way in single word verbal predicates. In this respect, the distinction between *SubjectExperiencer* (SubjExp) and *ObjectExperiencer* (ObjExp) single word verbal predicates has been established (Belletti & Rizzi 1988) based on the syntactic distribution of the verbal arguments and the associated Semantic Roles. The former project the Exᴘᴇʀɪᴇɴᴄᴇʀ of the emotion as their structural subject and the Tʜᴇᴍᴇ or the Sᴛɪᴍᴜʟᴜѕ as their structural object; the latter realise the Tʜᴇᴍᴇ or the Sᴛɪᴍᴜʟᴜѕ as the subject and the Exᴘᴇʀɪᴇɴᴄᴇʀ as their object. This information is of relevance to a number of NLP applications, and although it has not been encoded in the LG tables, it can be deduced easily. In fact, as it has been shown (Giouli & Fotopoulou 2014) for the single-word verbal predicates denoting emotion, the *N0* or *N1* complements with the [*+human*] restriction can be mapped onto the Exᴘᴇʀɪᴇɴᴄᴇʀ participant in the emotion event.

This is true for MWEs too; here the Exᴘᴇʀɪᴇɴᴄᴇʀ is realised not as a structural subject but in object position. In this sense, the non-fixed element that bears the semantic restriction [*+human*] corresponds unambiguously to the Exᴘᴇʀɪᴇɴᴄᴇʀ of the emotion. In the following examples, the Exᴘᴇʀɪᴇɴᴄᴇʀ of the emotion is expressed by the subject of the Greek and French expressions as shown in (35) and (36) respectively, or by the direct object as depicted in (37) and (38) below:

(35) *H   Άννα   πετάει από χαρά.*
      i   Ana   petai apo χara
      the Anna.sbj.exp flies   of joy
      'Anna is very happy.'

(36) *Anna   rayonne de joie.*
      Anna.sbj.exp shines   of joy
      'Anna is very happy.'

(37) to take one out of one's clothes 'to make someone very angry'
      o   Γιάννης με   έβγαλε   από τα ρούχα μου.
      o   Γianis me   evγale   apo ta ruχa   mu
      the John.sbj me.obj.exp took.out of the clothes mine
      'John made me very angry.'

(38) *Ce film   m'   a   ému   aux larmes.*
      this film.sbj me.obj.exp has touched in   tears
      'This film moved me to tears.'

Additionally, other *selectional restrictions* imposed on the variable elements of the verbal MWE are encoded. These restrictions further specify the type of complements (nominal, prepositional, sentential) that these predicates sub-categorise for. In this respect, prepositions selected by the MWE predicates are formally depicted and encoded.

5.3 Semantic classification

The semantic classification of the studied Greek and French MWEs was aimed at grouping them under pre-defined emotional concepts and at distinguishing semantically between expressions that are near synonyms. This was attempted following a schema defined for single-word Greek verbs denoting emotion (Giouli & Fotopoulou 2012) along three dimensions: (a) *emotion type* (b) *emotion polarity* (c) *emotion intensity* and (d) *aspect* of the emotion event. The semantic classification of verbal MWE predicates was performed separately by two experienced linguists in the form of primarily intuitive semantic grouping. At the next stage, discrepancies between the annotations thus obtained were discussed and resolved, whereas cases for which no agreement could be consolidated were left aside for future treatment. The outcome of this procedure was the definition of specifications that would be applicable for distinguishing between semantic classes.

Emotion is described as a set of two or more dimensions. The most common ones are *polarity*, i.e., *positive or negative connotation* of emotion and the *intensity* or strength of the emotion. The notion of semantic polarity, or the semantic orientation of words (whether they denote a positive or a negative emotion) has also been the focus of attention in many studies aimed at sentiment analysis (Esuli & Sebastiani 2006; Wilson et al. 2005) *inter alia*. In our approach, the encoding schema provides for the annotation of the *a priori polarity* of the emotion denoted, which subsumes one of the following values: (a) *positive*, i.e., predicates which express a pleasant feeling (b) *negative*, i.e., predicates which express an unpleasant feeling (c) *neutral*, i.e., predicates that denote an emotion that is neither positive not negative and (d) *ambiguous*, i.e., predicates expressing a feeling, the polarity of which is context-dependent (e.g., surprise).

