Addressing Template Composition and Parametric Features of Aspect and Modality in Ontological Semantic Acquisition of Phrasal Verbs

Julija Korostenskaja

Abstract. The article uses the theory of ontological semantics to propose methodological and practical solutions for the acquisition process of selected phrasal verbs. The *Knowledge Base Acquisition Editor* (KBAE) is adopted as a software for project implementation. Since in ontological semantics, each lexical unit is to be described in terms of its syntactic and semantic structure, the article provides an analysis of both components of representation. At the syntactic level, the pre-established seed templates are adjusted so as to avoid misleading interpretations and to account for causative constructions. At the semantic level, focus is placed on the application of the ontological parametric features of aspect and modality in the meaning representation of phrasal verbs. It is argued that the categories of aspect and modality help grasp the difference between the PV in question and its one-word counterpart. It is believed that the findings of the present study may have further research implications: first, an assessment of the frequency of causative constructions could help finalise the issue of Template 1, which has been found to be misleading. Second, a study of PVs conveying metaphorically extended meanings, which oftentimes turn out to be deducible, could foreground the feature of universality in languages. It could therefore be desirable to establish core PVs, the non-compositional meanings of which would pertain to the domain of cross-linguistically interpretable metaphoricity .

Keywords: ontological semantics, natural language processing, acquisition, parametric features, phrasal verbs.

Introduction

The present paper explores some of the issues that emerged during the acquisition process, i.e., computerised description of the syntactic structure and the meaning conveyed, of a selection of phrasal verbs in the framework of ontological semantics (OS), a theory of meaning which posits as its major goal applicability both to the study of natural language and computer processing. OS is a task-driven enterprise; consequently, answers to questions about what meaning is and how it is to be treated within the theory are sought using the method of deduction, i.e., the bottom-up approach: first, by analyzing a particular class of phenomena and then by formulating the ontological view (Nirenburg & Raskin, 2004).

To describe the meaning of both lexical items and grammatical structures, OS uses the ontology, a language-independent inventory of concepts. Meanwhile building up an enumerative, language-dependent lexicon is one of the most immediate practical tasks of the theory. All acquisition tasks, and consequently, the acquisition of phrasal verbs, are implemented by means of the Knowledge Base Acquisition Editor (KBAE), a computer software used "to expand… the ontology and lexicon" (Spartz, Malaia & Falk, 2005, p.2). The acquisition process is grounded on the principle of "complete coverage" whereby each sense is provided with a lexical entry of its own (Nirenburg & Raskin, 2004, p.274). Along the principle of complete coverage considerations of economy play an important role. The combination of these two principles favours the approach that lists all qualifying phrasal verb combinations irrespective of their inter-pretability thereby ensuring "single architecture and control environment", which is perceived as a prerequisite for quality output in computer applications (ibid., 10). Other theoretical premises of the theory have been presented elsewhere (Al-Hashimy, 2007; Televnaja, 2005; Malaia, 2005) and will not be discussed here due to practical constraints.

In this discussion, the term *phrasal verb* (PV) is used to refer to a verb-particle combination which satisfies at least one of the criteria specified in Section 1.2 below. It should be pointed out that that the choice of the term *phrasal verb* is stipulated by the fact that it seems to refer to the semantic domain the best, while the other frequently used alternatives, *verb-particle constructions and verb-particle combinations*, underscore the syntactic approach to the class. Another consideration has to do with the status of the particle in PV analysis. Since the boundary between the preposition and the particle cannot always be easily defined (see, e.g., Sroka (1972), Makkai (1972), Bolinger (1971), and, more recently, O'Dowd (1998)) and due to the fact that this distinction is redundant in the ontological description of PVs, it is not be reflected in the templates proposed.

The aim of this research was to supplement OS with a microtheory of phrasal verbs. The present paper addresses two of the most important issues that emerged during the acquisition process. At the syntactic level, it was the adjustment of one of the seed templates which appeared to produce misleading interpretations. At the semantic level, it was the application of the ontological parametric features of aspect and modality to capture the shades of meanings of PVs which differentiate them from their single-verb counterparts.

The current analysis is based on the findings obtained from the acquisition process of phrasal verbs selected from

Longman Pocket Phrasal Verbs Dictionary, range L-P inclusive, and later verified with the FrameNet selection of phrasal verbs. This approach is in line with the "rapid propagation" principle of OS (Nirenburg & Raskin, 2004, p.274). Given these theoretical premises, the present article will focus on two aspects pertinent to the acquisition of phrasal verbs in the KBAE: their syntactic representation, i.e., template adjustment, and their semantic representation, i.e., preferences made in the application of OS parameters.

