Role and Reference Grammar and Ontological Engineering[1]

Ricardo Mairal Usón
Universidad Nacional de Educación a Distancia
rmairal@flog.uned.es

Carlos Periñán Pascual
Universidad Católica San Antonio
jcperinan@pdi.ucam.edu

ABSTRACT

The aim of this paper is to discuss the advantages of a conceptual approach to meaning representation within the framework of the Lexical Constructional Model (LCM), a model of meaning construction elaborated in the research Group Lexicom (www.lexicom.es). In the last two years, the LCM has expanded its architecture so as to accommodate a multilingual lexico-conceptual knowledge base, FunGramKB (www.fungramkb.com). We claim that, among the numerous applications that can emerge from this work (e.g. intelligent agents for cross-linguistic information retrieval applications, intelligent question-answering systems etc.), a knowledge base like this favours a conceptualist approach over a lexical one.

KEYWORDS: Lexical representation, ontological semantics, knowledge engineering, conceptual logical structures.

1. Introduction

Recent research has shown that Linguistics as a science should embrace a wider perspective by collaborating with other scientific disciplines (cf. Jackendoff, 2007:253). In this regard, Enrique Alcaraz was a forerunner since he always thought that Linguistics should go well beyond the study of language proper and consequently sought to break the traditional frontiers that have parcelled linguistics with the aim of bringing fresh air and an stimulating dialogue among disciplines together with a deep interest in finding ways to integrate advances attained in other sciences. The foundation of the IULMA (http://www.iulma.es/) illustrates Enrique's willingness to engage in conversation with scientists coming from different disciplines with the aim of finding new venues, frameworks, ideas, theories etc. that could enrich the linguistic debate as such. The following quote from Jackendoff (2002: xiv) illustrates the very essence of Enrique's legacy: his open-mindedness, his firm commitment to find points of convergence among disciplines and to develop a healthy and honest attitude towards the game of science.

> We cannot afford the strategy that regrettably seems endemic in the cognitive sciences: one discovers a new tool, decides it is the only tool needed, and, in an act of academic (and funding) territoriality, loudly proclaims the superiority of this new tool over all others. My own attitude is that we are in this together. It is going to take us lots of tools to understand language. We should try to appreciate exactly what each of the tools we have is good for, [and to recognize when new and as yet undiscovered tools are necessary.]
> This is not to advocate a warm fuzzy embrace of every new approach that appears on the scene. Rather, what is called for is an open-mindedness to insights from whatever quarter, a willingness to recognize tensions among apparently competing insights, and a joint commitment to fight fair in the interests of deeper understanding. To my mind, that's what the game of science is about.

In connection with this, this paper offers a very preliminary discussion of the latest research done within the Lexicom group (www.lexicom.es), that is, the construction of a comprehensive model of

[1] Financial support for this research has been provided by the DGI, Spanish Ministry of Science and Innovation, grant FFI2008-05035-C02-01.

meaning construction together with a large-scale lexico-conceptual knowledge base, a scientific task that demands the collaboration of both linguists and knowledge engineers. We claim that linguistics as a science needs to speak to other sciences if ground-breaking advances in the structure of language want to be attained, not to mention the impact of linguists' expertise on other fields of knowledge.

2. The scientific context

With the advent of the Web 3.0, or the Semantic Web, new search engines are being developed so that these can ultimately understand natural language queries and consequently information extraction or retrieval applications can provide much more efficient results than those obtained at present. In connection with this, a great number of research groups are now working on the development of intelligent agents that facilitate a more efficient access to information. These projects demand the expertise of linguists and knowledge engineers, each team working on the linguistic and the computational side respectively.

We maintain that a delicate semantic description such as that posited in the Lexical Constructional Model (LCM)[2] could fill in a gap in computational lexicography where, recall *pace* Velardi *et alii* (1991), the following two approaches are distinguished: the conceptual content of lexical units[3] 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). Strictly speaking, the latter doesn't give a real definition of the lexical unit, but it describes its usage in the language via "meaning relations" with other lexical units.[4] Most current natural language processing (NLP) systems adopt a relational approach to represent lexical meanings, since it is easier to state associations among lexical units in the way of meaning relations than describing formally the conceptual content of lexical units. However, although large-scale development of deep-semantic resources requires a lot of time, effort and expertise, two main deficiencies in surface semantics can be definitely overcome: its expressive power is dramatically restricted, and redundancy is highly spread through the knowledge base, as has been shown by Periñán-Pascual and Arcas-Túnez (2007b).