Polarity identification results in a coarse – yet quite effective – classification of emotion expressions; a more fine-grained one was attempted on the basis of emotion types. Psychological considerations of sentiment claim that some emotions are more basic than others, therefore, they should be universal to all human languages. The identification of basic emotions is based upon specific functional and physiological criteria, yet languages are claimed to possess inventories that comprise a great number of emotion predicates that cannot be easily accommodated within such fairly straightforward schemes. To this end, different dimensions of emotion can be used to delineate senses. In the work presented here we adopted an extended version of the typological model defined by Plutchik (2001). The initial model comprises eight basic emotions: *anger*, *fear*, *sadness*, *disgust*, *surprise*, *anticipation*, *acceptance* and *joy*. On the basis of corpus evidence derived from a tri-lingual corpus (English, Greek, Spanish) annotated for sentiment (Giouli et al. 2013), the initial list of basic emotions was further extended with a set of complex emotions, such as *love* and *hate* or emotions of (self-)appraisal (e.g., *shame*, *respect*) that were not considered by Plutchik. To better account for the conceptual representation of the emotion vocabulary, the final set of emotion types includes 15 new classes, namely: *admiration*, *boredom*, *disappointment*, *envy*, *gratitude*, *hate*, *indifference*, *jealousy*, *love*, *relaxedness*, *remorse*, *resentment*, *respect*, and *shame*. Greek and French MWEs were assigned an emotion concept; this classification results in grouping Greek and French verbal MWEs under emotion concepts.

Moreover, to model the semantic distinction between near synonyms that occur within a semantic class such as φοβάμαι, fovame, 'to be scared', πανικοβάλλομαι, panikovalome, 'to panic', μου κόπηκαν τα ήπατα, mu kopikan ta ipata, 'to be very frightened', etc., entries were further coupled with the feature *inten-*

sity (or *strength*). The following values are provided for by the schema for the feature *strength*: *low*, *medium*, *high*, and *uncertain*. In fact, emotion verbal predicates have been shown to possess scalar qualities (Fellbaum & Mathieu 2012). In this respect, groups of verbs that were assigned the same emotion type were checked in order to identify different degrees of intensity of the same underlying emotion. In this respect, intuitive judgments of trained lexicographers were systematised and a number of linguistic tests were defined aimed at the consistent annotation and the ordering of predicates according to the intensity of the emotion they denote.

In both languages, intensity was proved to be dependent on the following aspects: (a) degree of fixedness (b) modifier selected (in SVCs) and (c) the *Vsup* selected. More precisely, the majority of verbal idioms were judged to express an emotional state or event of high intensity; these were further marked as not accepting any modifier. Similarly, the *Vsup* of an SVC seemed to have an impact on the value assigned to the feature intensity. Ultimately, a number of *Vsup* function as an intensifier of the emotion denoted. In this respect, the verbs *έχω* (εχο)/*avoir* 'to have', *νιώθω* (nioθo)/*éprouver* 'feel' and *αισθάνομαι* (esθanome)/*ressentir* 'to feel' in Greek and French respectively usually denote an emotion that bears the value *medium* for the feature *intensity*; on the contrary, when the verbs *πετάω* (petao) 'fly' and *rayonner* 'shine' are employed instead, the entire expression is marked as denoting the same emotion, yet with an intensity marked as *high*. Modification of the Greek and French expressions is permitted only when the *Vsup* that evokes a *medium* intensity of an emotion is employed as shown in (39) and (41); when the *Vsup* denoting an emotional state of *high* intensity is employed, modification is blocked as in (40) and (42):

(39) *Η   Άννα   νιώθει χαρά / μεγάλη χαρά.*
      i   Ana   nioθi   χara / meγali χara
      the Anna feels   joy   / big   joy
      'Anna is happy / very happy.'

(40) *Η   Άννα   πετάει από χαρά / *μεγάλη χαρά.*
      i   Ana   petai   apo χara / *meγali χara
      the Anna flies   of   joy   / big   joy
      'Anna is very happy.'

(41) *Anna éprouve de la   joie / une grande joie.*
      Anna feels   of the joy / a   big   joy
      'Anna is happy / very happy.'

(42) *Anna rayonne de joie / *rayonne d' une grande joie.*
      Anna shines of joy / shines   of a   big   joy
      'Anna is very happy.'