Classification of PVs in OS

Even a brief overview of literature suggests that phrasal verbs (hereinafter referred to as PVs), more often referred to as verb-particle constructions, have been viewed from a variety of perspectives. Thus PVs have been classified according to the meaning conveyed by the particle (Levin, 1993); prosodical features (Makkai, 1972; Sroka, 1972; Bolinger, 1971); the number of elements constituting a combination (Makkai, 1972), the entire meaning of the group (compositional vs non-compositional) (Pedersen & Nimb, 2000; Fellbaum, 1998); and the nature of the nominal group following the verb (Aarts, 1989). We have felt, however, that it is necessary to come up with a more general classification that would incorporate at least some of the above ideas and would be applicable to processing from the lexicographical viewpoint. We have found that, while the non-compositional combinations are analysed regardless of the approach and are unanimously attributed to the class of PVs, the more compositional cases (up till literal interpretations of combinations) have been treated as more common and, therefore, less interesting. To delineate the coverage of the work and provide a systematic, yet broad enough approach to PVs, we have decided to examine both compositional and non-compositional cases. Apart from purely compositional cases with literal meaning, we distinguish between two kinds of non-literal meanings, in a way similar to Bolinger (1971) and Makkai (1972). First, it is the metaphorical meaning, which may be viewed as an extension of the compositional (literal) meaning (see also discussion in Copestake & Briscoe, 1996); and second, it is the non-compositional meaning, in the sense as it is unanimously distinguished in all approaches. Within these three kinds of meaning, further distinctions may be drawn, the discussion of which is presented elsewhere (e.g., Baldwin, 2002; Korostenskaja, 2009).

Within our classification, we have argued for the separation of PVs characterised by aspectual, modality-coloured (in order to avoid ambiguity, we deliberately avoid the term **modal**), and metaphorical meaning, on the one hand, from non-compositional PVs, on the other, due to the following considerations. The meanings of the first three cases are predictable to a varying degree, whereas that of the last one is not. This broad coverage is not always accepted, but has to be considered by the ontology; which then should either serve more general purposes and provide a definition for any verb + particle combination; or treat the first three cases as purely compositional ones and thus neglect them, accounting for only those idiosyncratic instances that are indivisible in meaning and are non-compositional indeed.

After a preliminary analysis of PVs, we have found several semantic regularities that helped account for the notion of the (partial) non-compositionality of a combination. In our approach, the combination qualifies for falling under the category of phrasal verbs if at least one of the below conditions is satisfied:

1) the combination has a parametric aspectual feature;

2) the combination has a parametric modality feature;

3) the combination has a figurative meaning (due to either or both of the elements having a metaphorically extended meaning);

4) the combination has a non-compositional meaning (and may have both transparent and non-transparent structure).

In a way, the classification above reflects Bolinger's (1971, p.6) statement that "being or not being a phrasal verb is a matter of degree" by specifying what those degrees may be. Under the suggested approach, (non)-compositionality is no longer seen as an obligatory premise; it may or may not be present in the combination; moreover, any degree of it, hence its absence, is acceptable. The above criteria also account for the large group of combinations which are interpretable and have a transparent meaning structure.

Syntactic Representation of PVs in OS

The general ontological approach to PVs is that they present a class of multiword expressions characterised by discontinuous structure while their meaning is frequently non-compositional. The recurrent discontinuity of a combination was taken as a major criterion for the selection, thereby foregrounding structural regularities that depend on the admissible combinations within the group. After the regularities had been established, they were grouped into four basic **seed templates**. During the acquisition of new lexical items, in order to adapt to idiosyncrasies of the behavior of individual PVs at a low cost, the templates were allowed to undergo minor modifications, which meet the aforementioned "rapid propagation" principle. For example, to handle polysemous relations, when the semantic structure of a phrasal verb had more than one option for meaning representation, one template could be assigned a semantic structure consisting of two or several concepts. This is possible when senses suggested are too different to be given one sem-struc, but not different enough to be assigned different entries in the lexicon.

Initially, the following four different structural classes of PVs, resulting in four seed templates, were found:

1) a verb and a particle form a combination with an optional noun phrase preceding or following the particle. The possible combinations are VP NP P, VP P NP, VP P (e.g.: *hang sth up, hang up sth, hang up*).

2) a verb and a particle form a combination with a mandatory noun phrase preceding or following the particle. The possible combinations are VP NP P, VP P NP (e.g.: *mark down sth, mark sth down*).

3) a verb and a particle form a combination with a mandatory noun phrase following the particle. The order is strictly VP P NP (e.g.: *tell on smb*).

4) a verb and a particle form a combination. The order is VP P (e.g.: *pop off*).

For practical considerations, the categories of preposition and particle were indiscriminately referred to as **P** in the possible orders above and **prep** within the templates themselves. Throughout the acquisition process and the present article, the non-verbal element of the PV groups is referred to as particle, since it is the most neutral of the three terms that have been employed in relevant literature, the other two being prepositional adverbs, and adpreps. In the templates, in line with the goal of maintaining homogeneity with other applications within the KBAE, the same non-verbal element is referred to as preposition.