In connection with this, we have expanded and adapted previous work by Periñán-Pascual and Arcas-Túnez (2004, 2007a,b, 2008) and have developed an updated version of FunGramKB, a lexico-conceptual knowledge base that integrates very rich semantic and syntactic information that allows the development of natural language applications. In what follows, we will describe the linguistic and the technological contexts serving as a backdrop against which we will discuss the advantages of a conceptual approach for the anatomy of the lexicon as conceived in Role and Reference Grammar (RRG), a moderate functional theory of language (cf. Van Valin, 2005).

2.1. The linguistic context: levels of meaning construction

In the last few years Mairal and Ruiz de Mendoza (2009) and Ruiz de Mendoza and Mairal (2008) have developed the Lexical Constructional Model (LCM), a model that accounts for all those aspects involved in meaning construction, including those that go beyond core grammar, i.e. pragmatic implicature, illocutionary force and discourse coherence. Unlike some other existing linguistic models, the LCM is meant to provide an exhaustive, fine-grained description of meaning, which will then serve as input for the syntactic apparatus, the output of which will be mapped to phonological form.[5] The LCM distinguishes the following levels of meaning construction (cf. Ruiz de Mendoza and Mairal, 2008):

> Level 1 or argumental layer is concerned with those linguistic features producing core grammar characterizations.
> Level 2 or implicational layer includes the set of linguistic mechanisms accounting for heavily conventionalized situation-based low-level meaning implications.
> Level 3 or illocutionary layer is concerned with those linguistic aspects that account for conventionalized illocutionary meaning (situation-based high-level implications).

---

[2] We refer the reader to the webpage www.lexicom.es for a full description and coverage of the LCM.

[3] In this paper, the term "lexical unit" is used as a synonym of "predicate", i.e. content words to which morphosyntactic and semantic properties are assigned.

[4] EuroWordNet (Alonge *et alii* 1998, Vossen 1998) is one of the best-known examples of a multilingual "relational" database, which provides elaborate lexical networks by means of semantic relations between *synsets* (or clusters of synonymous words) within every language-dependent wordnet.

[5] For the syntactic apparatus, the LCM follows the RRG linking algorithm in the sense that theoretical notions such as macroroles, privileged syntactic arguments etc. are used for the description of the syntactic phase. In other words, we could say that the LCM provides a thorough semantic description that enriches the RRG semantic representation.

Level 4 or discoursive layer captures the set of linguistic mechanisms that are based on very schematic discourse structures.

These four different layers are interrelated by two cognitive processes: *subsumption* and *cueing.* In order to illustrate how the LCM provides fully-fledged semantic descriptions, let us see the format of a lexical entry as represented in the argumental module (cf. Ruiz de Mendoza and Mairal, 2008).

Lexical entries are represented in terms of 'lexical templates', which are originally a development of the logical structures in RRG (cf. Van Valin and LaPolla, 1997; Van Valin, 2004, 2005). Recall that RRG uses a decompositional system for representing the semantic and argument structure of verbs and other predicates (i.e. logical structures). The verb class adscription system is based on the *Aktionsart* distinctions proposed in Vendler (1967), and the decompositional system is a variant of the one proposed in Dowty (1979). Verb classes are divided into *states, activities*, *achievements*, *semelfactives*, and *accomplishments*, together with their corresponding causatives. Table 1 shows a representation of each verb class with their corresponding formalism (cf. Van Valin, 2005:45).

Table 1. RRG typology of verb classes.

VERB CLASS	LOGICAL STRUCTURE	EXAMPLE	INSTANTIATION OF LS
State	**predicate'** (x) or (x,y)	see	**see'** (x,y)
Activity	**do'** (x, [**predicate'** (x) or (x,y)]	run	**do'** (x,[**run**' (x)])
Achievement	INGR **predicate'** (x) or (x,y), *or* INGR **do'** (x, [**predicate'** (x) or (x,y)]	pop (burst into tears)	INGR **popped'** (x)
Semelfactive	SEML **predicate'** (x) or (x,y) SEML **do'** (x, [**predicate'** (x) or (x,y)]	glimpse, cough	SEML **see'** (x,y)
Accomplishment	BECOME **predicate'** (x) or (x,y), *or* BECOME **do'** (x, [**predicate'** (x) or (x,y)]	receive	BECOME **have'** (x,y)
Active accomplishment	**do'** (x, [**predicate$_1$'** (x, (y))] & INGR **predicate$_2$**' (z,x) or (y)	drink	**do'** (x,[**drink**' (x,y)]) & INGR **consumed'** (y)
Causative accomplishment	α CAUSES ß where α, ß are LS of any type	kill	[**do'** (x, ∅)] CAUSE [BECOME [**dead'** (y)]

