Developing a knowledge base for preposition sense disambiguation: A view from Role and Reference Grammar and FunGramKB

Diana Hernández-Pastor
Universidad Católica San Antonio
España

Carlos Periñán-Pascual
Universidad Politécnica de Valencia
España

Diana Hernández-Pastor: Modern Languages Department, Universidad Católica San Antonio, España.
| Correo electrónico: dhernandez@ucam.edu
Carlos Periñán-Pascual: Applied Linguistics Department, Universidad Politécnica de Valencia, España.
| Correo electrónico: jopepas3@upv.es

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Abstract

Prepositions represent a grammatical category of frequent use in many European languages. The combination of their semantics with other lexical categories usually makes them difficult to be computationally tractable. As far as natural language processing is concerned, some studies have contributed to make progress on the usage of prepositions. However, there still exists a need to develop a model that allows tackling the problems which result from the disambiguation of prepositional semantics. The goal of this paper is to describe a lexico-conceptual model which can store the knowledge required to disambiguate predicate prepositions, as well as how this model can be exploited by a parser to extract the semantic representation of a text. The theoretical foundation of this approach, which is grounded on the premises of Role and Reference Grammar and FunGramKB, is illustrated with temporal adjuncts expressed by prepositional phrases in English.

Keywords: preposition; knowledge base; disambiguation; Role and Reference Grammar; FunGramKB.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

1. Introduction

Prepositions represent a grammatical category with frequent presence in Romance and Germanic languages. Despite of this fact, they are considered a lexical class notably difficult to be managed. Many works have approached prepositions from a theoretical perspective. For example, there have been many attempts to explain the way humans categorize spatial usages of prepositions from a cognitive-grammatical point of view in different languages (cf. Bennett, 1975; Brugman, 1981; Lakoff, 1987; Herskovits, 1986, 1997; Tyler and Evans, 2003; Svenonius, 2008; Lindstromberg, 2010). This type of studies, whose results have usually been displayed as semantic networks (e.g. Tyler and Evans, 2003) or lists (e.g. Herskovits, 1986), tend to propose an overview of the semantic constraints characteristic of this grammatical class, as it is the case of prepositions combined with noun phrases (NPs). On the other hand, not much research has been developed for prepositions in natural language processing (NLP). Only a few works such as The Preposition Project (Litkowski and Hargraves, 2005, 2007) or PrepNet (Saint-Dizier, 2005, 2008) have taken on providing a robust description of the prepositional semantics for NLP applications. However, as Saint-Dizier remarked, there is still a need to build a repository that stores a syntactic and semantic formalization that allows tackling the problems of preposition sense disambiguation (PSD). This task is one of evident significance even for those NLP applications that require a simple way of understanding, e.g. information extraction or machine translation. In this sense, the main problem to be faced is the disambiguation of a lexical category that seems to be extremely polysemous, where the semantic interpretation of the preposition closely depends on the context, apart from the difficulty of identifying cross-linguistic regularities. The purpose of this paper is twofold. On the one hand, we do not only aim to design an effective model to store the necessary information to disambiguate prepositions, but also to describe the methodology to populate the knowledge base with that information. To illustrate, we focus on the temporal senses conveyed by those predicate prepositions functioning as peripheral adjuncts. From the NLP perspective, this research, which is grounded in the theoretical model of Role and Reference Grammar (RRG) (Van Valin and Lapolla, 1997; Van Valin,

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

1992, 2005, 2008) and FunGramKB (Periñán-Pascual and Arcas-Túnez, 2004, 2005, 2007, 2010a, 2010b; Mairal-Usón and Periñán-Pascual, 2009; Periñán-Pascual and Mairal-Usón, 2009, 2010), contributes to the development of ARTEMIS (Automatically Representing TExt Meaning via an Interlingua-based System), a proof-of-concept computer application which is able to automatically provide a semantic representation of a text under the format of a conceptual logical structure (CLS) (Periñán-Pascual, 2013; Periñán-Pascual and Arcas-Túnez, 2014). The structure of this paper is organized as follows: section 2 presents the state of the art in the field of PSD in NLP, whereas the theoretical framework for this study is discussed in section 3; section 4 proposes a lexico-conceptual model for prepositions and section 5 describes the analysis of the temporal domain of some English prepositions according to this model; finally, section 6 presents some conclusions.

2. Preposition disambiguation in NLP

The problem of ambiguity has always been a troubling task in linguistics, cognitive science and NLP. Ambiguity has been traditionally analyzed into two categories, i.e. lexical ambiguity, where a given word can have several meanings, and structural ambiguity, where a given sentence can have several syntactic structures. Both types of ambiguities finally result in semantic ambiguity, since the utterance is semantically misinterpreted. In the case of prepositions, ambiguity has usually been illustrated as structural, as it has to do with their relational character with other lexical classes. For example, in (1), an NLP application would find trouble to solve the ambiguity triggered by the preposition *with*:

(1) a. Tom ate yoghurt <u>with a spoon</u>.
     b. Tom ate yoghurt <u>with cereals.</u>

The problem arises here with the function of the prepositional phrase (PP). The disambiguation task would be based on determining which governor of the PP is, whether the verb *ate* or the noun *yoghurt*. In (1a) the PP depends on the verb, as Tom makes use of the spoon to eat the yoghurt; whereas in (1b) the preposition *with* introduces a PP dependent on the noun, as it is the yoghurt which contains the cereals. PP attachment is the

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

area of syntax that tries to explain this sort of syntactic ambiguity, having as its main goal to determine which the governor of the preposition is. Most of research has focused on this type of structure, i.e. a noun which is preceded by a verb and followed by a PP. Nevertheless, the focus of our study is aimed at those predicate prepositions functioning as peripheral temporal adjuncts of the clause as in (2):

(2)   Most people come to my farm in the afternoon.

For the purpose of developing a knowledge base which stores both the syntactic and semantic information of prepositions from a lexico-ontological perspective, we need to thoroughly analyze how the problem should be faced.

This field of research has usually focused on the development of models which are aimed at human-processing strategies based on the analysis of syntactic trees independent from the lexical or discursive context (Baldwin et al., 2009). However, the need to formalize the knowledge about prepositions for NLP systems can become an arduous labour. For example, since prepositions are often considered as one of the most polysemous lexical categories, it tends to be difficult for their semantic knowledge to become formalized in a machine-tractable way. Few works (cf. Harabagiu, 1996; Dorr, 1997; O'Hara and Wiebe, 2003; Alam, 2004; Boonthum et al., 2005; Litkowski and Hargraves, 2005, 2007; Saint-Dizier, 2005, 2008; Tratz and Hovy, 2009) have taken the risk to approach the description of preposition semantics for their disambiguation. For example, Harabagiu (1996) used WordNet (Fellbaum, 1998) to work on the disambiguation of PPs, specifically those introduced by *of*, by applying an approach based on inferential heuristics. However, the semantic description in WordNet is only based on relational meaning, as it describes the usage of lexical units in the language through meaning relations with other lexical units[1]. Dorr (1997) designed The English LCS Lexicon, which provides an inventory of 165 prepositions with 497 senses where each of the preposition senses is

---

[1] In computational lexicography, there are two main approaches to formalize meaning (Velardi et al., 1991): the cognitive content of a lexical unit can be described by means of semantic features or primitives (i.e. conceptual meaning), or through associations with other lexical units in the lexicon (i.e. relational meaning).