Since this section is concerned specifically with eliminating inaccuracies of representation contained within template 1, the initial model of the template is provided below:

(1) *(verb-particle-v1*
(anno
(def "...")
(ex "...")
(comments "..."))
(syn-struc
(1
((np ((root $var1) (cat np)))
(root $var0) (cat v)
(np ((root $var2) (cat np)))
(prep ((root up) (cat prep))))
(2
((np ((root $var1) (cat np)))
(root $var0) (cat v)
(prep ((root up) (cat prep)))
(np ((root $var2) (cat np) (opt +))))) ;; final np optional
(sem-struc
((1 2)
(EVENT
(agent (value ^$var1) (sem HUMAN))
(theme (value ^$var2) (sem EVENT OBJECT))
(...))))
)

As can be seen, there are three types of information given in the template: the general information on the phrasal verb, the syntactic information, and the semantic information. The upper two parts of the template are the lexical entry itself and the annotation anno. Every lexical entry is schematically presented as verb-particle-X; where X points to the numerical order of senses distinguished for the combination in question. Notably, if the same verb-particle combination expresses two distinct meanings which are stored in two structurally different templates (e.g., one use with the interchangeable order and one use with the strict order), the number in the opening line

for both PVs will be the same. Designed specifically for the human consumer, the annotation provides the definition as well as acquirer's specific comments on use peculiarities of the PV in question. All computationally relevant information is presented in the syntactic and the semantic structures.

As has already been mentioned, upon closer analysis, we have discovered that Template 1 could produce misleading interpretations of the PVs it describes. First of all, PVs falling under Template 1 appear very seldom: out of all PVs in the selection, only *open out*, *open up*, *play back*, *pull round*, *pull through*, and *pay out* fit the template. Besides, much of the flexibility of these PVs to appear with or without the final noun group (encoded as ";; final np optional" in the relevant var^ 2 line of the syn-struc of the template) depends on whether the combination is capable of producing a causative reading. If it is, then, provided that the grammatical subject causes the action, the structure needs an object argument. On the other hand, the structure does not need the direct object position to be filled when it has a non-causative reading. In this case, the patient of the causative structure is the grammatical subject. In the result of this "systematic" or "constructional" polysemy, of which causative constructions are an instance (Copestake & Briscoe, 1996, p.17ff.), we end up having a misleading representation of a single PV which in fact is two polysemous PVs, the non-causative PV[1] and the causative PV[2]. The following instance of *open up* exemplifies this confusion:

(2) *(open-up-v3*
(anno
(def "if a new shop, business etc opens up or is opened up, someone starts it".
(ex "This guy opened up several expensive restaurants and night clubs".
(comments "pronoun issue; can only go between verb and postposition"))
(syn-struc
(1
(np ((root $var1) (cat np)))
(v (root $var0) (cat v))
(np ((root $var2) (cat np)))
(prep ((root up) (cat prep))))
(2
(np ((root $var1) (cat np)))
(v (root $var0) (cat v))
(prep ((root up) (cat prep)))
(np ((root $var2) (cat np) (opt +))))) ;; final np optional
(sem-struc
((1 2)
(FOUND-ORGANIZATION
(agent (value ^$var1) (sem HUMAN) (relaxable-to ORGANIZATION))
(theme (value ^$var2) (sem CORPORATION))
(aspect (phase begin) (iteration 1) (scope FOUND-ORGANIZATION)))))

Used in its initial form, Template 1 accounts for the optionality of the direct object; however, it does not account for the change in the grammatical subject. The examples of the two possible uses of *open up* are given below:

(3) *John opened up a new restaurant on Thursday.* (causative reading);
A new restaurant opened up on Thursday. (non-causative reading);

To compare, Rappaport et al. (1993, p.44) suggests treating causative constructions as a complex event with the following relationship between causative and non-causative structures of the verb meaning for BREAK:

(4) Causative BREAK: [x cause [y become BROKEN]]
Non-causative BREAK: [y become BROKEN].

In the cases like (3) and (4) above, there are two interpretations available: somebody opens/ breaks something, or something opens/ breaks itself, without the specification of the driving force, i.e., the *agent*. Moreover, PVs categorised under Template 1 also seem to be represented by verbs which take direct objects of restricted semantics, e.g., *hang up* [the phone], *pay out* [the money], *play back* [music when referring to machines]. In all of these cases, the direct object is dependent on the verb to the extent that it cannot appear in the subject position. In this way, Template 1 may have one of the three possible functions:

1) It may convey the interchange between the agent and the patient in the causative construction. Omission of the final *np* signifies the change of structure from the causative reading into the non-causative reading. If the omission does take place, the template becomes misleading since it allows for cases where the obligatory grammatical direct object will not be present, while the subject position will be taken not by the patient, but by the agent of the action, thereby producing structures like *John opened up* (in the sense 'became expansive, began to talk freely');

2) It may convey redundancy of the direct object position when the verb appears in its intransitive use (i.e., the non-causative reading), e.g., *pull through*;

3) It may convey the mandatory use of the direct object (with restricted semantics) when the verb appears in its transitive causative use e.g., *play back*.

To sum up, the ambiguity of Template 1 lies in the fact that it allows for both causative and non-causative structures and may therefore prompt incorrect results. Although the semantic structure of the template does show the case role relationship of the arguments with the verb, it does not capture causativity. To account for the change in the meaning of the entire PV, we have two options: either to introduce a new parameter responsible for taking over causativity (just like *cause* in the scheme above), or to divide the meaning of the verb between two different templates. We have chosen the latter option aiming to avoid the addition of new ontological parameters and resources as long as the existing ones can cope with an emerging problem (see also Moon, 2000, p.100).