RRG maintains that states and activities are primitives and thus form part of the logical representation of other predicates; by way of example, an accomplishment is either a state or activity predicate modified by the telic operator BECOME. However, Van Valin and Wilkins (1993) and Van Valin and LaPolla (1997) all claim quite explicitly that state and activity atomic predicates need further semantic decomposition and thus provide a first approach for the predicate *remember* and speech act verbs respectively.[6]

In an attempt to provide logical structures with a more robust semantic decomposition, we decided to develop the notion of lexical template. A lexical template then enriches RRG logical structures by including a semantic component that captures those pragmatic and semantic parameters that are idiosyncratic to the meaning of a word. For example, if we want to account for the semantic differences between *cautivar* ('captivate'), *arrebatar* ('seize'), *arrobar* ('entrance'), *embelesar* ('enrapture'), *extasiar* ('send into an ecstasy'), *hechizar* ('bewitch') from the domain of feeling in Spanish, we would certainly need some mechanism that allows us to discriminate and encode those meaning elements that differentiate one predicate from others. Here is the basic representational format for a lexical template:

(1)
predicate: [SEMANTIC MODULE<qualia>] [AKTIONSART MODULE <semantic primes>]

A lexical template consists of two modules: (i) the semantic module, and (ii) the logical representation or *Aktionsart* module, each of which is encoded differently. The right hand part of the representation includes the inventory of logical structures as developed in RRG with the proviso that the predicates used as part of the meaning definition are now conceptual units that are linked to the ontology. The semantic module consists of a number of internal variables, i.e. world knowledge elements of semantic structure, which relate in very specific ways to the external variables that account for those arguments that have a

[6] For a discussion of the exact details of the formalism of the first lexical templates, we refer the reader to Van Valin and Wilkins (1993), Van Valin and LaPolla (1997), Mairal and Faber (2002, 2007), Mairal and Ruiz de Mendoza (2008).

grammatical impact. We have used Pustejovsky's (1995) quale as part of the notational device to represent this semantic information. Consider the following example[7]:

(2)
fathom:
EVENTSTR: **know'** (x, y)
QUALIASTR: {Q_F: **MANNER** : MagnObstr **think'** (x, y)
Q_T: Culm **know'** (x,y <ALL>)}

The *Aktionsart* module specifies that this predicate is a state predicate that takes two variables (*x, y*). This state structure is in turn defined by a primitive predicate **know**', which, together with the primitive **think'**, are the two defining predicates for the complete lexical domain of cognitive verbs. The semantic module is expressed in terms of two *qualia*: the *formal* and the *telic*. The formal *quale* describes the great difficulty involved in carrying out the process of thinking, i.e. it includes the semantic attributes by means of which *fathom* is semantically distinguished within the larger set of cognition predicates in English[8]. The telic, as encoded in Q_T: Culm **know'** (x,y), specifies the culmination of the process of acquisition of knowledge, that is, the final process of understanding something.

Hence, the LCM represents lexical entries with a formalism that includes a description that goes well beyond those aspects that are grammatically relevant and thus include all those semantic and pragmatic parameters that, although these are not syntactically visible, are also part of the meaning of a word (cf. Levin and Rappaport, 2005:chapter 7). Thus, if the output of the LCM is a fine-grained delicate description of all aspects involved in meaning construction, then the LCM offers a very nice framework for the development of NLP applications based on a deep semantic approach.

2.2. The technological context: The architecture of a knowledge base

FunGramKB is a multipurpose lexico-conceptual knowledge base for NLP systems, mainly those requiring language understanding. FunGramKB is made up of two information levels, which in turn consist of several independent but interrelated modules:

Lexical level (i.e. linguistic knowledge):

- The lexicon stores morphosyntactic, pragmatic and collocational information about lexical units.[9]
- The morphicon helps our system to handle cases of inflectional morphology.

Conceptual level (i.e. non-linguistic knowledge):

- The ontology is presented as a hierarchical catalogue of all the concepts that a person has in mind when talking about everyday situations. Here is where semantic knowledge is stored in the form of meaning postulates.
- The cognicon stores procedural knowledge (e.g. how to fry an egg, how to buy a product, etc.) by means of conceptual proto-macrostructures, i.e. script-like schemata in which a sequence of stereotypical actions is organised on the basis of temporal continuity, and more particularly on Allen's temporal model (Allen, 1983; Allen and Ferguson, 1994).
- The onomasticon stores information about instances of entities and events, such as Bill Gates, Taj Mahal, or 9/11. This module stores two different types of schemata (i.e. snapshots and stories), since instances can be portrayed synchronically or diachronically.