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

described through a representation based on the model of lexical conceptual structures (LCS) by Jackendoff (1983, 1990, 1996). O'Hara and Wiebe (2003) formulated a semantic labeling classification for PPs in the databases Penn Treebank (Marcus et al., 1993) and FrameNet (Fillmore et al., 2001) using traditional word collocations. Alam (2004) analyzed the preposition *over* on the basis of two factors, i.e. meaning determined by the complement and meaning determined by its head. More particularly, she devised two decision trees, i.e. one for head and another for complement, upon ontological categories which allowed her to provide the semantic analysis of the preposition. Boonthum et al. (2005) worked on the disambiguation of *with* by determining the possible meaning of the preposition according to the LCS descriptors of the verb and the relationships between the categories designed for head and complement classified into one or more WordNet ontological categories. Saint-Dizier (2005) designed PrepNet, where a description of preposition semantics for NLP applications can be found. This classification is inspired on thematic role classifications with the aim of creating "with rather broad senses with several levels of stratification, and for most 'standard' usages […] a relatively fine-grained description of prepositions" (Saint-Dizier, 2008: 2). However, despite of the fact that he intended to devise a generic inventory, the result was a quite stratified inventory difficult to be computationally implemented. Despite all these studies, Hartrumpf et al. (2006) recalled that linguistic studies on preposition semantics have not obtained much impact on NLP applications. Litkowski and Hargraves (2005) presented The Preposition Project, which stores information of English preposition senses to be used in NLP with FrameNet[2]. This project attempts to analyze preposition behaviour through the comparison and contrast of the description made in a collegiate dictionary such as the Oxford Dictionary of English (2003) and the treatment received by this lexical category in FrameNet and the traditional English Grammar (Quirk et al., 1985). In a nutshell, their methodology seeks to identify and characterize English prepositions by providing an account of semantic roles. Likewise, Litkowski and Hargraves (2007) designed the SemEval-2007 task 6 to disambiguate prepositions and investigate the

---

[2] The database is publicly available at http://www.clres.com/prepositions.html for further research.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

linguistic constraints of this grammatical category[3]. In this task, a set of over 25,000 instances was developed, covering 34 of the most frequent English prepositions, with the aim of determining whether the general behaviour of individual prepositions is sustained with a set of occurrences extracted from FrameNet. Initial results in the SemEval-2007 task 6 was of great help to make progress on this domain, in such a way that researchers like Popescu et al. (2007), Ye and Baldwin (2007), Yuret (2007), Dahlmeier et al. (2009), or Tratz and Hovy (2009) published results making use of statistical techniques to identify reliable features associated to PSD. Moreover, as an attempt to improve the preposition analysis, three corpora were developed (Litkowski, 2013a): the SemEval Corpus, which was built from FrameNet, the Oxford English Corpus, which was derived from a set of examples for senses in the Oxford Dictionary of English, and the CPA corpus of prepositions, which was created from a set of sentences in the British National Corpus following the methodology of the Corpus Pattern Analysis project for verbs (Hanks, 2004). Nevertheless, Litkowski (2013b) highlighted the need to change the course of PSD research, motivated by the importance of prepositions in semantic role classification in combination with other lexical units, where the focus is now on the semantic relations rather than on the prepositions *per se*. In this line, Litkowski (2014) presented a new resource for the analysis of prepositions, i.e. The Pattern Dictionary of English Prepositions, in which 304 prepositions linked to 81,509 sentences were analyzed from the data stored in The Preposition Project, resulting in the construction of an inventory of the senses discovered through the aforementioned corpora. Indeed, this project serves to store information about the prototypical syntagmatic patterns with which prepositions are associated. In the same vein, Tratz and Hovy (2009) already framed a supervised classification approach for PSD, deriving linguistically-motivated features from both sides of the preposition. For example, the more in-depth research by Tratz and Hovy (2011) presented a parser containing four modules, one of them being a PSD system which achieved 85.7% accuracy on the SemEval-2007 task. This is actually a fast and accurate

---

[3] SemEval (Semantic Evaluation) takes the form of workshops organized by ACL-SIGLEX where different tasks are carried out to evaluate semantic analysis systems.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

parser that works on a dependency-tree conversion of the Penn Treebank and its associated fine-grained dependency labels.

Nonetheless, there is still a need to describe prepositions syntactically and semantically from a *deeper perspective* that enables their formalization to be applied to cross-linguistic NLP settings. In this line, our research also seeks to characterize prepositional meaning by providing an account of semantic roles, but the deeper perspective is achieved by grounding the model on the framework of both a linguistic theory (i.e. RRG) and a knowledge engineering project (i.e. FunGramKB).

3. Theoretical framework of the analysis

3.1. Role and Reference Grammar

RRG is one of the most relevant functional models of language in current linguistics. This linguistic theory advocates that the study of language must be done bearing in mind semantic and communicative functions. In this way, RRG attempts to develop a linguistic framework that allows the description of all languages. The main components of the RRG model are the lexicon, the syntactic inventory and the parser, as shown in figure 1.
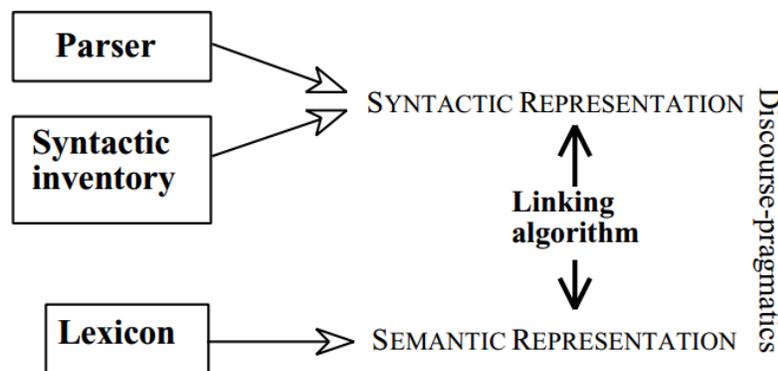

Figure 1. General architecture of RRG (Van Valin, 2005: 131)

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

The first two modules are connected by a linking algorithm which is bidirectional, that is, this algorithm links the syntactic to the semantic representation and the semantic to the syntactic representation in such a way that it accounts for aspects of the communicative situation in which the speaker and the receiver are involved.

In RRG, every lexical unit is formally represented by a logical structure, so it constitutes the nucleus of the semantic representations in the lexicon. Therefore, RRG requires a universal semantic metalanguage that allows the conceptual representation of predicates in the form of logical structures. Moreover, all features of the clause structure must be universal, in such a way that it should be possible to carry out the comparative analysis of structures of different languages. In fact, these two assumptions become the main pillars of the RRG model, which (i) is concerned with the integration of syntax, semantics and pragmatics in grammatical systems, and (ii) tries to achieve a high degree of typological adequacy, that is, the search of common aspects in the clause structure in all languages. It is in this search of typological adequacy where RRG rejects the standard formats of clause structure (e.g. grammatical relations) and puts forward a new clause structure in which universal elements are differentiated from those which are language-dependent. In this context, RRG proposed the Layered Structure of the Clause, whose semantically-motivated constituents are the nucleus, the core and the periphery.
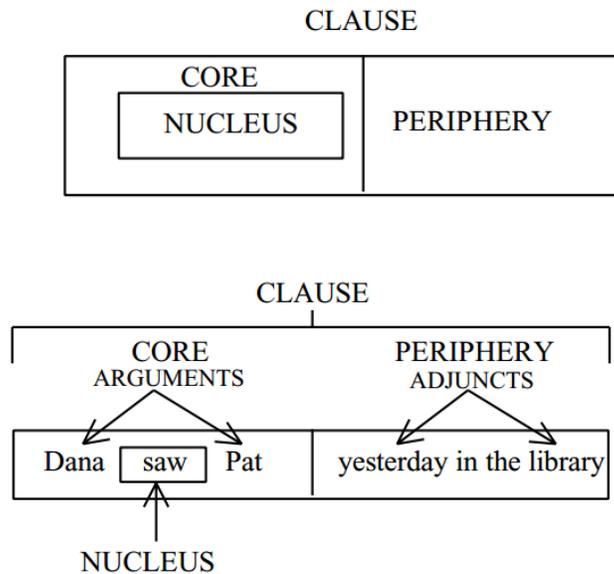
ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Figure 2. Components of the Layered Structure of the Clause (Van Valin, 2005: 4)

As shown in figure 2, the sentence *Dana saw Pat yesterday in the library* consists of a core, with the nucleus and two core arguments, and a periphery, with two peripheral adjuncts. According to Van Valin (2005), the elements contained in the core, the nucleus and the periphery may happen in any order providing that a given language allows it.