Finally, the encoding schema also provides values for the feature *aspect*, i.e., the perspective taken on the internal temporal organization of the emotion event. Different values of *aspect* distinguish different ways of viewing the internal temporal constituency of the same event. The schema adopted provides the values *inchoativeAspect*, *terminativeAspect*, *durativeAspect* and *frequentiveAspect*. The encoding at this level, however, has been finalised only for the Greek MWEs.

6 Discussion

At the final stage of our study, an examination of the interplay between syntactic, semantic and lexical features of the studied MWEs was performed. Moreover, cross-lingual similarities and differences were identified. As has already been mentioned, our working hypothesis was that despite idiosyncrasies, MWEs that pertain to a given semantic class share features that are characteristic for this class; moreover, these features can be even attested cross-linguistically. As has already been mentioned in §5.1 above, MWE identification and classification employs lexical and morphosyntactic besides semantic criteria (Gross 1982; 1998a; Lamiroy 2003). However, they do not apply in all cases in a uniform way, and the variability attested brings about the notion *degree of fixedness* (Gross 1996). On the one hand, fixed expressions bear a meaning that cannot be computed based on the meaning of their constituents and the rules used to combine them. SVCs, on the other hand, have a rather transparent meaning due to the presence of the *Nsent* which retains its original sense. However, a number of problems are posed and the limits between SVCs and verbal fixed expressions (see also §5.1) are in some cases fuzzy: despite the semantic transparency entailed by the *Nsent*, the overall structure is often susceptible to a number of constraints as shown in examples (43) and (44) below:

(43) *Φωτίστηκε το πρόσωπο του Νίκου από χαρά.*
      fotistikie   to prosopo tu Niku apo χara
      was.lit.up the face.NOM the Nikos by   happinesss
      'Nikos' face lit up with happiness.'

(44)   * *Φωτίστηκε ο   Νίκος   από χαρά.*
       fotistikie   o   Nikos   apo χara
       was.lit.up   the Nikos.ɴᴏᴍ by   hapiness

According to a study on verbal MWEs (Balibar-Mrabti 1995), expressions like the one depicted in (43) are defined as *semi-fixed* ones. In this respect, the verbal MWEs under study were found to be placed along the continuum *fixed*, *semi-fixed* and SVCs. Consequently, the class of semi-fixed expressions constitutes a grey zone, the intermediate mentioned in §5.1 and §5.2.2. However, in this work, we opted for classifying semi-fixed expressions that comprise a predicative noun *Nsent* as SVCs.

One step further, the correlation between the features *non-compositionality/ fixedness* and the attributes *polarity* and *intensity* was examined. Our underlying assumption was that the degree of fixedness of the relevant expressions and the polarity/intensity of the emotion denoted are highly correlated. In this respect, the focus was placed on the values assigned for the feature *intensity* of the emotion denoted and their correlation to the aspects of MWE category (i.e., fixed expression or SVC). The majority of the considered Greek MWEs, that is 410 expressions, were attributed the value *Negative* for the feature *Polarity*, whereas only 169 were encoded as *Positive* and 133 as *Neutral*. Of these, 97 MWEs denote *anger*, 73 denote *fear*, and 105 denote *sadness*; 90 expressions were identified as expressing *joy* and 30 a *surprise* event. The remaining expressions are distributed across the remaining conceptual categories. Another interesting remark concerns verbal idiomatic *non-compositional* expressions; most of the expressions (260) that have been assigned the value *negative* for the feature *polarity* are also encoded as being of type *fixed* (as opposed to 150 expressions classified as SVCs). Additionally, fixed expressions were – in most cases – attributed a value *high* for the feature *intensity*. Of the approximately 300 *fixed* expressions, 210 are assigned the value *high* for the feature *intensity*. On the contrary, SVCs in both languages do not constitute a uniform class, and the overall emotion intensity denoted depends largely on the *Vsup* selected rather than the *Nsent* itself. Three cases are identified:

- The *Vsup* is selected by all *Nsent* predicates; these verbs[4] adhere to a productive and relatively open paradigmatic axis, and syntactic variability is allowed to some extent. In these cases, the *intensity* of the emotion denoted

---

[4]For example, *έχω* (eχo)/*avoir* 'to have', *νιώθω* (nioθo)/*éprouver* 'to feel' and *αισθάνομαι* (es-θanome)/*resentir* 'to feel'.

is determined on the basis of the semantics of the *Nsent*; any possible modifier functions as an intensifier of the emotion denoted.