[7] We refer the reader to Mairal and Ruiz de Mendoza (2008) for further discussion on the format of lexical templates and quale.
[8] Note that 'Magn Obstr' are two lexical functions that have been formulated within the framework of Mel'cuk's Explanatory and Combinatorial Lexicology (ELC) (cf. Aloson Ramos, 2002).
[9] We refer the reader to Periñán and Mairal (2009) for a detailed description of the features and values in FunGramKB lexical entries.

It is important to note that while the lexical module captures language specific properties, the conceptual level is language independent (cf. Section 3 for the notion of 'universality'). This means that knowledge engineers build just one ontology, one cognicon and one onomasticon, while linguists have to develop one lexicon and one morphicon for German, one lexicon and one morphicon for Spanish and so on for each of the languages involved in the knowledge base[10].

The FunGramKB ontology distinguishes three different conceptual levels, each one of them with concepts of a different type:

(i) Metaconcepts (e.g. #ABSTRACT, #COLLECTION, #MOTION, #POSSESSION, #TEMPORAL etc) constitute the upper level in the taxonomy, as a result of the analysis of the most relevant linguistic ontologies, i.e. DOLCE, Generalized Upper Model, SIMPLE, SUMO etc. The FunGramKB ontology is actually split into three subontologies, since subsumption (IS-A) is the only taxonomic relation permitted, and therefore each subontology arranges lexical units of a different part of speech: i.e. #ENTITY for nouns, #EVENT for verbs, and #QUALITY for adjectives and some adverbs.

(ii) Basic concepts, which are preceded by symbol +, are used as defining units which enable the construction of meaning postulates for basic concepts and terminals, as well as taking part as selection preferences in thematic frames: e.g. +FAR_00, +FEEL_00, +HOT_00, +MONEY_00, +SAY_00, +WINDOW_00 etc.

(iii) Terminals, which are headed by symbol $, are those concepts which have no definitory potential to take part in meaning postulates: e.g. $AVENUE_00, $DAGGER_00, $GLEAM_00, $JULY_00, $OBSOLETE_00, $SENILE_00 etc.

In FunGramKB, basic and terminal concepts are not stored as atomic symbols but are provided with semantic properties such as the thematic frame and the meaning postulate. Both of them are conceptual schemata, since they employ concepts—and not words—as the building blocks for the formal description of meaning. Thus, thematic frames as well as meaning postulates become language-independent semantic knowledge representations.

Every event and quality in the ontology is assigned one thematic frame, i.e. a prototypical conceptual construct which states the number and type of participants involved in the cognitive situation portrayed by the concept. Consider the thematic frame of +OPEN_01, to which lexical units such as *open* [Eng] or *abrir* [Spa] are linked:

(3)
TF: (x1)Agent (x2: +DOOR_00 ^ +WINDOW_00)Theme (x3)Location (x4)Origin (x5)Goal

Thematic frames can also include those selectional preferences typically involved in the cognitive situation.[11] Thus, thematic frame (1) describes a prototypical cognitive scenario in which "entity$_1$ (Agent) moves entity$_2$ (Theme)—being typically a door or window—from one place (Origin) to another (Goal), there also being a place (Location) along which entity$_2$ moves". It should not be forgotten that, although one or more subcategorization frames can be assigned to a single lexical unit, every concept is provided with just one thematic frame.

Alternatively, a meaning postulate is a set of one or more logically connected predications (e_1, e_2... e_n), which are conceptual constructs carrying the generic features of the concept[12]. Meaning postulates integrate the information stated in a thematic frame. For example, the basic concept + WEAVE_00, which belongs to the metacognitive dimension #TRANSFORMATION > #CREATION, specifies a thematic frame such that an entity (Theme), typically a human being, creates another entity (Referent), typically a cloth or a basket:

(4)
TF: (x1: +HUMAN_00)Theme (x2: +CLOTH_00 ^ +BASKET_00)Referent

This information is fully integrated into the meaning postulate, which presents a semantic representation of the basic concept +WEAVE_00:

[10] English and Spanish are fully supported in the current version of FunGramKB, although we have just begun to work with other languages, i.e. German, French, Italian, Bulgarian and Catalan.
[11] Selectional preferences are stated when they can exert some predictive power on the participant. A protocol is currently being developed in order to determine coherent membership criteria for the selectional preferences of participants in thematic frames.
[12] Periñán-Pascual and Arcas-Túnez (2004) describe the formal grammar of well-formed predications for meaning postulates in FunGramKB.