Unlike other linguistic theories, RRG views prepositions as elements activated by systematic principles underlying the semantics of the verb. To the best of our knowledge, little work has been done on prepositions in the RRG framework. An exceptional case is the outstanding work by Jolly (1991). She proposed that "prepositions did not have to be idiosyncratically specified with each verbal lexical entry but rather would be predicted by some general principles related to the semantics of the verb and the assignment of arguments to morphosyntactic positions within a clause" (Jolly, 1992: 275). Accordingly, she made a classification of prepositions as far as their functionality is concerned, laying out three different types of PPs within RRG:

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

(i)   PPs introduced by predicative prepositions functioning as peripheral modifiers of the clause (i.e. adjuncts): e.g. John walks every day in the park.

(ii)   PPs introduced by non-predicative prepositions (i.e. oblique core arguments): e.g. Peter gave the book to Mary.

(iii) PPs introduced by predicative prepositions marking verb arguments (i.e. argument-adjuncts in the core): e.g. Leslie put the book on the box.

The scope of our paper focuses only on those predicative PPs which work as adjuncts in the periphery of the RRG clause, that is, those PPs that do not construe with the verb and are not part of the core of the clause. In these adjuncts, the preposition functions as a predicator, being the nucleus of the phrase and licensing its own NP-argument, as shown in figure 3.
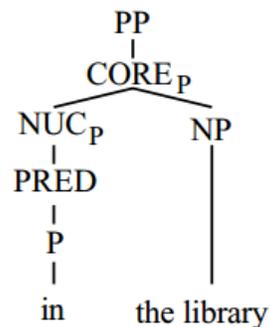

Figure 3. Structure of a predicative PP in English (Van Valin, 2005: 23)

According to Van Valin (2005: 21), these prepositions "contribute substantive semantic information to the clause in which they occur, both in terms of their own meaning and the meaning of the argument that they license". Furthermore, it is important to note that, according to Rauh (2010), adpositions which function as predicates contain a layered structure similar to that existing for the clause, so a core, a nucleus and a periphery can also be identified. All this can be summarized in the fact that some semantic description of

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

prepositions and their predicated elements may be useful, always after having dealt with the syntactic and pragmatic context in which they appear.

3.2. FunGramKB

FunGramKB is a lexico-conceptual knowledge base that implies a paradigmatic change in natural language understanding systems. The main purpose of FunGramKB is its application in NLP systems, providing modifications in the design of the linguistic model by leaving apart a lexicist approach and adopting an ontological one which entails redundancy reduction as well as information increase in the knowledge base. The lexical and grammatical levels of FunGramKB are defined within linguistic theories that allow the system to get the syntactic-semantic generalizations that account for explanations as well as predictions about linguistic phenomena. Particularly, RRG and the Lexical Constructional Model (LCM) (Mairal-Usón and Ruiz de Mendoza, 2006; Ruiz de Mendoza and Mairal-Usón, 2007, 2008) provide a suitable framework for the FunGramKB lexical and grammatical modules. Undoubtedly, the cornerstone of FunGramKB is the ontology, as here is where the conceptual description is made for every lexical unit introduced in the knowledge base.
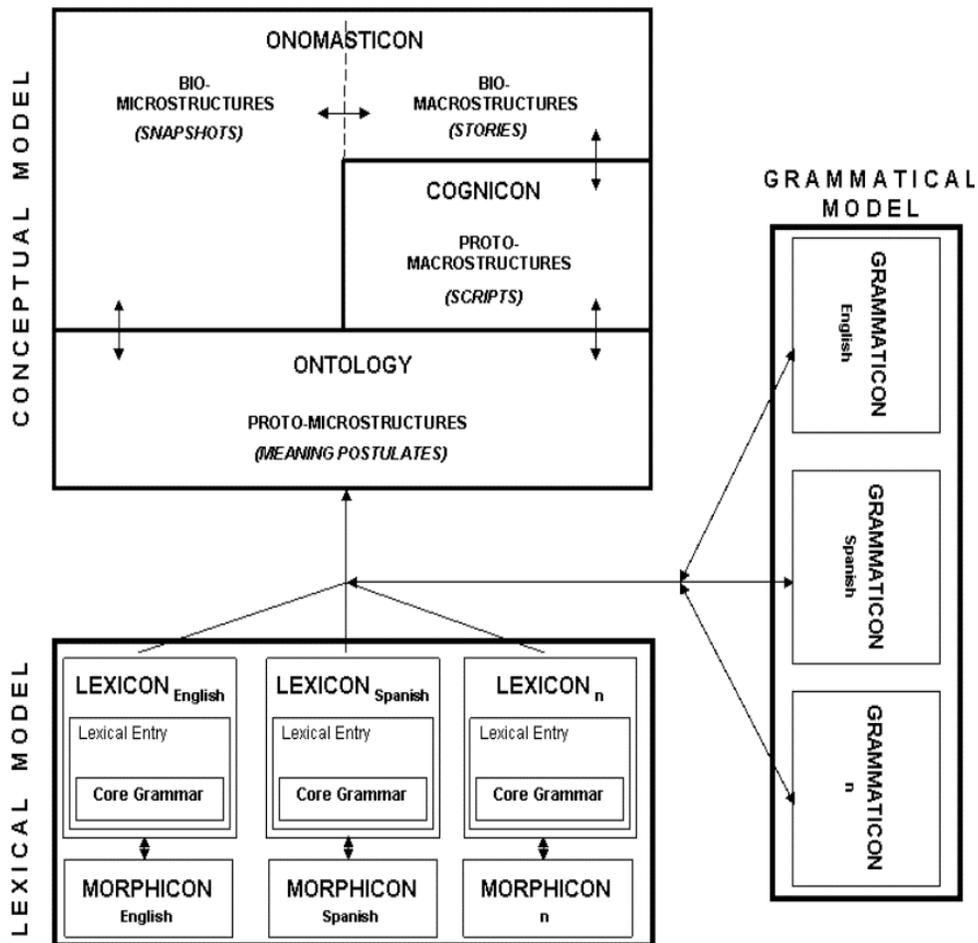
ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Figure 4. Architecture of FunGramKB (Periñán-Pascual and Arcas-Túnez, 2011: 3)

According to Periñán-Pascual (2012), and much in line with cognitive psychology theories, the three modules that make up the conceptual level correspond to the different types of knowledge that model human processing. By the same token, it is the interaction among these three conceptual modules (e.g. ontology, cognicon and onomasticon) which allows text comprehension in FunGramKB. In this context, semantic knowledge in FunGramKB is described from a deep perspective. In other words, semantic descriptions result from defining interwoven semantic features in the form of ontological units. More

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

particularly, for the construction of meaning representations, FunGramKB has developed COREL (Conceptual Representation Language—Periñán-Pascual and Mairal-Usón, 2010).

As stated previously, the cornerstone of the architecture of FunGramKB is the ontology, which consists of three distinct conceptual levels, being each one populated by different types of concepts (Periñán-Pascual and Arcas-Túnez, 2010a):

(i)   Metaconcepts, preceded by # (e.g. #TEMPORAL, #INSTRUMENT, etc.), shape the cognitive dimensions in the upper level of the ontology.

(ii) Basic concepts, preceded by + (e.g. +PERIOD_00, +OCCURRENCE_00, etc.) and located in the middle level of the ontology, constitute the defining units which take part in the construction of meaning postulates and as selectional preferences in thematic frames.

(iii) Terminal concepts, preceded by the symbol $ (e.g. $JANUARY_00, $HELIPORT_00, etc.), are distributed at the leaf level of the ontology.