In light of the above, we have suggested that Template 1 should be eliminated. Consequently, the causative and the non-causative readings have to be split into two different, but already existing templates. This will help account for the the

causative feature and will make it possible to formally treat the relevant cases as instances of homography. Notably, other structures are not potentially affected by such an error due to the fixed word order in the combination they describe. Thus Template 2, the only other template which allows the interchange between the particle and the direct object does not involve the grammatical subject in the variation and is not misleading. Therefore, we suggest dividing the syntactic representation of the problematic PVs initially categorised under Template 1 between Template 2 to cover the causative reading; and Template 4, to cover the non-causative reading. In the cases where there is no causative reading and all three orders are possible, i.e., with the verb itself used transitively (when the direct object may either precede or follow the particle), and with the verb used intransitively, the representation can be adequately handled either by adding an extra line to Template 2 stating the optionality of the final np (which is justified), or by using the existing resources and dividing the information between two templates to account for the transitive and intransitive uses of a verb, in the same manner as in the case of causative versus non-causative structures.

The change in the syn-struc will naturally be reflected by the sem-struc of each of the entries. If used transitively, the phrasal verb will have two case roles, the agent and the patient, both being expressed by ontological concepts. If used intransitively, the agent's SEM slot will only have the default filler HUMAN, but will not be formally represented by a variable noun; the patient will express the concept of ^$var1. Below we provide the sem-strucs of both the agentive (and hence transitive and causative) and the non-agentive (and hence intransitive and non-causative) instances of *open up*:

(5) Agentive use:

(sem-struc
((1 2)
(FOUND-ORGANIZATION
(agent (value ^$var1) (sem HUMAN) (relaxable-to ORGANIZATION))
(theme (value ^$var2) (sem CORPORATION))
(aspect (phase begin) (iteration 1) (scope FOUND-ORGANIZATION)))))

(6) Non-agentive use:

(sem-struc
(FOUND-ORGANIZATION
(agent (sem HUMAN))
(theme (value ^$var1) (sem CORPORATION))
(aspect (phase begin) (iteration 1) (scope FOUND-ORGANIZATION))))

Semantic Representation of PVs in OS

The ontological parametric features of aspect and modality, which express attitudinal information relevant to the main concept of the structure and which can be conveyed by lexical units and syntactic structures alike, have proved extremely valuable in the meaning representation of PVs. As has been stated above, these features have been used as

classification criteria for PVs since they help account for their (partial) non-compositionality.

Aspectual information of the verb, under the term Aktionsart, has been frequently discussed in literature (e.g., Bolinger, 1971; Brinton, 1988; Comrie, 1976). Within OS, the aspectual information or value is described in terms of phase and iteration. The phase refers to the stage of the development of an action and has four values: begin, continue, end, and begin-continue-end, with the latter describing instantaneous actions. Iteration refers to the number of times an action takes place and gives the numerical value or, if the number is not contained explicitly in the semantics of the PV, the value multiple.

In the semantic description of PVs, aspectual information was frequently employed to help express complex concepts, which otherwise seemed highly problematic, since in the latter case the decision would have to be made as to which part of the complex event was less important and could therefore be sacrificed. For example, aspectual information was provided to represent the semantic structure of a group of PVs with the main verb *pull* and referring to the change in motion, as in *pull in* expressing two events: driving and stopping. The fact that the vehicle stops is essential in the semantic structure of the phrasal verb; and, since one can say that stopping involves the final phase in the range begin-continue-end, the cessation of motion was rendered via the end phase in the aspect line of the sem-struc with the main concept for the phrasal verb identified as DRIVE. The template is given below:

(7) *(pull-in-v1*
(anno
(def "to drive to the side of the road or to a place where you can stop your car")
(ex "I'm going to pull in at the next garage — I want to check the tires".)
(comments "pronoun issue: can only go between verb and postposition".))
(syn-struc
(np ((root $var1) (cat np)))
(v (root $var0) (cat v))
(prep ((root in) (cat prep))))
(sem-struc
(DRIVE
(agent (value ^$var1) (sem HUMAN) (relaxable-to WHEELED-ENGINE-VEHICLE))
(theme (sem WHEELED-ENGINE-VEHICLE))
(destination (sem PLACE))
(aspect (phase end) (iteration 1) (scope DRIVE))))

Aspectual information is very often followed by some modality information. The next template exemplifies this combination by storing the aspectual continue phase and relevant values of the potential and the epiteuctic modalities in the sem-struc of the PV *push ahead*:

(8) *(push-ahead-v1*
(anno
(def "to continue doing something, even though there are problems")

(ex "I'd like you all to push ahead and get the job done as soon as possible".)
(comments "...")
(syn-struc
(np ((root $var1) (cat np)))
(v (root $var0) (cat v))
(prep ((root ahead) (cat prep))))
(sem-struc
(WORK-ACTIVITY
(agent (value ^$var1) (sem HUMAN))
(theme (sem WORK-ACTIVITY))
(aspect (phase continue) (scope WORK-ACTIVITY))
(modality (type potential) (value 0.5) (scope WORK-ACTIVITY))
(modality (type epiteuctic) (value 0.5) (scope WORK-ACTIVITY))))

As can be seen, the semantic structure expresses two aspects of parametric information: (a) the stage of the process, and (b) the (partial) success of an action. The implication is that, although the action may be requiring much effort and its performability is questioned (which is rendered by the potential modality here), it has nevertheless (so far) been successful, the latter being rendered by the epiteuctic modality. A more detailed description of modalities is provided below.