(5) MP: +(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +HAND_00 | +MACHINE_00)Instrument (f2: +THREAD_00 ^ +STRING_00)Means (f3: +HUMAN_00)Beneficiary)

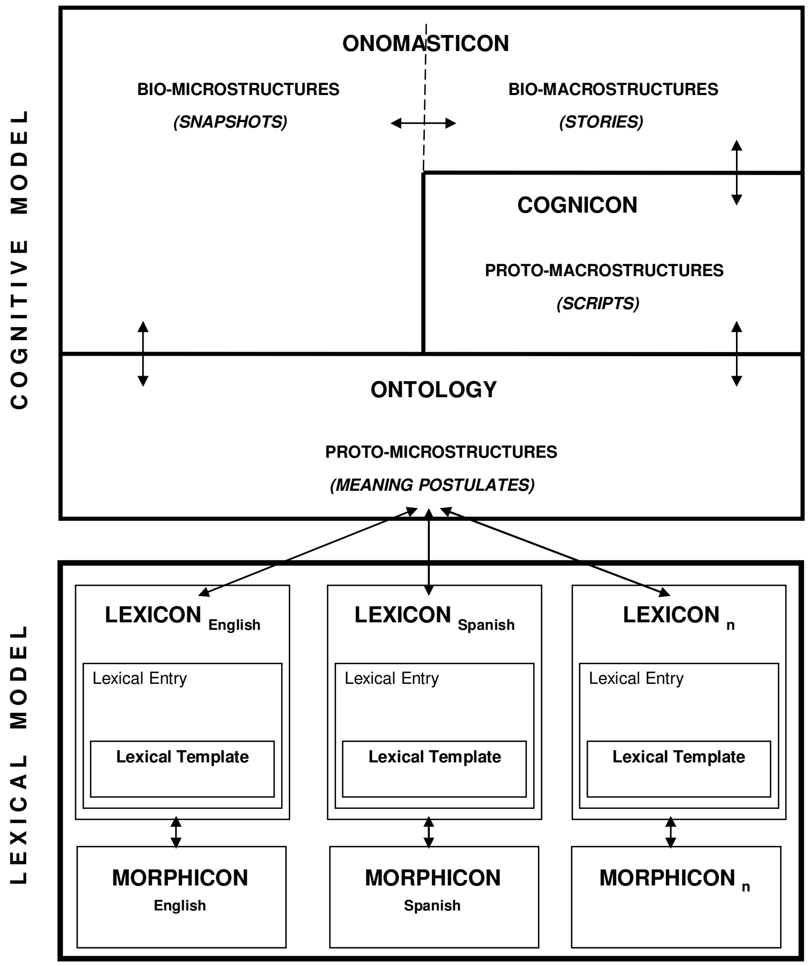
The genus of the meaning postulate is the superordinate +CREATE_00, which is modified by three differentiae (f1), (f2) and (f3) that express the instrument, the means and the beneficiary respectively. The definition would go as follows: an entity creates another entity for someone (Beneficiary) with an Instrument (typically a hand or a machine), and by using a string or a thread (Means).

In essence, FunGramKB is provided with a universal, linguistically-motivated and general-purpose ontology. At this stage, a further issue that arises is to explore whether a conceptual approach offers a more viable framework for meaning construction than a lexical approach.

3. Ontological semantics in RRG

As shown in Figure 1, FunGramKB adopts a conceptualist approach, whereby the ontology is the pivotal component that feeds the lexicon. In this section, we will deal with how the ontology and the lexicon interact, and more particularly, how conceptual properties such as thematic frames and meaning postulates interact with information stored in lexical entries, resulting in the construction of *Conceptual Logical Structures* (CLS)[13]. We claim that an ontological semantic approach minimizes redundancy and maximizes informativeness in the lexicon. In what follows we will highlight some of the advantages

Figure 1. FunGramKB modules.



[13] This type of approach is very similar in many respects to Jackendoff's (1990; 2002) universal approach since the different lexica are largely based on one single ontology. Thus, the role of a CLS is to serve as a bridge between the more abstract level as represented in the ontology and the particular idiosyncrasies as coded in a given linguistic expression.

The benefits of a conceptual approach to the classical model of logical structure can be summarized as follows. Firstly, lexical representations in the form of CLSs now become real language independent representations, since these are made of concepts and not words. Let us clarify what we mean by 'universal' here. CLSs are not made of words, i.e. lexical material, but of concepts. Recall that our ontology takes the form of a universal concept taxonomy, where "universal" means that every concept[14] we imagine has, or can have, an appropriate place in this ontology (Corcho, Fernández López and Gómez Pérez 2001). Instead of adopting a strong approach like that represented by the Natural Semantic Metalanguage approach (cf. Goddard and Wierzbicka, 2002), which identifies a reduced inventory of semantic primitives that are used to represent meaning, FunGramKB posits an inventory of basic concepts which are not all deemed as universal primitives, unless a so-called 'stepwise conceptual decomposition'[15] is applied to CLSs. Our ultimate aim is to have an inventory of concepts, regardless of whether these are universal or not, that can be used to define any word in any of the European languages that are claimed to be part of the ontology we have designed, i.e. any word can be defined by using the basic concepts we have in the ontology, which does not necessarily mean that the actual design of the ontology corresponds to a universal representation of the world. Hence, CLSs are lexically-oriented schemata containing elements of a universal semantic metalanguage, which means that these representations have cross-linguistic validity.