Every ontological unit in FunGramKB is assigned two conceptual properties: thematic frame and meaning postulate. The thematic frame denotes the prototypical participants taking part in the cognitive situation described by a concept, whereas the meaning postulate defines the unsituated common-sense knowledge involved in that cognitive situation. To illustrate, the conceptual entry of +GIVE_00 is presented in (3):

(3) Concept: +GIVE_00
     Thematic   Frame: (x1:   +HUMAN_00   ^   +ANIMAL_00)Agent   (x2:
     +CORPUSCULAR_00)Theme (x1)Origin (x3: +HUMAN_00 ^+ANIMAL_00)Goal
     Meaning Postulate: +((e1: +TRANSFER_00 (x1)Agent (x2)Theme (x1)Origin (x3)Goal
     (f1:   +HAND_00)Instrument)(e2:   +BE_02   (x1)Theme   (x3)Location   (f2:   m
     +NEAR_00)Position))[4]
     Lexical units: give, hand (Eng); dar, entregar (Spa); cedere (Ita); donner (Fre).

---

[4] The English translation of the meaning postulate of +GIVE_00 is "somebody transfers something to somebody else with his/her hand, and both animated entities are very close to one another".

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Moreover, the relation between linguistic and conceptual knowledge is harmonized by CLSs, which serve as "the interface between the semantic structure and the syntactic representation of sentences" (Periñán-Pascual and Arcas-Túnez, 2010b: 3). As demonstrated by Periñán-Pascual and Arcas-Túnez (2014), this model of CLS is an enhanced version of the logical structure presented in RRG. For example, (4) presents the CLS of the verb *give*:

(4) $<[_{AKT}{}^{ACC} [+GIVE\_00 (x\text{-Agent, } y\text{-Theme, } z\text{-Goal})]]>$

It can be noted that the Aktionsart operator (i.e. AKT) together with an "argument pattern" headed by the event replaces the semantic skeleton originated by the RRG decompositional system[5]. Indeed, the core grammar in the lexicon contains those attributes whose values allow the system to build automatically the CLSs of lexical units. For the sake of clarity, the core grammar of *give* is shown in (5):

(5) Aktionsart: Accomplishment
     Lexical template: x-Agent, y-Theme, z-Goal

In the lexical template, every variable is assigned a thematic role from the thematic frame of the concept to which the verb is linked (i.e. +GIVE_00). These thematic roles actually play a key role in the CLS, since they allow the lexico-conceptual linkage by integrating the core grammar of the verb with the thematic frame of the corresponding event in the ontology, and, in turn, each participant of the thematic frame is co-indexed with a participant in the meaning postulate of the concept. In this way, the NLP system can reach a deeper level of comprehension by deriving the meaning postulate of the event (i.e. conceptual level) from the CLS of the verb (i.e. lexical level).

---

[5] In this argument pattern, the variables x, y and z will not be instantiated with predicates but with ontological concepts.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

4. Prepositions in FunGramKB

As stated in section 3.1, the goal of this paper is aimed at those predicative PPs which work as peripheral adjuncts. Indeed, the motivation of this line of research is grounded on the fact that the other two prepositional types were already explored in FunGramKB. On the one hand, oblique core arguments, such as those in (6), are dealt by the lexical model of FunGramKB.

(6) The meal consisted of vegetarian dishes.
   He relies on his family.

Oblique core arguments involve functional prepositions, which are characterized by their lack of semantic content. For example, the prepositions *of* and *on* are uniquely selected by the verbs *consist* and *rely* respectively in (6). As a result, the thematic role of the given PP is not assigned by the preposition, but by the verb. Consequently, the most suitable way of addressing this issue in FunGramKB is by means of the core grammar in the lexical entry of the verb, which contains those attributes whose values allow the system to build the CLS, e.g. the most representative Aktionsart in which the verb can occur, its prototypical arguments, or the prepositions involved in these arguments, among many other features.

On the other hand, argument-adjuncts in the core of the clause, such as those in (7), are dealt by the grammatical model of FunGramKB, namely the constructicon in the argumental layer.

(7)   The bottle floated into the cave.
   Bill broke the vase into pieces.

Both peripheral adjuncts and argument-adjuncts illustrate what has been coined as semi-lexical prepositions, since they provide descriptive semantic information. From the FunGramKB approach, however, these two prepositional types cannot be treated under the

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

same category. Whereas the thematic role of peripheral adjuncts is determined by the preposition, argument-adjuncts belong to subcategorization patterns whose thematic role grids are provided by the constructional schemata of the verb. For example, in the case of (7), the constructional schema assigned to the motion construction enables *float* to mark *into the cave* as Goal, and the constructional schema assigned to the resultative construction enables *break* to mark *into pieces* as Result. For this reason, as far as prepositions are concerned, the pending issue is how to model the knowledge required in FunGramKB to determine the thematic role of PP-adjuncts.

4.1. The lexical entry for prepositions

The prepositions of our study can be treated as predicates with arguments, so we first considered designing a core grammar similar to those for events. However, this hypothesis was soon rejected on the premise that prepositions lack their own Aktionsart, and consequently they cannot be used to generate their own CLS. Therefore, we started to design a new model of lexical entry for prepositions, where the system could store syntactic information related to every sense of the preposition.

The first step was to review the implications for prepositions of being predicative or not. In this regard, we considered what Baldwin et al. (2009) stated as the two most relevant syntactic features of prepositions for NLP applications, i.e. selection and valence. On the one hand, selection is the feature of prepositions that determines if a given preposition is selected by a governor or not. Therefore, oblique core arguments and argument-adjuncts in RRG involve selected prepositions, since the presence of the preposition is dependent on the lexical and constructional semantics of the verb respectively. By contrast, peripheral adjuncts involve unselected prepositions, which do not depend on the verb. On the other hand, valence refers to the transitivity of the preposition, i.e. the ability of the preposition to select an object (or internal argument) in order to construct PPs. Considering that our objective is to analyze those PPs that function as peripheral adjuncts, we can assume that we work on unselected prepositions with transitive valence. The implications of these two features for the knowledge base are further analysed in the remainder of this section.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

It is beyond discussion that the unselected preposition contributes semantically to the PP in which it appears, as shown in (8):

(8) Jane and Tom talked about the situation in the workshop.

In turn, this PP behaves as a relatively independent element that also contributes to the semantics of the overall clause by offering a location for the action denoted by the main verb. This perspective can lead to the misunderstanding that PP-adjuncts may be totally independent from the rest of the clause. However, in spite of their autonomy, adjuncts help to shape the aspectual meaning of the clause beyond the semantic level of the meaning of the preposition. In this regard, we agree with Cortés-Rodríguez (2014) that aspectual features should be distributed into the four RRG layers of the sentence, in such a way that the aspectual information provided in one layer could be extended to the next layer by means of the lexico-conceptual linkage. According to his proposal, the interpretation of aspectual features within the RRG framework would take place at the predicate (e.g. Aktionsart), nucleus (e.g. aspect operator), core (e.g. incremental complement) and clause (e.g. tense operator) layers of the sentence. One of the key issues of this proposal is the sharp distinction between L[exical]-(a)telicity, which refers to the aspectual feature of the Aktionsart, and S[yntactic]-(a)telicity, which refers to the aspectual feature of the core; whereas the former characterizes the event-type of a given predicate, the latter describes the specific event of a clause. For example, the event in (9) is interpreted as expressing unbounded time, since it is derived from a predicate whose Aktionsart is an activity in the lexical entry (i.e. L-atelic).

(9) Larry painted.

However, a quantified complement with the same verb, as in (10), marks the culmination of the action, resulting in an active accomplishment which licenses a telic temporal PP-adjunct (i.e. S-telic).

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

(10)   Larry painted three portraits in two hours.

In these cases, FunGramKB should be halfway between the lexical approach of RRG and Cortés-Rodríguez's constructional approach. On the one hand, the lexical entry of the preposition should include those aspectual markers with which the preposition could be in consonance. These markers would be part of feature-based structures in the form of attribute-value matrices (AVMs) within lexical rules, which are automatically derived from the Lexicon and the Ontology. On the other hand, as these aspectual markers could be merged through non-monotonic unification operations with other AVMs from syntactic and constructional rules in the grammar of the language, we claim that telicity shifts could be considered a syntactic phenomenon in ARTEMIS.

