

Representing Constructional Schemata in the FunGramKB Grammaticon¹

Ricardo Mairal & Carlos Periñán-Pascual

Recent research into FunGramKB has focused on the development of a proof-of-concept prototype, ARTEMIS, which intends to automatically provide a semantic representation of a text under the format of a *conceptual logical structure* by viewing the RRG linking algorithm from a conceptual perspective. However, little has been said about the format of the Grammaticon, the place where constructional schemata are claimed to reside within FunGramKB. With this in mind, the aim of this chapter is to discuss the format of constructional schemata in ARTEMIS. ARTEMIS's explanatory scope is not confined to argument constructions, as has been the case in RRG so far and most construction grammar approaches, but takes a step forward to account for those meaning dimensions that have a long tradition in pragmatics and discourse analysis, that is, the non-propositional dimension of meaning. In so doing, ARTEMIS resorts to the Lexical Constructional Model, a comprehensive model of meaning construction. The primary aim of this chapter is to discuss the format of these four level schemata and their representation in a natural language engineering project like ARTEMIS.

1 Introduction

FunGramKB Suite is an online knowledge-engineering environment for the semi-automatic construction of a multipurpose lexico-conceptual knowledge base for natural language processing (NLP) systems, i. e. FunGramKB (Periñán-Pascual & Arcas 2004, 2007, 2010, Periñán-Pascual & Mairal 2009, 2010a, Mairal & Periñán-Pascual 2009). On the one hand, FunGramKB is multipurpose in the sense that it is both multifunctional and multilingual. In other words, FunGramKB has been

¹ Financial support for this research has been provided by the DGI, Spanish Ministry of Education and Science, grant FFI2011-29798-C02-01.

designed to be reused in various NLP tasks (e. g. information retrieval and extraction, machine translation, dialogue-based systems, etc.) and with several natural languages. The English and Spanish lexica are fully supported in the current version of FunGramKB, although we also work with other languages, i. e. German, French, Italian, Bulgarian and Catalan. On the other hand, the knowledge base is lexico-conceptual, because it comprises three general levels of information: lexical, grammatical and conceptual.² Recent research into FunGramKB Suite has resulted in 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) (Perrián-Pascual 2013b, Perrián-Pascual & Arcas 2014). This research is based on previous work, which has explored the methodological repercussions of viewing the Role and Reference Grammar (hereafter, RRG) linking algorithm from a conceptual perspective. A brief description of some of the most relevant working proposals, together with some of the most representative references, is outlined below.

a) The RRG linking algorithm is claimed to have a conceptual grounding such that there exists a specific knowledge base, i. e. FunGramKB, which interfaces with the different linguistic modules of the RRG linking algorithm. Figure 1 illustrates this cognitive turn:

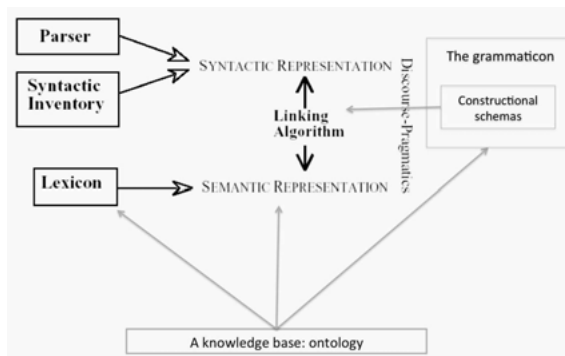


Figure 1: The RRG architecture within a conceptual framework.

² For further information about this previous background work, we refer the reader to the following websites where some of the most relevant literature can be downloaded: www.fungramkb.com and www.lexicom.es

The resulting semantic representations have a cognitive status, concurring with the assumption that primes in RRG standard *logical structures* are thought to be ontologically driven. However, nothing is said about the internal structure of the ontology that supports the cognitive nature of these primes. This is why this proposal comes to fill in an existing gap in the RRG apparatus.