- The *Vsup* selection is subject to lexical restrictions, and syntactic variability is not allowed.[5] In this case, the *Vsup* contributes to the intensity and/or some aspectual meaning of the emotion denoted. The overall intensity of the emotion expression is determined on the basis of the semantics of the *Nsent*, and the *Vsup* functions as an intensifier.

- The *Vsup* selection is extremely limited or unique, and a strong lexicalization is attested; syntactic variability is not allowed and the *Vsup* is an intensive or aspectual variant that has a strong impact on the intensity of the emotion denoted:

(45) *με τρώει η ζήλια   / *στενοχώρια / *λύπη*
       me troi   i   zilia   / stenoχoria / lipi
       me eats   the jealousy.NOM / worry.NOM   / regret.NOM

       'to be devoured by jealousy'

(46) *être rongé   par la jalousie*
       to.be gnawed by   the jealousy

       'to be devoured by jealousy'

7 Cross-lingual considerations

Research on idioms reported in Villavicencio et al. (2004) shows that there is remarkable variation in MWEs across languages. Similar variations are attested in the data used in the current research. As one might expect, there is no one-to-one correspondence between syntactic patterns in the two languages. It is worth looking at SVCs and fixed expressions separately here.

Greek and French SVCs present a number of similarities in terms of the underlying syntax and semantics. In some cases, even a direct lexico-syntactic correspondence is observed for a cross-lingual MWE pair with similar semantics as illustrated in (47) and (48) below. Furthermore, semantic transparency in SVCs implies more correspondences at least at the level of syntactical patterns – we have demonstrated this with examples (8) and (9). As one might expect, differences between the Greek and French expressions are limited to basically those

---

[5]For example, *ανατριχιάζω* (anatriχiazo)/*frissoner* 'to shiver', *λάμπω* (labo)/*briller* 'to shine', *λιώνω* (liono)/*fondre* 'to dissolve', etc.

that exist in general between the two languages, i.e., usage of determiners and the indefinite article, case marking for NPs in subject and object position in Greek as opposed to PP complements in French, etc.

(47) to give to the nerves 'to cause anger'
      δίνω στα   νεύρα
      δino sta   nevra
      give to.the nerves
      'to cause anger'

(48) donner sur les nerfs
      to.give on the nerves
      'to cause anger'

In other cases, Greek and French SVCs share the same syntactic structure and underlying semantics, yet their lexical composition is different. The differences are attested both in the lexical choice of the *Vsup* and/or the overall structure of the verbal expression. For example, the French verb *nager* 'to swim' seems to be more productive than its Greek counterpart *πλέω* (pleo) 'to sail' as shown in (49) and (50) below. The latter is only employed in a rather fixed configuration and selects only one *Nsent*, showing, thus, a limited (or even fixed) distribution:

(49) *nager   dans le bonheur/ la joie/ l' optimisme/ l' amour*
      to.swim in   the happiness/ the joy/ the optimism/ the love
      'to be very happy/ happy/ very optimistic/ in love'

(50) *πλέω σε πελάγη ευτυχίας/   *στην ευτυχία/   *στην αισιοδοξία/ *στην*
      pleo se pelaγi eftiçias/   stin eftiçia/   stin esioδoksia/ stin
      sail in seas   happiness.GEN/ in.the happiness/ in.the optimism/ in.the
      αγάπη
      aγapi
      love
      'to be very happy/ happy/ optimistic/ full of love'

Being *conceptual metaphors* (usually obsolete), *fixed expressions* present in some cases considerable similarities in both lexical choice and structure cross-linguistically. Again, differences are limited to the usage of determiners, argument realization, selection of prepositions, etc. Often, the lexicalised nominal element (that assumes the function of the direct object) denotes a part of the

body (*Npc*) as exemplified below. These expressions open a slot that is filled by a variable noun in genitive case in Greek and a PP complement in French (*à N* 'to N'). This element is usually realised as a cliticised pronoun – in both Greek (51) and French (52)) – and it designates the beneficiary of the event expressed by the predicate (Leclère 1976; Fotopoulou 1993a). This genitive (in Greek) and PP (in French) is a specific case with semantic and syntactic features; Leclère (1976) has offered the term *datif étendu*) for this genitive:

(51) μου κόβονται τα ήπατα 'my liver is cut'
  του   κόπηκαν τα   ήπατα
  tu   kopikan ta   ipata
  he.GEN cut   the liver.PL.NOM

  'to be frightened'

(52) *lui casser   les pieds*
  him to.break the feet

  'to get on one's nerves'

In some cases, similarities are attested in terms of argument structure. For example, the Greek verbal expression depicted in (53) and its French counterpart shown in (54) are encoded as entries in Greek and French tables. Each table features MWEs that share the same properties and lexico-syntactic constraints; this means that the resulting tables are to a large extent homogenous. Therefore, correspondences between homogenous LG tables in Greek and French can be obtained and mappings of MWEs from one language to the other are feasible.