Within the OS theory, modality refers to the agent's attitude expressed in the sentence and is subdivided into seven types: epistemic, epiteuctic, deontic, volitive, potential, evaluative, and saliency. Each of these types may have a value ranging from 0, referring to the negative realisation of an event (such as impossibility, undesirability, failure, etc.) to 1.0 referring to the maximum positive realization of the event expressed through an ontological concept (such as certainty, desirability, success, etc) (Nirenburg & Raskin, 2004, 249ff.). With respect to PVs, modality proved invaluable in identifying their pragmatic context and consequently rendering their meaning more precisely, thus contrasting PVs with their one-word counterparts.

As is known, PVs can often be distinguished not only by their less formal tone, but also by a more vivid representation of a particular concept, which can be successfully rendered through the relevant modality with a corresponding value. The acquisition process and subsequent revisions have shown, however, that there are two methodological decisions to be made with respect to the coverage of modality as an ontological parameter:

1) whether to allow modality to denote different ranges in attitudinal characteristics of the main concept and thus to bring the differences between synonymous expressions to a finer grain;

2) whether to apply modality to the concept expressed by var2 in the sem-struc and thus treat modality as a heterogeneous semantic parameter.

For example, regarding the first problem, we have had two solutions to propose with respect to the PV *pop in*: first, to account for its informality and colloquialism; and second, to render the meaning of temporariness. Analyzing the pragmatic

content of the given PV, we have come to the conclusion that there is a difference between opinionism on the one hand and informality on the other. Although both are relevant in identifying the exact meaning of the combination, we have not been able to establish a finer scale of granularity other than to refer to evaluative modality on both occasions, but to assign each instance a different value. Thus, evaluation would be set as one of the semantic features for PVs like *look down on smb* and *pop in*, both having low value of the evaluative modality. While this problem might be regarded as irrelevant in the acquisition of other kinds of lexical units, we believe that it is one of the tenets in the acquisition of PVs because of the specificity of the latter. To distinguish between low estimate and informality, we have used different values of the same modality. In the sem-struc of *look down on smb*, the evaluative modality has been assigned the value below 0.4, while in the sem-struc of *pop in* the value assigned is below 0.7, thereby representing the difference between low opinion toward the theme on the hand and on the other, the informal tone of the PV. The semantic structures of both PVs are given below in (8) and (9) respectively:

(9) *(look-down-on-v1*
    <...>
    (sem-struc
    (RESPECT
    (agent (value ^$var1) (sem HUMAN))
    (theme (value ^$var2) (sem OBJECT))
    (modality (type evaluative) (value < 0.4) (scope RESPECT))))
)
(10) *(pop-in-v1*
    <...>
    (sem-struc
    (COME
    (agent (value ^$var1) (sem HUMAN))
    (theme (sem SOCIAL-ROLE))
    (modality (type evaluative) (value < 0.7) (scope COME))))
)

Regarding the second problem, we leave the question open whether modality as such should aim only at the specification of the main event or the theme that is affected by the semantics of the verb. Although Nirenburg & Raskin (2004) attribute modality only to events (stored in the slots for the main concept of the sem-struc), in the cases where the evaluative and the saliency modalities are to be employed, we have often found that the concept expressed by ^$var2 could just as well serve the scope for the established value of modality. For example, lack of respect in *look down on* spreads onto the grammatical object in ^$var2; as a result, the latter can be characterised by the low value as well, thereby pointing to the direction of the spread as well as the relevant area of modality and producing the following line in the sem-struc:

(11) *(modality (type evaluative) (value < 0.4) (scope OBJECT))))*

As a consequence, we have had to account for the evaluation of the theme on the one hand, and for the evaluation of the PV itself, on the other. In the case of *look down on*, we believe that a qualitative semantic representation should

consider two factors: first, it should establish the numeric value of the influence of the main event on the theme (by pointing out that *look down on* is coded in the ontology within the concept RESPECT with a low value for the concept expressed in the theme. Secondly, the PV itself has to be placed on a certain scale in the hierarchy among the synonymous verbs and other multi-word expressions which all share the main meaning component RESPECT. We believe that modality is capable of grasping the two differences provided that a) it is either more strictly defined in the inventory of modalities than it is now, possibly, allowing for certain flexibility depending on the application; or b) it is treated as a larger concept and a cover term allowing for finer differentiation within its structure. At the present stage, and as our acquired verbs reveal, while maintaining the tendency to assign modality to the main concepts of the sem-struc, we were often inclined to rather assign a particular value of modality to the concept expressed in var2, justifying this choice by the fact that the semantic representation of the phrasal verb gained more precision.