b) From the preceding discussion, it follows that standard RRG lexical representations, i. e. *logical structures*, are now conceived in terms of *conceptual logical structures* (CLS) (cf. Mairal, Perinián-Pascual & Pérez 2012, Mairal 2012, Perinián-Pascual & Arcas 2014) for a detailed explanation of the format of this new type of lexical representation). This implies that primes are substituted for conceptual units that are part of the FunGramKB Ontology. Each conceptual unit, in turn, has its own distinguishing properties, i. e. a *thematic frame* and a *meaning postulate*³, so that the resulting representation provides access to knowledge which goes beyond the grammatically relevant aspects of meaning. For example, a predicate like ‘marchitar’ (*wither*) in Spanish is linked to the terminal⁴ conceptual unit \$WITHER_00 in the Ontology. In terms of RRG, this is a one-place predicate which designates an accomplishment:

(1) BECOME **withered**’ (x)

Hence, this representation only captures those aspects that are grammatically relevant, while nothing is said about those features that go beyond syntax. If conceptual units from an ontology are used instead of lexical units, then the resulting representation will have access to the information provided by the thematic frame and meaning postulate of the conceptual unit to which the predicate in the lexicon is linked to. In the case that concerns us here, as noted above, in the FunGramKB Spanish Lexicon the predicate ‘marchitar’ is linked to the terminal conceptual unit \$WITHER_00, which includes a thematic frame (2a) with two arguments, the second of which is restricted by a number of selectional preferences. Moreover, this conceptual unit is provided with a meaning definition (2b) such that

³ For a detailed description of the technicalities of these two notions, *thematic frames* and *meaning postulates*, we refer the reader to Perinián-Pascual & Arcas (2010), Perinián-Pascual & Mairal (2010a) and Mairal (2012).

⁴ FunGramKB, unlike some other existing ontological engineering projects, follow a four-level classification of concepts: metaconcepts, basic concepts, terminal concepts and subconcepts, each occupying a different place in the conceptual hierarchy and represented by a different notational system (cf. Perinián-Pascual & Arcas 2010, Perinián-Pascual & Mairal 2010a, 2011).

someone dries something (usually a plant, flower or leaf) and as a result this entity becomes small and weak and begins to die. On the whole, an ontological framework provides an enriched version of standard logical structure (see Van Valin & Mairal (in press) for further elaboration of this issue)⁵:

(2) \$WITHER_00

- a. (x1)Theme (x2: +PLANT_00 ^ +FLOWER_00 ^ +LEAF_00) Referent
- b. +(e1: +DRY_00 (x1) Theme (x2) Referent (f1: (e2: +BECOME_00 (x2)Theme (x3: +SMALL_00 & +WEAK_00) Attribute)) Result (f2: (e3: ing +DIE_00 (x2) Theme)Result
Entity1 dries entity2, typically a plant, flower or leaf, and consequently entity2 becomes small and weak and starts to die.

c) The final output is a fully specified semantic representation that is built around the notion of *aktionsart* and consists of conceptual units and operators (as those distinguished in Van Valin & LaPolla 1997, Van Valin 2005)⁶, all of which offer a very comprehensive picture of a natural language user's linguistic knowledge of an input text (see Perinián-Pascual & Mairal 2009, 2012):

(3) a. Ron destroyed the building

- b. <_{IF}DEC <_{TNS}PAST <_{ASP}PERF <_{CONSTR-L1}KER2 <[_{AKT}ACC [+DESTROY_00 (%RON_00-Theme, \$BUILDING_00-Referent)]]>>>>

Hence, a semantic representation expressed in terms of a CLS is now connected up to a knowledge base from where it is possible to retrieve world knowledge information via a reasoning engine. By using artificial intelligence techniques, e. g. graph-based algorithms, our semantic knowledge as expressed in a CLS can be further enriched by information coming from different modules of the knowledge base (i. e. the Ontology, the Cognicon and the Onomasticon)⁷.

⁵ Note that the properties of conceptual units are expressed in COREL (Conceptual Representational Language). COREL-formatted schemata, which can be computationally considered as a combination of conceptual graphs and frames, are modelled through propositional representations consisting of logically-connected predications. We refer the reader to Perinián-Pascual & Arcas (2004) and Perinián-Pascual & Mairal (2010a) for a full description of its technicalities.

⁶ As noted in (3b), there are two operators, AKT and CONSTR-L1, which are not part of Van Valin's (2005) inventory. These will be discussed later in Section 4.1.