Finally, concerning the second feature of predicative prepositions, we also intended to include information about their internal argument in the lexical entry. As the semantic head of the PP, the preposition is involved in co-occurrence restrictions with the arguments of the verb, so the semantic content of the internal argument of the preposition is constrained. As exemplified by Rauh (1993: 108), the preposition *at* allows a spatially interpretable complement in (11a), but the preposition *above* cannot take a complement marked as temporal in (11b).

(11)   a. Bill arrived at the station.
        b. *Bill stayed above an hour.

However, we soon realized that an ontological approach to selectional preferences is more efficient to minimize redundancy in a multilingual setting.

Consequently, we decided a single criterion for the design of the lexical model of prepositions:

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

- CRITERION #1: The lexical entry should state the way in which the preposition interacts with aspectual features at the predicate, nucleus, core and clause layers of the sentence.

4.2. Towards an ontology of relations

The FunGramKB ontology is currently organized into three subontologies: #ENTITY for nouns, #EVENT for verbs and #QUALITY for adjectives. The goal of this section is to describe the methodological underpinnings for the construction of the subontology of #RELATION, which is linguistically instantiated by prepositions.

First, as stated by Hindle and Rooth (1993), neglecting the information regarding the object of the preposition could prevent the system from achieving the expected results. For example, the two sentences in (12) show that not only the Aktionsart of the clause and the argument structure may help disambiguate a given PP-adjunct but also the object of the preposition can help to trigger the appropriate sense of the preposition, e.g. temporal in (12a) and locative in (12b).

(12) a. Britain reopened its embassy in December.
       b. Britain reopened its embassy in Teheran.

In these examples, the sense of the preposition can only be determined by the selectional preference of the relation, being this information obtained from the FunGramKB conceptual modules. Thus, the conceptual referent of *December* in (12a) becomes $DECEMBER_00 in the ontology, and that of *Teheran* in (12b) becomes %TEHERAN_00 in the onomasticon. As seen in the conceptual paths of (13a) and (13b), the former is categorized as Time, whereas the latter is categorized as Place.

(13) a. +TIME_00 >> +PERIOD_00 >> +MONTH_00 >> $DECEMBER_00
       b. +PLACE_00 >> +ARTIFICIAL_AREA_00 >> +TOWN_00 >> +CITY_00 >> $CAPITAL_00 >> %TEHERAN_00

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Second, we decided that the conceptual entry of the relation should be devoid of any meaning postulate. This decision is motivated by the role of prepositions in the CLS and the COREL scheme. In ARTEMIS, the CLS serves as the pivotal representation between the input text and the COREL scheme, which in turn serves as the pivotal representation between the CLS and the automated reasoner. Periñán-Pascual (2013) explained this lexico-conceptual linkage with the sentence "The juice froze black in the refrigerator", whose CLS and COREL scheme are shown in (14) and (15) respectively:

(14) $<_{IF}$ DECL $<_{Tense}$ past $<_{CONSTR-L1}$ RESI $<_{CONSTR-L1}$ INCH $<_{AKT}$ ACC [+FREEZE_00 (+JUICE_00-Referent,   +BLACK_00-Result)]   (+REFRIGERATOR_00-Location) >>>>>

(15)   +(e1: past +FREEZE_00 (x1)Theme (x2: +JUICE_00)Referent (f1: (e2: +BECOME_00 (x2)Theme (x3: +BLACK_00)Attribute))Result (f2: +REFRIGERATOR_00)Location)

When some kind of reasoning with the input is required, the CLS should be transduced into a COREL scheme, so that it can be enriched by the conceptual knowledge from the FunGramKB cognitive modules. For example, the COREL scheme (15) can be extended to (16) through the meaning postulate of +FREEZE_00.

(16)   +(e1: past +COOL_00 (x1)Theme (x2: +JUICE_00)Referent (f1: +MUCH_00)Quantity (f2: (e2: +BECOME_00 (x2)Theme (x3: +SOLID_00 & +BLACK)Attribute))Result (f3: +REFRIGERATOR_00)Location)

But what about prepositions in this mapping of the input text into the CLS, and in turn into the COREL scheme? As can be seen in (14), prepositions do not become formalized as conceptual units, just like noun, verbs and adjectives do, but as thematic roles. Therefore, the meaning postulate of a relation would play no role at all in the

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

FunGramKB ontology, since relations do not contribute to the extension of COREL schemata after the lexico-conceptual linkage.

Third, we intended to devise a three-layered ontological model for relations, including metaconcepts, basic concepts and terminals. However, since there are no meaning postulates in this subontology, there is no point in making a difference between basic and terminal concepts[6], so the logical procedure was for the latter to disappear.

Four, the only taxonomic relation permitted in the FunGramKB ontology is subsumption, i.e. the IS-A relation (Periñán-Pascual and Arcas-Túnez, 2010b). However, and in the absence of further evidence, we thought that at least for now there should be no ontological depth below the metaconceptual layer. This idea is supported by the fact that one of the mechanisms to extend the meaning postulate of a concept is by means of inheritance, which "involves the transfer of one or several predications from a superordinate concept to a subordinate one in the ontology" (Periñán-Pascual and Arcas-Túnez, 2005: 241). As there are no meaning postulates in the subontology of relations, then there is no point to relate conceptual units through subsumption, since no predication can be inherited. It is important to note, however, that metaconcepts should not be deemed as conceptual units but as universal cognitive dimensions which facilitate the integration and exchange of information with other ontologies. Therefore, hierarchization is justified only in the case of metaconcepts.

Consequently, we decided the following criteria for the development of the subontology of relations:

- CRITERION #2: the conceptual entry of the relation should contain a thematic frame including the selectional preference(s) of its internal argument.
- CRITERION #3: the subontology of relations consists of two types of concepts, i.e. metaconcepts and basic concepts.

---

[6] Recall that the borderline between basic and terminal concepts is based on the presence or lack of definitory potential to take part in meaning postulates.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

- CRITERION #4: subsumption-based hierarchization only takes place in the metaconceptual level.

5. Analyzing the temporal domain of English prepositions

To illustrate the methodological underpinnings of the lexico-conceptual model devised in section 4, we are going to consider the temporal senses of prepositions *at*, *by*, *for*, *in*, *on*, *since*, *until* and *within*.

FunGramKB adopts a conceptualist approach to language processing, where the ontology is the component which feeds the lexicon. In other words, the lexical entry of a given word cannot be populated unless that word has been assigned to a conceptual unit in the ontology. Therefore, we moved onto the conceptual modelling stage, going from the macrostructure (i.e. the ontological organization of the domain) to the microstructure (i.e. the entry of basic concepts).

First of all, the subontology should be provided with a metaconceptual level, so that basic concepts can be placed in the appropriate cognitive dimension (cf. CRITERION #3). For this purpose, we created #RELATION as the root metaconcept of this subontology of relations. Then, since relations always take the form of thematic roles in the CLS of a sentence, the most logical approach was to take the inventory of thematic roles for FunGramKB satellites as metaconceptual categories subsumed by #RELATION (cf. CRITERION #4). Table 1 shows the metaconcepts which resulted from these thematic roles.

#BENEFICIARY	#GOAL	#PURPOSE
#COMPANY	#INSTRUMENT	#QUANTITY
#COMPARISON	#LOCATION	#REASON
#CONDITION	#MANNER	#SCENE
#DURATION	#MEANS	#SPEED
#FREQUENCY	#POSITION	#TIME

Table 1. Metaconceptual dimensions in the #RELATION subontology

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Once we focused on the #TIME dimension, we determined to create deeper metaconceptual levels, so that finer temporal dimensions would reflect the way that "the senses of prepositions tend to be closely related to one another" (O'Hara and Wiebe, 2009: 151). To avoid having to reinvent the wheel, we reviewed scientific literature to look for a taxonomy of temporal prepositions to be computerized for natural language understanding and to provide an adequate account of the RRG aspectual distinctions. Our ontological modelling of the #TIME dimension was finally grounded on Brée's research (Brée, 1992; Brée and Pratt-Hartmann, 2002), whose framework of analysis "is likely to be reflected in the use of temporals in all languages, or at least all European languages" (Brée, 1992: 347).