As for the second feature, a conceptualist approach like the one maintained here offers an elegant solution for the choice of the correct prime. In connection with this, we claim that a prime or an undefinable is always associated to the concept to which the lexical unit is linked. This allows us to avoid the problem of having to regard as undefinable predicates which can be further semantically decomposed, e.g. defining the predicate *redden* in terms of BECOME **red'**, or *popped* in terms of INGR **popped'**, or activity predicates like *sing* or *drink* in terms of **do'** (x,[**drink**' (x)]) or **do'** (x,[**sing**' (x)]). The innovation here with respect to the original RRG proposal resides in finding a systematic procedure to identify the correct prime together with a uniform framework for decomposing semantically every predicate until we arrive at the undefinable elements. Consider the Spanish predicate *arreglar,* which has two potential meanings: (i) put into a proper or systematic order (e.g. "arrange the flowers in the vase"), and (ii) restore by replacing a part or putting together what is torn or broken (e.g. "She offered to repair my TV set"; "Will you repair my shoes please"). Each sense will be linked to a different concept, +TIDY_00 and +REPAIR_00 respectively, which will be represented by the following meaning postulates:

(6) +TIDY_00
+(e1: +DO_00 (x1)Theme (x3)Referent (f1: (e2: +BE_01 (x2)Theme (x4: $TIDY_D_00)Attribute))Result)

(7) +REPAIR_00
+(e1: +CHANGE_00 (x1)Theme (x2)Referent (f1: (e2: past +DAMAGE_00 (x3)Theme (x2)Referent))Reason (f2: (e3: pos +OPERATE_00 (x4)Theme (x2)Referent))Result)

Both would be part of the #TRANSFORMATION dimension and the two relevant participants are a Theme (i.e. entity that transforms another entity) and a Referent (i.e. the entity that is transformed by another entity). Then, (6) has the following interpretation: an entity does something to another entity with the result that a different entity is now tidy. Additionally, (7) is interpreted as follows: an entity x_1 changes an entity x_2 because an entity x_3 has damaged x_2 with the result that an entity x_4 operates on x_2. The following two CLSs will be automatically generated:

(8) [+DO_00 (x_{Theme}, 0)] CAUSE [BECOME +TIDY_00 ($y_{Referent}$)]

(9) [+DO_00 (x_{Theme}, 0)] CAUSE [BECOME +REPAIR_00 ($y_{Referent}$)]

In sum, a conceptualist approach opens the door to the ability to link primes as posited in CLSs with conceptual units in the ontology. From this it follows that the use of concepts as building blocks in

[14] Terms such as "class", "category" or "semantic type" are often used in ontology engineering to refer to elements such as FunGramKB "concepts". However, we prefer the latter, since it better describes the domain of processing in the two-tier model of our NLP knowledge base, i.e. lexical level and conceptual level.

[15] This task is partly founded on Dik's (1989) 'stepwise lexical decomposition' principle, which establishes a way of interrelating lexical entries where the *definiens* in a meaning postulate can be converted into the *definiendum* of another meaning postulate. The fact that Dik's functional model is provided with a lexical decomposition process enables the construction of meaning postulates in a simple fashion, as well as minimizing information redundancy. In FunGramKB, the stepwise conceptual decomposition is similar to Dik's principle, apart from the fact that the main building blocks of our meaning representations are not lexical units but concepts. Therefore, the stepwise conceptual decomposition is defined as the process in which conceptual units in a predication are substituted by their respective meaning postulates until a meaning representation composed of metaconceptual primitives is reached.

lexical meaning representations removes the problem of semantic ambiguity and, since the semantic knowledge repository is shared by all lexica, this also provides a consistent procedure to successfully reach standardized *definiens*.

Thirdly, we maintain that CLSs together with the meaning postulates and the thematic frames in the ontology provide a full-grained semantic decomposition provided that meaning postulates contain the world-knowledge information of the concept that lexical entry is linked to: every subcategorised element in the CLS of a lexical unit is referenced through thematic roles to a participant in the thematic frame of the concept to which that lexical unit is linked, and, in turn, every participant in that thematic frame is referenced through co-indexation to a participant in the meaning postulate of that concept. To illustrate, let us consider the CLSs for the verb *write* in the instance (10a).