(53) *βγαίνω από τα ρούχα μου*
  vɣieno apo ta ruχa   mu
  get.out from the clothes mine

  'to be very angry'

(54) *sortir   de ses gonds*
  to.get.out of one's pumps

  'to be very angry'

Additionally, there are many verbal idiomatic expressions which have no direct or precise equivalent in the other language and they correspond to a single word verbal predicate, as shown in the Greek example (55) which is attributed the French verb *gâcher* 'to spoil':

(55) to me he/she/it takes it out sour 'he/she/it makes it unpleasant to me'

    του   *το βγάζω*   *ξινό*

    tu   to vɣazo   ksino

    he.GEN it take.out sour

    'to make unpleasant'

Semantically almost equivalent expressions that still present differences in aspectual meaning and/or the intensity of the emotion have been identified in the Greek and (to a large extent) in the French data. Sense discrimination and the alignment of Greek and French MWEs can be enhanced on the basis of the values assigned to those emotion-related attributes: a set of MWEs are classified under the same emotion concept, yet sense discrimination is further enhanced on the basis of the values assigned to emotion-related attributes.

8 Conclusions and future research

MWEs pose challenges with respect to their identification, analysis and representation both to linguistic theory and to applications. In this study, we aimed at consolidating the typologies of emotion MWEs in Greek and French and at finding cross-lingual analogies and asymmetries. The syntactic, lexical and semantic properties of the Greek and French verbal constructions were systematically examined, by taking also into account the semantic properties of the semantic field, namely the features intensity and polarity of the emotion denoted. We have shown that, despite existing idiosyncrasies, in both languages the MWEs in the semantic field of emotion share properties. Moreover, syntactic, semantic and lexical features of emotion MWEs seem to have an impact on the semantics of the expression in terms of emotion-related features. Future work will be oriented towards (a) investigating the properties of semi-fixed expressions, taking into account the degree of fixedness (b) studying the aspectual variants of SVCs in both languages (c) revising the coding used in the emotion Lexicon according to new studies and data and (d) populating the lexical resource with new expressions.

Acknowledgements

Abbreviations

Asent	ajdective denoting emotion	Nsent	noun denoting emotion
caus	cause	Ppv	pre-verbal position
exp	experiencer	refl	reflexive pronoun
incho	inchoative	svc	support verb construction
e	zero element	Vsent	verb denoting emotion
lg	Lexicon-Grammar	Vsup	support verb

References

Balibar-Mrabti, Antoinette. 1995. Une étude de la combinatoire des noms de sentiment dans une grammaire locale. *Langue Française* 105. 88–97.

Belletti, Adriana & Luizi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6. 291–352.

Calzolari, Nicoletta, Charles J. Fillmore, Ralph Grishman, Nancy Ide, Alessandro Lenci, Catherine MacLeod & Antonio Zampolli. 2002. Towards best practice for multiword expressions in computational lexicons. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*. Las Palmas, Canary Islands.

Constant, Matthieu & Aggeliki Fotopoulou. 2016. *A systematic study on the fixedness degree of verbal multiword expressions: Application to Modern Greek and French.* Tech. rep. PARSEME's 6th general meeting in Struga, FYR Macedonia. Selected posters. 7-8 April 2016.

Copestake, Ann, Fabre Lambeau, Aline Villavicencio, Francis Bond, Timothy Baldwin, Ivan Andrew Sag & Dan Flickinger. 2002. Multiword expressions: Linguistic precision and reusability. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*. Las Palmas, Canary Islands: Association for Computational Linguistics.

Esuli, Andrea & Fabrizio Sebastiani. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*.

Fellbaum, Christiane & Yannick Yvette Mathieu. 2012. Scalar properties of emotion verbs and their representation in WordNet. In *Proceedings of the 6th International Conference of the Global WordNet Association (GWC-2012)*, 100–105. Matsue, Japan.