To understand the semantics of temporal PP adjuncts, three notions of time should be considered[7]:

(i)   the time of occurrence of the event of the predication (TOE),

(ii)   the time from which TOE is computed, i.e. the time of reference (TOR), and

(iii)   the length of time introduced by the argument of the temporal preposition (t).

Given these temporal parameters, we could model the whole #TIME dimension, including metaconcepts and basic concepts, as shown in figure 5. The description of each metaconcept is provided below.

---

[7] The notions (i-ii) are inspired on Reichenbach's (1947) *point of the event* and *point of reference*. For example, in a statement such as *When John arrived, Mary had gone out*, the point of the event is the moment when Mary went out and the point of reference is the moment when John arrived.
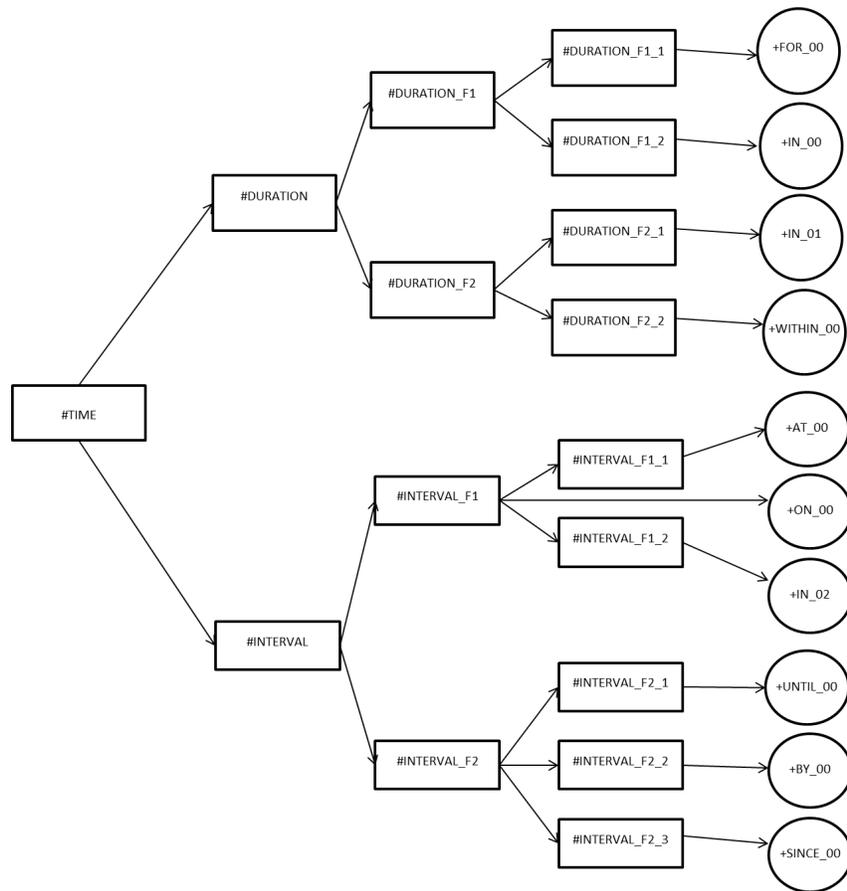
ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

Figure 5. The #TIME dimension in the #RELATION subontology

- #DURATION: *t* is a time without an actual location on the time axis.
  - #DURATION_F1: the TOE is the same as *t*, i.e. TOE = *t*.
    - #DURATION_F1_1: the predication does not imply a terminal point of the event.
    - #DURATION_F1_2: the predication implies a terminal point of the event.
  - #DURATION_F2: the TOE is located with respect to the TOR.
    - #DURATION_F2_1: the TOE is located at a specific time later than the TOR, where TOE = TOR + *t*.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

- ▪ #DURATION_F2_2: the TOE is located in a range of time with the TOR at one end and t at the other, so that TOE < TOR + t.
- #INTERVAL: t is a time with an actual location on the time axis.
  - ○ #INTERVAL_F1: the TOE is the same as t, i.e. TOE = t.
    - ▪ #INTERVAL_F1_1: t is a time point.
    - ▪ #INTERVAL_F1_2: t is a period.
  - ○ #INTERVAL_F2: the TOE is located in a range of time with the TOR usually at one end and t at the other.
    - ▪ #INTERVAL_F2_1: t begins at the TOR by default, and the TOE must last the whole t[8].
    - ▪ #INTERVAL_F2_2: t begins at the TOR, and the TOE can last the whole t or be located at some point of t.
    - ▪ #INTERVAL_F2_3: t always ends at the TOR.

The next step involved the lexicalization of each relation, as shown in table 2.

Relations	Lexical units	Example
+FOR_00	for	We sat in the coach for three hours.
+IN_00	in	He walked two miles in eight days.
+IN_01	in	I'll have some breakfast ready in a few minutes.
+WITHIN_00	within	The train will leave within an hour.
+AT_00	at	The funeral will be carried out this afternoon at 3.00.
+ON_00	on	John arrived on Sunday morning.
+IN_02	in, during	My neighbors were moving

---

[8] This default may be overridden using *from*, as in *Peter will be in Paris from the end of March until June*.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

		furniture in the evening. Sandstorms are common during the winter.
+UNTIL_00	until, till	They serve food until 7pm.
+BY_00	by	The work must be finished by May 1.
+SINCE_00	since	Arnold has been a member of parliament since 1987.

Table 2. Lexical units and examples of temporal relations

After the ontological organization of this domain, we developed the conceptual entry for each one of the basic concepts (cf. CRITERION #2). Although this type of conceptual entry only requires information about the thematic frame of the relation, the traversing path from the root metaconcept to a given basic concept can be deemed to provide the meaning of the target concept, as each metaconcept in this subontology describes a semantic feature. In (17), we can see some of the selectional preferences of the concept +IN_02 by way of example.

(17) (f1: +AFTERNOON_00 ^ +CENTURY_00 ^ +EVENING_00 ^ +MONTH_00 ^ +MORNING_00 ^ +SEASON_00 ^ +YEAR_00…)Time

Finally, the ontological modelling was followed by the lexical modelling, where we developed the lexical entry of each of the prepositions linked to any of the above relations (cf. CRITERION #1). To illustrate this stage, let us consider the preposition *in* in the following examples:

(18) John built two houses in 3 months.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

(19)   John ran to the park in an hour.

(20)   The train arrives in 10 minutes.

(21)   We shall run in 5 minutes.

In the above examples, the argument of the temporal adjunct is perceived as #DURATION. On the one hand, the preposition in the sentences (18) and (19) should be conceptualized as +IN_00, i.e. the TOE is the same as *t* and the predication implies a terminal point of the event. As explained in section 4.1, the aspectuality of specific events results from the interaction of aspectual features from the three layers, in such a way that the features in upper layers override those in lower layers. In the case of (18), two features which provide aspectual information are introduced. First, the predicate's Aktionsart describes an atelic event-type. Second, the aspectuality of the core describes a telic event, since the quantified complement brings an end-boundary to the specific event. In the case of (19), the predicate *run* corresponds to an L-atelic activity, but the temporal adjunct encodes a bounded path with the complement, resulting in a S-telic active accomplishment. Both examples demonstrate that the aspectuality of the event in the core overrides the atelicity of the Aktionsart assigned from the lexical entry. Therefore, the lexical entry of the preposition *in* with the sense of +IN_00 should include the following template of constraints:

(22)   [[HEAD: in]
        [RELATION: +IN_00]
        [CONSTRAINTS:
            [CLAUSE:
                [CORE:
                    [NUC:
                        [PRED:
                            [AKT: ACT]]]

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

[ARGS:

[[TYPE: y][NUM: NOT null]] ||

[[TYPE: y][QNT: NOT null]] ||

[[TYPE: y][DEF: NOT null]] ||

[[TYPE: y][PREP: to]]]]]]]]

[AKT-SHIFT: ACA]]