(10) a. John wrote a letter to Peter.
b. [+DO_00 (%JOHN_00$_{\text{Theme}}$, [+WRITE_00 (%JOHN_00$_{\text{Theme}}$, +LETTER_00$_{\text{Referent}}$)]) & INGR +EXIST_00 (+LETTER_00$_{\text{Referent}}$)] PURP +HAVE_00 (%PETER_00$_{\text{Goal}}$ +LETTER_00$_{\text{Referent}}$)[16]

This verb designates an active accomplishment which is further modified by a default argument (*to Peter*) (Pustejovsky 1995: 63), which functions as an argument adjunct with a purposive meaning (cf. Van Valin and LaPolla, 1997: 382-384). The CLS (10b) can be enriched by the semantic properties of the main concepts involved, i.e. +WRITE_00 and +LETTER_00 as specified in (11).

(11) +WRITE_00
Thematic Frame:
(x1: +HUMAN_00)Theme
(x2: +WRITING_00 ^ +IMAGE_00)Referent

Meaning Postulate:
+(e1: +CREATE_00 (x1)Theme (x2)Referent (f1: +PENCIL_00 ^ +PEN_00 ^ +MACHINE_00)Instrument (f2: +PAPER_00 | +WALL_00)Location (f3: +ON_00)Position)

+LETTER_00
Meaning Postulate:
+(e1: +BE_00 (x1: +LETTER_00)Theme (x2: +DOCUMENT_00)Referent)
*(e2: +PUT_00 (x3: +HUMAN_00)Agent (x1)Theme (x4)Origin (x5: +ENVELOPE_00)Goal (f1: +IN_00)Position (f2: (e3: +SEND_00 (x3)Agent (x5)Theme (x3)Origin (x6: +HUMAN_00)Goal))Purpose)

In this case, and when all constructs are fully integrated, the FunGramKB reasoner can infer that "John probably used a pencil, a pen, or a machine to write the letter" and that "he probably put the letter inside an envelope in order to send it to Peter".

4. Concluding remarks

Recent research in the Lexicom research group has shown that the LCM, a comprehensive model for meaning construction, can accommodate a lexical conceptual knowledge base, which will eventually serve for the development of NLP applications, i.e. the design of intelligent agents in the framework of the semantic web, the development of intelligent question-answer systems that go beyond factoid questions, cross-linguistic information retrieval applications etc. This type of research evidences the urgent need to collaborate with other disciplines and then cultivate more fertile challenges for linguistic research.

In this context, Enrique was also a leader for all of us. He has left us with an invaluable intellectual legacy, i.e. an impressive and brilliant scientific work as represented in his extensive scientific production. But, Enrique's legacy goes well beyond his actual linguistic production and steps into a more qualitative sphere where he succeeded in transmitting an overwhelming enthusiasm and love for the study of language in all its different manifestations. Moreover, his open-mindedness towards new ideas and theories together with an attitude firmly seated in rigorous criteria to recognize tensions among different competing frameworks and apprehend the best of both worlds made Enrique a prestigious scholar and a great colleague and friend. His attitude towards Linguistics, and we'd dare say, towards life is the best tribute that will always preserve in our memory. Thanks, Enrique, for teaching us what the game of science should consist of.

[16] Note that this representation does not account for the rest of the semantic operators (i.e. tense, aspect, modality, illocutionary force etc.) that are also part of the conceptual logical representation of this sentence.

Bibliography

Allen, J. F. (1983): “Maintaining knowledge about temporal intervals”. Communications of the ACM 26 (11), pp. 832-843.

Allen, J. F. and G. Ferguson. (1994):. “Actions and events in interval temporal logic”. *Journal of Logic and Computation* 4 (5), pp. 531-579.

Alonge, A., N. Calzolari, P. Vossen, L. Bloksma, I. Castellón, M. A. Martí and W. Peters. (1998): “The linguistic design of the EuroWordNet database”. *Computers and the Humanities* 32, 2-3. 91-115.

Alonso Ramos, M. (2002): “Colocaciones y contorno en la definición lexicográfica”, *Lingüística Española Actual* 24, 1: 63-96.

Corcho, O, M. Fernández López and A. Gómez Pérez. (2001): *Technical Roadmap v. 1.0*, IST-OntoWeb Project, Madrid, Universidad Politécnica de Madrid. In: http://www.ontoweb.org/download/deliverables/D11_v1_0.pdf

Dik, S.C. (1989): *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Berlin /New York: Mouton de Gruyter. Revised edition by Kees Hengeveld, 1997.