The fact that modality may spread over var2 is also reflected in our treatment of the saliency modality. Saliency has been exemplified by the verb *emphasize* defined roughly as "attach importance to something" (Raskin, from the interview with the author, 25 September, 2003). We believe that a verb like *ignore* may be placed at the other end of the saliency scale and allow for the possible variation in between. Then verbs like *neglect* and *ignore* could be defined not only in terms of the epistemic modality, but also in terms of the degree of their influence on the *theme* as suggested by the treatment of *emphasize*. The acceptance of such treatment would permit a more precise description of the semantics of *pass off* defined as "ignore something or pretend it is not very serious" (Longman, 2001, p.250), for example. This approach to representing the meaning of the given PV would refer not so much to the mere failure to discern a certain idea expressed by the theme (if using the epistemic modality), but rather to the agent's intentional behaviour with respect to the theme. Coming back to the issue of coverage, there is still one more question to be answered, and that is the way we should mark modality in the template: whether as spreading over the main concept of the sem-struc, or over its theme. If it is the main concept, we establish the proportional relationship between *IGNORE* and its "amount" stored in *pass off*. If it is the theme, we establish the "weight" of importance of the concept expressed by the theme as "seen" by the verb. Below we provide both versions: the full template in (11) and, for space considerations, the semantic structure of the relevant verb in (12):

(12) *(pass-off-v1*
    (anno
    (def "to ignore something or pretend it is not very serious, because you do not want it to hurt you".)
    (ex "laughing lightly, Claire passed the remark off as a joke".)
    (comments "pronoun issue; can only go between verb and preposition".))
    <..>

```
(sem-struc
((1 2)
(IGNORE
(agent (value ^$var1) (sem HUMAN))
(theme (value ^$var2) (sem EVENT))
(modality (type saliency) (value > 0.6) (scope IGNORE))))
)
```
(13)
```
(sem-struc
((1 2)
(IGNORE
(agent (value ^$var1) (sem HUMAN))
(theme (value ^$var2) (sem EVENT))
(modality (type saliency) (value < 0.4) (scope EVENT))))
)
```

As can be seen, depending on whether we attribute the saliency modality to the main event of the PV or its theme, we have given it different values. In the first case, we have established the strength of IGNORE as expressed in *pass off*. We imply that the full value of IGNORE being 1 is (at least) partially realised in *pass off*, which, under this approach, refers to 'complete or partial ignoring' of the event of the theme formally found in the range beyond 0.6. In the second version, however, the saliency modality has been assigned a value from the other end of the spectrum, which suggests that the importance of the theme as rendered by the verb is low. Our own opinion is that the saliency modality is better realised when applied to the theme, since otherwise it becomes (almost) indistinguishable from the epistemic modality exemplified by verbs and phrases like *doubt*, *assert*, *be certain*, etc. We also believe that the epistemic and the saliency modalities are epistemologically different. We believe that the epistemic modality does not so much weigh the situation (the latter being expressed in the theme) but rather, taken in its absolute value, estimates the degree of relevance or feasibility. In contrast, the saliency modality takes the action itself in its absolute value (expressed in the main concept which, in its turn, can be gradable), but questions the saliency and content of the theme. With these distinctions made and by definition, referring to "the importance the speaker attaches to a component of text meaning" (Nirenburg & Raskin, 2004, p.253), in our analysis, the saliency modality should be theme-oriented rather than event-oriented. Consequently, this double applicability of the concept of modality on the whole may prompt the proposal of several semantic structures. The example above is an illustration of this case.

What this discussion also suggests is that verbs like *neglect* and *doubt* may require a specification not only by the epistemic modality, but also by the saliency modality, as in, for example, *neglect the answer* meaning roughly "not give credit" and opposite to "emphasize the answer" on the one hand; and on the other, having a meaning similar to "deny" and opposite to "assert". We have no doubt that there is a counter-argument to our viewpoint on this issue, especially from the perspective of the circularity of a semantic description, and leave the question open, concluding with one more example from the complex concept domain.

The PV *pass by* has two relevant components of meaning to be represented. On the one hand, there is a certain EVENT that takes place under normal circumstances, unaffected in any way by the human expressed through the patient role (it has to be pointed out here that, throughout the acquisition process, we have operated with concepts available in the KBAE and, although we hesitated accepting an opinion that *pass by* stands for (any) EVENT, we understand the logic behind it: anything may happen, and, this generalization presupposes that the positive value expressed through the concept EVENT equals occurrence). On the other hand, we have to account for the fact that, by definition, the patient either does not notice the event at all or is unaffected by it, which is not so much IGNORE, i.e., deliberate failure to notice, but rather involuntary, so that the two events never appear in the same place in a timeline and occur independently form each other. Below we provide our earlier version of the semantic structure, rejected in further analysis:

(14)
```
(sem-struc
EVENT
(agent ( sem nothing))
(theme (value ^$var1) (sem EVENT))
(patient (value ^$var2) (sem HUMAN))
(modality (type saliency) (value < 0.3) (scope HUMAN))))
```

In this case, we treat the unaffected object argument as the patient, since the verb refers to the degree of success of the influence rather than to the relationship between the event and the person affected by it. The question that is of interest from the methodological point of view is whether the object argument, the unaffected patient, is a patient or should rather be referred to as theme. However, we have to find a way to mark the **degree** of conception by the object argument of the event. We propose the following approach. Our suggested semantic representations have often contained a combination of the epiteuctic modality and the epistemic modality, in order to point out the overcoming of certain obstacles to a particular event. By using each of these modalities with a value 1 we refer to the independence of the event expressed in the subject argument or the object argument (rather than the absence of difficulties in taking place as such). To show the absence of the relation between the subject and the object, we have used the saliency modality with a low value, which then establishes the patient's/ speaker's perspective on the event, and which is in compliance with the definition of the modality. The final semantic representation of *pass by* is given below:

(15)
```
(pass-by-v2
(anno
(def "if an event passes you by, it happens without your noticing it much or being affected by it")
(ex "My mother lived in a remote village, and all the excitement of the 1960s passed her by".)
<…>
(sem-struc
(EVENT
(agent (sem nothing))
(theme (value ^$var1) (sem EVENT))
(patient (value ^$var2) (sem HUMAN))
```

(modality (type potential) (value 1) (scope EVENT))
(modality (type epiteuctic) (value 1)(scope EVENT))
(modality (type saliency) (value < 0.3) (scope EVENT))))
)

Conclusions

The present study was concerned with some of the methodological issues entailed in the acquisition of a selection of English phrasal verbs in the KBAE. At the syntactic level, the templates were adjusted so as to account for causative constructions and to produce correct interpretations. At the semantic level, the ontological parametric features of aspect and modality were used in order to capture the meaning of PVs more accurately and at the same time, to contrast it with the meaning of the relevant one-word counterparts.

The findings of the present study may be used in further investigations into the semantics of the given selection of PVs at the more practical level, by incorporating corpora into analysis. First of all, it would be interesting to assess the frequency of causative constructions, thereby finalising the issue of Template 1. Other corpus-based approaches could focus on the pragmatic functions of PVs as well as on PVs conveying metaphorically extended meanings. These directions are also in line with recent research tendencies into verbal semantics and corpus-based reconsiderations (e.g., Liu, 2011; Close&Bas, 2010). Lastly, the implications of the present study may be considered at the broader, cross-linguistic level. It is believed that, although non-compositionality is indeed a characteristic feature of many PVs, the metaphorically extended meanings they convey may nevertheless be sometimes logically deducible due to the feature of universality in languages. It could therefore be desirable to establish **core** PVs, the non-compositional meanings of which would pertain to the domain of cross-linguistically interpretable metaphoricity.

References

1. Al-Hashimy, A. S. H., 2007. Acquisition of Lexical Units in Ontological Semantics. Proceedings of the International Symposium on Computational Intelligence and Intelligent Informatics [online], ISCIII '07. 28-30 March, 2007, pp.233–237 Available at: http://www.zuj.edu.jo/conferences/ICIT07/PaperList/Papers/407amal.pdf. [Accessed: April 2012].

2. Aarts, B., 1989. Verb-preposition Construction and Small Clauses in English. Journal of Linguistics, 25, pp.277-290. Dordrecht: Kluwer Academic Publishers. http://dx.doi.org/10.1017/S0022226700014109

3. Baldwin, T., 2002. Verb-particle Extraction. (notes) [online], Available at: http://mwe.stanford.edu/20020730.pdf. [Accessed September 2012].

4. Bolinger, D., 1971. The Phrasal Verb in English. Cambridge, MA: Harvard University Press.

5. Brinton, L., 1988. The Development of English Aspectual Systems. Cambridge: Cambridge University Press.

6. Close, J., Bas, A., 2010. Current Change in the Modal System of English: a Case Study of *must*, *have to* and *have got to*. *In*: Lenker, U., Huber, J., and Mailhammer, R (eds.), The History of English Verbal and Nominal Constructions (Volume 1) of English Historical Linguistics, 2008: Selected Papers from the Fifteenth International Conference on English Historical Linguistics (ICEHL 15), Munich 24-30 August 2008 [online], Amsterdam: John Benjamins, pp.165-182. Available at: http://www.ucl.ac.uk/english-usage/projects/verb-phrase/Current-change-modal-system.pdf. [Accessed March 2012].

7. Comrie, B., 1976. Aspect: an Introduction to the Study of Verbal Aspect and Related Problems. Cambridge: Cambridge University Press.

8. Copestake, A., Briscoe, T., 1996. Semi-productive Polysemy and Sense Extension. Lexical Semantics: the Problem of Polysemy. Pustejovsky, J. and Boguraev, B. (eds.) Oxford: Oxford University Press, pp.15-69.

9. Fellbaum, C., 1998. Towards a Representation of Idioms in Word Net. Proceedings of the COLING/ACL Workshop on Usage of WordNet in Natural Language Processing Systems [online], Montreal. Available at: http://citeseer.ist.psu.edu/bhandari02machine.html. [Accessed March 2011].

10. Korostenskaja, J., 2009. English Phrasal Verbs in Ontological Semantics: Lexical Acquisition and Semantic Representation. Saarbrücken: VDM Verlag.

11. Levin, B., 1993. English Verb Classes and Alternations: A Preliminary Investigation. Chicago: University of Chicago Press.

12. Liu, D., 2011. The Most Frequently Used English Phrasal Verbs in American and British English: a Multicorpus Examination. TESOL Quarterly 45.4, Chicago: Wiley-Blackwell, pp.661-688..