On the other hand, the preposition in the sentences (20) and (21) should be conceptualized as +IN_01, i.e. the TOE is located at a specific time later than the TOR, where TOE = TOR + t. As stated by Filip (1999: 23), "with achievements and activities the time indicated by the time-span adverbial is calculated from a contextually given reference point", and +IN_01 is a relation whose TOE is located with respect to the TOR. Moreover, this relation cannot be combined with perfective aspect or past tense operators, since the preposition focuses on the beginning of the event. Therefore, the lexical entry of the preposition *in* with the sense of +IN_01 should include the following template of constraints:

(23)   [[HEAD: in]

[RELATION: +IN_01]

[CONSTRAINTS:

[CLAUSE: [TNS: pres || fut]

[CORE:

[NUC:

[ASP: NOT perf]

[PRED:

[AKT: ACT || ACH]]]]]]]]

At first sight, we might think that, instead of placing these constraints in the lexical entries of the preposition, they could have been assigned to the ontological entry of the relation. In other words, should CRITERION #1 have been treated as other ontological

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

guideline? As shown in figure 4, the information about the linguistic behaviour of lexical units (i.e. lexicon) and the semantic knowledge of conceptual units (i.e. ontology) are clearly distributed in different modules of the FunGramKB architecture. In this framework, preposition constraints take the form of lexico-syntactic features where some of their values can be language-dependent, so they must be described in lexical entries. Evidence to support this assertion can be found in the example (24), described by Slabakova (2000: 748):

(24)   Tja z-gotvi jadene za 3 cˇasa.
         'She cooked some food in three hours'.

In this Bulgarian sentence, although the complement of the verb is a mass noun, a lexically selected morpheme on the verb signals telicity, in such a way that "mass nouns or bare plurals in form take on an interpretation of specified quantity or number" (Slabakova, 2000: 749). Another example which demonstrates that Aktionsart markers are language-dependent is illustrated in (25), where the optional reflexive clitic *se* occurs with a specified cardinality object to reinforce the telicity of the clause:

(25)   Juan se comió una hamburguesa en 10 minutos.
         'John ate a hamburger in ten minutes'.

Therefore, one of the senses of the prepositions *in* in English, *za* in Bulgarian and *en* in Spanish is conceptualized as +IN_00 in the ontology, but the template of constraints for each preposition cannot contain the same aspectual markers, so it makes more sense to include them in their corresponding lexical entries.

6. Conclusions

This paper presented the theoretical underpinnings for the development of a subontology of relations together with a lexical model of prepositions within the framework

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

of RRG and FunGramKB. In particular, we described not only the principles which structure the ontological organization of relations but also the model of lexico-syntactic constraints which can determine the linguistic contextual cues interacting with the semantics of prepositions. In this case, special emphasis was given to aspectual features in the predicate, nucleus, core and clause levels of the sentence. Since oblique core arguments and argument-adjuncts have already been explored in FunGramKB, this research focused on the predicative PPs which work as adjuncts in the periphery of the RRG clause. The outcome of this work is intended to be used with ARTEMIS, a parser which provides the semantic representation of a text under the format of a CLS.

The theoretical foundation of this lexico-conceptual model was exemplified with the temporal domain of English prepositions. However, we aim to expand this analysis to other semantic domains (e.g. location) and to prepositions in other languages in future research. In this latter case, we do not rule out the possibility that our subontology for the temporal domain could be reviewed in order to accommodate it to other languages, since the structuring of ontologies is usually guided by a process of negotiation.

7. Acknowledgement

8. References

ALAM, Yukiko Sasaki, 2004: "Decision Trees for Sense Disambiguation of Prepositions: Case of Over" in *Proceedings of the HLT-NAACL, Computational Lexical Semantic Workshop Conference*, Boston: MA, 52-59.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

BALDWIN, Timothy, Valia KORDONI and Aline VILLACENCIO, 2009: "Prepositions in Applications: A Survey and Introduction to the Special Issue", *Computational Linguistics* 35(2), 119-147.

BENNETT, David C., 1975: *Spatial and Temporal uses of English Prepositions: An Essay in Stratificational Semantics*, London: Longman.

BOONTHUM, Chutima, Shunichi TOIDA and Irwin LEVINSTEIN, 2005: "Sense disambiguation for preposition 'with'" in *Proceedings of the Second ACL-SIGSEM Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, 153-162.

BRÉE, David S., 1992: "Words for time" in Françoise MACAR, Viviane POUTHAS and William J. FRIEDMAN (eds.): *Time, Action and Cognition: Towards Bridging the Gap*, Dordrecht: Kluwer, 337-348.

BRÉE, David S. and Ian E. PRATT-HARTMANN, 2002: "Temporal semantics of prepositions in context" in Susanne FEIGENBAUM and Dennis KURZON (eds.): *Prepositions in their Syntactic, Semantic and Pragmatic Context*, Amsterdam/Philadelphia: John Benjamins, 75-115.

BRUGMAN, Claudia M., 1981: *Story of Over*. MA. Thesis, University of Berkeley, California.

CORTÉS-RODRÍGUEZ, Francisco J., 2014: "Aspectual features in Role and Reference Grammar: a layered proposal", *Revista Española de Lingüística Aplicada* 27(1), 23-53.

DAHLMEIER, Daniel, Hwee TOU NG and Tanja SCHULTZ, 2009: "Joint learning of preposition senses and semantic roles of prepositional phrases" in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1*, 450-458.

DORR, Bonnie J., 1997: "Large-scale dictionary construction for foreign language tutoring and interlingual machine translation", *Machine Translation* 12(4), 271-322.

FELLBAUM, Christiane, 1998: *WordNet: An Electronic Lexical Database*, Cambridge, MA: MIT Press.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

FILIP, Hana, 1999: *Aspect, Eventuality Types and Nominal Reference*, New York: Routledge, Taylor & Francis Group.

FILLMORE, Charles J., Charles WOOTERS and Collin F. BAKER, 2001: "Building a large lexical databank which provides deep semantics" in *Proceedings of the Pacific Asian Conference on Language, Information and Computation*, Hong Kong, 3-25.

HANKS, Patrick, 2004: "Corpus Pattern Analysis" in *Proceedings of the 11th EURALEX International Conference*, Lorient, 87-98.

HARABAGIU, Sanda M., 1996: "An Application of WordNet to Prepositional Attachment" in *Proceedings of the 34th annual meeting on Association for Computational Linguistics*, Association for Computational Linguistics, 360-362.

HARTRUMPF, Sven, Hermann HELBIG and Rainer OSSWALD, 2006: "Semantic interpretation of prepositions for NLP applications" in *Proceedings of the Third ACL-SIGSEM Workshop on Prepositions*, Trento, 29-36.

HERSKOVITS, Annette, 1986: *Language and Spatial Cognition. An Interdisciplinary Study of the Prepositions in English*, Cambridge: Cambridge University Press.

HERSKOVITS, Annette, 1997: "Language, Spatial Cognition, and Vision" in Oliviero STOCK (ed.): *Temporal and Spatial Reasoning*, Dordrecht: Kluwer Academic Press, 155-202.

HINDLE, Donald and Mats ROOTH, 1993: "Structural ambiguity and lexical relations", *Computational linguistics* 19(1), 103-120.

JACKENDOFF, Ray, 1983: *Semantics and Cognition*, Cambridge: MIT Press.

JACKENDOFF, Ray, 1990: *Semantic Structures*, Cambridge: MIT Press.

JACKENDOFF, Ray, 1996: "Conceptual Semantics and Cognitive Semantics", *Cognitive Linguistics* 7, 93-129.

JOLLY, Julia, 1991: *Prepositional analysis within the framework of role and reference grammar*, Berlin: Peter Lang.

JOLLY, Julia, 1992: "Preposition Assignment in English" in J. R. VAN VALIN (ed.): *Advances in Role and Reference Grammar*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 275-310.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

LAKOFF, George, 1987: *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*, Chicago: The University of Chicago Press.

LINDSTROMBERG, Seth, 2010: *English Prepositions Explained*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

LITKOWSKI, Ken C., 2013a: "The Preposition Project Corpora", *Technical Report* 13-01, Damascus, MD: CL Research.