Dowty, D. (1979): *Word meaning and Montague Grammar*. Dordrecht: Foris.

Goddard, C, and A. Wierzbicka (eds.) (2002): *Meaning and Universal Grammar*, Amsterdam: John Benjamins.

Jackendoff, R. (1990): *Semantic Structures*. Cambridge (Mass.): MIT Press.

Jackendoff, R. (2002): *Foundations of Language*. Oxford, Oxford University Press.

Jackendoff, R. (2007): “A Whole ot of Challenges for Linguistics” *Journal of English Linguistics*, 2007; 253-262.

Levin, B. and M. Rappaport Hovav. (2005): *Argument Realization*, Research Surveys in Linguistics Series, Cambridge University Press, Cambridge, UK.

Mairal Usón, R. and P. Faber. (2002): “Functional Grammar and lexical templates”. In Mairal, R. and M.J. Pérez Quintero (eds.): *New perspectives on argument structure in Functional Grammar*, pp. 41-98. Berlin /New York: Mouton de Gruyter.

Mairal Usón, R. and P. Faber. (2007): “Lexical templates within a functional cognitive theory of meaning”. *Annual Review of Cognitive Linguistics* (5), 137-172. Ámsterdam: John Benjamins.

Mairal Usón, R. and F. Ruiz de Mendoza. (2008): “New challenges for lexical representation within the Lexical-Constructional Model (LCM)” *Revista Canaria de Estudios Ingleses*. Vol. 57. pp. 137-158 (2008). Universidad de La Laguna.

Mairal Usón, R. and F. Ruiz de Mendoza (2009): “Levels of description and explanation in meaning construction”. In C. S. Butler and J. Martín Arista (eds.) *Deconstructing Constructions*. Ámsterdam/ Philadelphia: John Benjamins, pp. 153-198.

Mel’cuk, I. (1989): “Semantic primitives from the viewpoint of the Meaning-Text Linguistic Theory.” *Quaderni di Semantica* 10 (1): 65-102.

Periñán-Pascual, C. and F. Arcas-Túnez. (2004): “Meaning postulates in a lexico-conceptual knowledge base”, *15th International Workshop on Databases and Expert Systems Applications*, IEEE, Los Alamitos (California), 38-42.

Periñán-Pascual, C. and F. Arcas-Túnez.. (2007a): “Cognitive modules of an NLP knowledge base for language understanding”, *Procesamiento del Lenguaje Natural* 39, 197-204.

Periñán-Pascual, C. and F. Arcas-Túnez (2007b): “Deep semantics in an NLP knowledge base”, *12th Conference of the Spanish Association for Artificial Intelligence*, 279-288.

Periñán-Pascual, C. and F. Arcas-Túnez. (2008): “Modelling OLIF frame with EAGLES/ISLE specifications: an interlingual approach”. *Procesamiento del Lenguaje Natural* 40: 9-16.

Periñán-Pascual, C. and R. Mairal (2009): “The FunGramKB lexical model” *Revista Española de Lingüística Aplicada.*

Procter, P. (ed.) (1978): *Longman Dictionary of Contemporary English*. Harlow (Essex): Longman.

Pustejovsky, J. (1995): *The Generative Lexicon*. Cambridge, Mass.: MIT Press.

Ruiz de Mendoza Ibáñez, F. and R. Mairal, (2008): “Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model”. *Folia Linguistica* 42/2 (2008), 355–400.

SIMPLE Specification Group. (2000): “Specification SIMPLE Work Package 2”. Linguistic Specifications Deliverable D2.1.

Van Valin, R. D. Jr. (2004): “Lexical representation, co-composition, and linking syntax and semantics”. Unpublished ms. University at Buffalo. (Available from the RRG website).

Van Valin, R. D. Jr. (2005): *Exploring the Syntax-Semantics Interface*, Cambridge: Cambridge University Press.

Van Valin, R. D. Jr., and D. P. Wilkins. (1993): “Predicting Syntactic Structure from Semantic Representations: remember in English and its equivalents in Mparntwe Arrernte·. In R. D. Van Valin Jr. (ed.): *Advances in Role in Reference Grammar*. Amsterdam/Philadelphia: John Benjamins, pp. , 499-534

Van Valin, R. D. Jr., and R. LaPolla. (1997): *Syntax, Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Velardi, P., M.T. Pazienza and M. Fasolo. (1991): “How to encode semantic knowledge: a method for meaning representation and computer-aided acquisition”. *Computational Linguistics* 17, 2. 153-170.

Vendler, Z. (1967): *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

Vossen, P. (1998): “Introduction to EuroWordNet”. *Computers and the Humanities* 32, 2-3. 73-89.