13. Makkai, A., 1972. Idiom Structure in English. The Hague: Mouton.

14. Malaia, E., 2005. Digital Identity Management in Ontological Semantics: Methodology and Practice of Domain (Ph.D. thesis), CERIAS Tech Report [online], Purdue University, USA. Available at: http://www.cerias.purdue.edu/apps/ reports_and_papers/view/2863. [Accessed March 2012].

15. Moon, R., 2000. Lexicography and Disambiguation: the Size of the Problem. Computers and Humanities [online], 34:1-2, pp.99–102. Available at: http://www.kluweronline.com/issn/0010-4817//contents. [Accessed January 2012].

16. Nirenburg, S., Raskin, V., 2004. Ontological Semantics. Cambridge, MA: MIT Press.

17. O'Dowd, E. M., 1998. Prepositions and Particles in English: A Discourse-Functional Account. New York, Oxford: Oxford University Press.

18. Pedersen, B. S., Nimb, S., 2000. Semantic Encoding of Danish Verbs in SIMPLE — Adapting a Verb-framed Model to a Satellite-framed Language [online]. Available at: http://cst.dk/simple/lrec/nylrec.doc. [Accessed August 2012].

19. Raskin, V., 2003, September 25. Personal interview.

20. Rappaport, M., Laughren, M., and Levin, B., 1993. Levels of Lexical Representation. Semantics and the Lexicon. Pustejovsky, J. (ed.) Dordrecht: Kluwer Academic Publishers, pp.37–54.

21. Spartz, J., Malaia, E., and Falk., C., 2005. Methodology and Tools for Ontological Semantic Acguisition, CERIAS. *In*: Raskin, V. and Spartz, J. (eds.). Proceedings of the 4th Midwest Computational Linguistics Colloquium 2007 [online]. Purdue University, West Lafayette, Indiana. April 28, 2007. Available at: http://cllt.osu.edu/mclc/paper/methodology_spartz.pdf. [Accessed April 2012].

22. Sroka, K. A., 1972. The Syntax of English Phrasal Verbs. The Hague: Mouton.

23. Televnaja, J., 2005. Representation of English Phrasal Verbs in Ontological Semantics. Kalbotyra, 55(3), Vilnius University Press.

Data Sources

1. (KBAE) Knowledge Base Acquisition Editor, 2004.

2. Longman Pocket Phrasal Verb Dictionary, 2001. Longman.

List of Abbreviations

anno	annotation
cat	category
def	definition
ex	example
KBAE	Knowledge Base Acquisition Editor
np	noun phrase
opt	optional
OS	Ontological Semantics
P	preposition and particle [used in the representation of structural classes of pvs]
prep	preposition and particle [used in templates]
pv	phrasal verb

sem	semantic	*syn-struc*	syntactic structure
sem-struc	semantic structure	*var*	variable
smb	somebody	*vp*	verb phrase
sth	something		

Julija Korostenskaja

Schemų sandaros bei parametrinių veikslo ir modalumo bruožų analizė aprašant frazeologinius veiksmažodžius ontologinėje semantikoje

Santrauka

Straipsnyje analizuojamas frazeologinių veiksmažodžių kompiuterizuotas aprašymas ontologinės semantikos teorijoje, panaudojant programą Knowledge Base Acquisition Editor (KBAE). Kadangi kiekvieno kalbos vieneto aprašymas KBAE programoje susideda iš dviejų dalių, sintaksinės ir semantinės, straipsnyje pateikiamos pagrindinės analizuojamų vienetų apdorojimo metodologinės problemos ir aptariami jų praktiniai sprendimai. Sintaksiniame lygmenyje parodomas vienos iš išankstinių frazeologinių veiksmažodžių aprašymo schemų klaidingumas bei, siekiant tinkamai aprašyti visas leistinas sintaksines kombinacijas,. siūlomas alternatyvus aprašymo būdas, panaudojant kitas iš anksto išvystytas schemas. Semantiniame lygmenyje pateikiami sprendimai, padedantys detaliai aprašyti frazeologinio veiksmažodžio reikšmę bei pabrėžti jos skirtumą nuo paprasto veiksmažodžio reikšmės, pasitelkiant parametrinius ontologinės semantikos parametrus, t. y. veikslo (aspect) bei modalumo (modality) kategorijas. Manoma, kad šios kategorijos sudaro ne tik charakteringą ir skiriamąjį frazeologinių veiksmažodžių bruožą, bet ir patikslina kai kurių jų atstovų reikšmės metaforiškumą. Tikimasi, kad tyrimo rezultatai gali būti panaudoti praktinio pobūdžio natūralios kalbos studijose, pasitelkiant tekstynų duomenis, siekiant patikslinti analizuojamų frazeologinių veiksmažodžių reikšmes, pragmatinį kontekstą, vartojimo dažnumą bei nustatyti frazeologinių veiksmažodžių metaforiškumo suvokimo bei loginio pagrindimo galimybes tarpkalbiniame lygmenyje.

About the author

Julija Korostenskaja, Associate Professor, Vilnius University, Institute of Foreign Languages, Vilnius University.
Research interests: semantics, phonology, teaching English as a foreign language, academic writing.
Address: Vilnius University, Institute of Foreign Languages, Universiteto Street 5, LT-01513 Vilnius.
E-mail: julija.korostenskaja@uki.vu.lt