LITKOWSKI, Ken C., 2013b: "Preposition Disambiguation: Still a Problem", *Technical Report* 13-02, Damascus, MD: CL Research.

LITKOWSKI, Ken C., 2014: "Pattern Dictionary of English Prepositions" In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, Maryland, USA, 1274-1283.

LITKOWSKI, Ken C. and Orin HARGRAVES, 2005: "The Preposition Project" in *Proceedings of the Second ACL-SIGSEM Workshop on The Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications*, Colchester, 171-179.

LITKOWSKI, Ken C. and Orin HARGRAVES, 2007: "The SemEval-2007 Task 06: Word Sense Disambiguation of Prepositions" in *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, 24-29.

MAIRAL-USÓN, Ricardo and Carlos PERIÑÁN-PASCUAL, 2009: "The anatomy of the lexicon within the framework of an NLP knowledge base", *Revista Española de Lingüística Aplicada* 22, 217-244.

MAIRAL-USÓN, Ricardo and José Francisco RUIZ DE MENDOZA, 2006: "Internal and external constraints in meaning construction: The lexicon grammar continuum" in Laura ALBA JUEZ and M. Teresa GUBERT (eds.): *Estudios de Filología Inglesa: Homenaje a la Dra. Asunción Alba Pelayo*, Madrid: UNED, 219-237.

MARCUS, Mitchell, Mary Ann MARCINKIEWICZ and Beatrice SANTORINI, 1993: "Building a large annotated corpus of English: The Penn Treebank", *Computational Linguistics* 19 (2), 313-330.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

O'HARA, Tom and Janyce WIEBE, 2003: "Preposition semantic classification via Penn Treebank and FrameNet" in *Proceedings of the Seventh Conference on Natural Language Learning -03*, Edmonton, 79-86.

O'HARA, Tom and Janyce WIEBE, 2009: "Exploiting Semantic Role Resources for Preposition Disambiguation", *Computational Linguistics* 35 (2), 151-184.

PERIÑÁN-PASCUAL, Carlos, 2012: "The situated common-sense knowledge in FunGramKB", *Review of Cognitive Linguistics* 10 (1), 184-214.

PERIÑÁN-PASCUAL, Carlos, 2013: "Towards a model of constructional meaning for natural language understanding" in Brian NOLAN and Elke DIEDRICHSEN (eds.): *Linking Constructions into Functional Linguistics: The Role of Constructions in a Functional Grammar*, Amsterdam and Philadelphia: John Benjamins Publishing Company.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2004: "Meaning postulates in a lexico-conceptual knowledge base" in *15th International Workshop on Databases and Expert Systems Applications*, IEEE, Los Alamitos (California), 38-42.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2005: "Microconceptual-Knowledge Spreading in FunGramKB" in *Proceedings of the 9th IASTED International Conference on Artificial Intelligence and Soft Computing* Anaheim-Calgary-Zurich: ACTA Press, 239-244.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2007: "Cognitive modules of an NLP knowledge base for language understanding", *Procesamiento del Lenguaje Natural* 39, 197-204.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2010a: "The Architecture of FunGramKB" in *Proceedings of the 7th International Conference on Language Resources and Evaluation*, Malta, 2667-2674.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2010b: "Ontological commitments in FunGramKB", *Procesamiento del Lenguaje Natural* 44, 27-34.

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2011: "Introduction to FunGramKB", *Anglogermánica Online* 8, 1-15.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

PERIÑÁN-PASCUAL, Carlos and Francisco ARCAS-TÚNEZ, 2014: "The implementation of the FunGramKB CLS Constructor in ARTEMIS" in Carlos PERIÑÁN-PASCUAL and Brian NOLAN (eds): *Language Processing and Grammars: The Role of Functionally Oriented Computational Models*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

PERIÑÁN-PASCUAL, Carlos and Ricardo MAIRAL-USÓN, 2009: "Bringing Role and Reference Grammar to natural language understanding", *Procesamiento del Lenguaje Natural* 43, 265-273.

PERIÑÁN-PASCUAL, Carlos and Ricardo MAIRAL-USÓN, 2010: "La gramática de COREL: un lenguaje de representación conceptual", *Onomazein* 21, 11-45.

POPESCU, Octavian, Sara TONELLI and Emanuele PIANTA, 2007: "IRST-BP: Preposition Disambiguation based on Chain Clarifying Relationships Contexts" in *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, 191-194.

QUIRK, Randolph, Sidney GREENBAUM, Geoffrey LEECH and Jan SVARTVIK, 1985: *A Comprehensive Grammar of the English Language*, London: Longman.

RAUH, Gisa, 1993: "On the grammar of lexical and non-lexical prepositions in English" in Cornelia ZELINSKY-WIBBELT (ed): *The Semantics of Prepositions: From Mental Processing to Natural Language Processing*, Berlin: Mouton de Gruyter, 99-150.

RAUH, Gisa, 2010: *Syntactic categories: Their identification and description in linguistic theories*, Oxford: Oxford University Press.

REICHENBACH, Hans, 1947: *Elements of Symbolic Logic*, New York: Macmillan & Co.

RUIZ DE MENDOZA, Francisco José and Ricardo MAIRAL-USÓN, 2007: "Levels of semantic representation: Where lexicon and grammar meet", *Interlingüistica* 17, 26-47.

RUIZ DE MENDOZA, Francisco José and Ricardo MAIRAL-USÓN, 2008: "Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model", *Folia Linguistica* 42 (2), 355-400.

SAINT-DIZIER, Patrick, 2005: "PrepNet: A framework for describing prepositions: Preliminary investigation results" in *Proceedings of IWCS 6*, Tilburg, 145-157.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

SAINT-DIZIER, Patrick, 2008: "Syntactic and Semantic Frames in PrepNet" in *Third International Joint Conference on Natural Language Processing*, Hyderabad, 763-768.

SLABAKOVA, Roumyana, 2000: "L1 transfer revisited: the L2 acquisition of telicity marking in English by Spanish and Bulgarian native speakers", *Linguistics* 38, 739-70.

SVENONIUS, Peter, 2008: "Projections of P" in Anna ASBURY, Jakub DOTLACIL, Berit GEHRKE, and Rick NOUWEN (eds.): *Syntax and Semantics of Spatial P*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 63-84.

TRATZ, Stephen and Eduard HOVY, 2009: "Disambiguation of preposition sense using linguistically motivated features" in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume: Student Research Workshop and Doctoral Consortium*, Boulder, 96-100.

TRATZ, Stephen and Eduard HOVY, 2011: "A Fast, Accurate, Non-Projective, Semantically-Enriched Parser" in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Edinburgh, 1257-1268.

TYLER, Andrea and Vyvyan EVANS, 2003: *The Semantics of English Prepositions. Spatial Scenes, Embodied Meaning and Cognition*, Cambridge: Cambridge University Press.

VAN VALIN, Robert D. Jr., 1992: *Advances in Role and Reference Grammar*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

VAN VALIN, Robert D. Jr., 2005: *Exploring the Syntax-Semantics Interface*, Cambridge: Cambridge University Press.

VAN VALIN, Robert D. Jr., 2008: *Investigations of the Syntax-Semantics-Pragmatics Surface*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

VAN VALIN, Robert D. Jr. and Randy J. LAPOLLA, 1997: *Syntax, Structure, Meaning and Function*, Cambridge: Cambridge University Press.

VELARDI, Paola, Michela FASOLO and Maria Teresa PAZIENZA, 1991: "How to Encode Semantic Knowledge: A Method for Meaning Representation and Computer-aided Acquisition", *Computational Linguistics* 17-2, 153-170.

ONOMÁZEIN 33 (junio de 2016): 251-288
Diana Hernández-Pastor y Carlos Periñán-Pascual:
Developing a knowledge base for preposition sense disambiguation: A view from Role…

YE, Patrick and Timothy BALDWIN, 2007: "MELB-YB: Preposition Sense Disambiguation Using Rich Semantic Features" in *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, 241-244.

YURET, Deniz, 2007: "KU: Word Sense Disambiguation by Substitution" in *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)*, Prague, 207-213.