Fotopoulou, Aggeliki. 1993a. Traitement du cas génitif dans une classification des phrases à compléments figés du grec moderne. *Lingvisticae Investigationes* XVII(2). 259–280.

Fotopoulou, Aggeliki. 1993b. *Une classification des phrases à compléments figés en grec moderne. Étude morphosyntaxique des phrases figées.* Saint-Denis: Université Paris VIII dissertation.

Fotopoulou, Aggeliki, Marianna Mini, Mavina Pantazara & Argyro Moustaki. 2009. La combinatoire lexicale des noms de sentiments en grec moderne. In Iva Novacova & Agnes Tutin (eds.), *Le lexique des émotions*, 81–103. Grenoble: ELLUG.

Giouli, Voula & Aggeliki Fotopoulou. 2012. Emotion verbs in Greek. From Lexicon-Grammar tables to multi-purpose syntactic and semantic lexica. In *Proceedings of the XV Euralex International Congress (EURALEX 2012)*, 485–492. Oslo: University of Oslo.

Giouli, Voula & Aggeliki Fotopoulou. 2014. Linguistically motivated language resources for sentiment analysis. In *Proceedings of the Workshop for Lexical and Grammatical Resources for Language Processing. Collocated with the 25th International Conference on Computational Linguistics (COLING 2014)*, 39–46. Dublin: Dublin City University.

Gross, Maurice. 1975. *Méthodes en syntaxe. Régime des constructions complétives.* Paris: Hermann.

Gross, Maurice. 1982. Une classification des phrases "figées"du français. *Revue Québécoise de Linguistique (RQL)* 11(2). 151–185.

Gross, Maurice. 1995. Une grammaire locale de l'expression des sentiments. *Langue Française* 105. 70–87.

Gross, Maurice. 1996. Les formes *être Prép X* du français. *Lingvisticae Investigationes* 20(2). 217–270.

Gross, Maurice. 1998a. La fonction sémantique des verbes supports. *Travaux de linguistique* 37. 25–46.

Gross, Maurice. 1998b. Les limites de la phrase figée. *Langage* 90. 7–23.

Lamiroy, Béatrice. 2003. Les notions linguistiques de figement et de contrainte. *Lingvisticae Investigationes* 26(1). 1–14.

Leclère, Christian. 1976. Datifs syntaxiques et datif éthique. *Méthodes en grammaire française* 7396.

Leeman, Danielle. 1991. Hurler de rage, rayonner de bonheur: remarques sur une construction en de. *Langue française* (91). 80–101.

Levin, Beth. 1993. *English verb classes and alternations: A preliminary investigation.* Chicago: University of Chicago Press.

Mathieu, Yannick Yvette & Christiane Fellbaum. 2010. Verbs of emotion in French and English. In *Proceedings of the Fifth Global WordNet Conference.* Mumbai, India.

Mathieu, Yvette Yannick. 1999. Les prédicats de sentiment. *Langages* 136. 41–52.

Mathieu, Yvette Yannick. 2005. A computational semantic lexicon of French verbs of emotion. In James G. Shanahan, Yan Qu & Janyce Wiebe (eds.), *Computing attitude and affect in text*, 109–123. Dordrecht: Springer.

Mini, Marianna. 2009. *Linguistic and psycholinguistic study of fixed verbal expressions with fixed subject in Greek: A morphosyntactic analysis, lexicosemantic gradation and processing by elementary school children.* Patras, Greece: University of Patras dissertation.

Moustaki, Argyro, Mavina Pantazara, Aggeliki Fotopoulou & Marianna Mini. 2008. Comment traduire les noms d'émotion. Etude contrastive entre le grec moderne et le français. *Discours. Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics* (3).

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

Pantazara, Mavina, Aggeliki Fotopoulou, Marianna Mini & Argyro Moustaki. 2008. La description des noms de sentiments du grec moderne. *Lingvisticae Investigationes* 31(2).

Plutchik, Robert. 2001. The nature of emotions. *American scientist* 89(4). 344–350.

Tutin, Agnès, Iva Novakova, Francis Grossmann & Cristelle Cavalla. 2006. Esquisse de typologie des noms d'affect à partir de leurs propriétés combinatoires. *Langue française* (2). 32–49.

Villavicencio, Aline, Timothy Baldwin & Benjamin Waldron. 2004. A multilingual database of idioms. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)*, 1127–1130. Lisbon, Portugal.

Wilson, Theresa, Janyce Wiebe & Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 347–354.