⁷ The cognitive level in FunGramKB, as extensively discussed elsewhere (cf. Perinián-Pascual & Arcas 2007, Perinián-Pascual 2012, 2013), consists of three modules: the Ontology where semantic knowledge is stored in terms of a hierarchy of concepts; (ii) the Cognicon where procedural knowledge is

d) In Figure 1, we note that a specific place has been allocated for constructional schemata given their prominent role in the lexical-grammatical interface⁸. However, in our view these constructional schemata have still a very linguistic descriptive flavor, which makes it a bit difficult to make use of when one is confronted with a natural language processing application. In connection with this, we believe that constructional schemata can be enriched if these are based on a conceptual framework. In this regard, there are two specific aspects that are worth mentioning: (i) the first one is concerned with the format of the constructional schemata; a more formalized scheme than the present RRG constructional schemata is needed; (ii) the second has to do with the fact that the semantics of standard RRG constructional templates can be enhanced via information coming from the Ontology (cf. Van Valin & Mairal in press, Perrián-Pascual 2013b, Mairal, 2012).

In essence, this previous work, which underlines the methodological advantages of driving RRG into a conceptually-oriented paradigm, has been influential in the first stages of the development of ARTEMIS. Notwithstanding, we still noted that an NLP application which aims to provide a semantic representation of an input text automatically cannot be silent about non-propositional aspects of meaning, since these are highly influential in meaning construction. As stated in the RRG literature, semantic interpretations only deal with the propositional dimension of meaning, that is, with ‘who did what to whom’, whereas nothing is said about the non-propositional dimension of meaning. For example, the following *wh*-questions not only seek information about a particular item in a given state of affairs but also all of them seem to suggest that the situation the speaker is asking is wrong, that is, the speaker is expressing his concern about the propositional content⁹:

coded in terms of a number of different scripts; (iii) the Onomasticon, where episodic knowledge is described. For a preliminary introduction on how a CLS can be enriched by using artificial intelligence techniques, we refer the reader to Perrián-Pascual and Mairal (2009, 2010b, 2012) which provided some insights in the area of machine translation and cross-linguistic information retrieval.

⁸ See Van Valin (2005, 2013), Van Valin and Mairal (in press), Nolan (2011) and Diedrichsen (2011) for recent discussions on the role of constructions within RRG.

⁹ Following the recommendation of an anonymous reviewer, it is fair to note that not all *Wh*-questions explain disapproval, but English abounds with examples of constructions based on *Wh*-questions that do express disapproval or at least some type of (usually negative) emotional reaction on the part of the speaker:

- (1) Who do you think you are to talk to me like that?
- (2) Where (on earth) have you put my glasses?

- (4) a. Who has been messing up with the bulletin board?
- b. Where have you been all night?
- c. What is the child doing with the carving knife in the kitchen?
- d. When was that order issued?

The semantic representations of these instances should be able to capture this non-propositional dimension of meaning, that is, to search for a type of representational mechanism that allows making explicit whatever is implicit. In connection with this, a lot of relevant work has been done within the Lexical Constructional Model (LCM), a comprehensive model of meaning construction that organizes constructional schemata around four levels of analysis: level-1 or argumental constructions, level-2 or implicative constructions, level-3 or illocutionary constructions and level-4 or discourse constructions (cf. Ruiz de Mendoza & Mairal 2008, Mairal & Ruiz de Mendoza 2009, Ruiz de Mendoza 2013), among others; see also Butler (2009, 2013) for an assessment of the LCM). As shown below, the LCM provides the analytical tools to deal with those aspects of meaning that go beyond grammar and have a long tradition in pragmatics and discourse studies.

In what follows, the primary aim of this chapter is to focus on how constructional schemata are actually dealt with within ARTEMIS. After a brief presentation of the computational architecture of ARTEMIS in Section 2, Section 3 contex-

-
- (3) Why should JÓHN do that? (with stress prominence on "John")
 - (4) What is the child doing RÍGHT now? (with stress prominence on "right")
 - (5) When will she finally ARRÍVE? (with stress prominence of "arrive")
 - (6) Who's been messing with my laptop?
 - (7) What've you been doing (in the kitchen)?
 - (8) Who WÍLL then? (with added stress prominence on "will")
 - (9) What the heck are you talking about?
 - (10) Whenever is she going go learn?

These sentences make use of various linguistic resources to signal the speaker's emotional reaction:

- a) Prosodic marking
- b) Hedges like "ever", "the heck", "on earth"
- c) Specialized constructional forms, like "What's X Doing Y?" and "Who's been V-ing Y?", "(non subject Wh-'s X been V-ing Y

Such marking points to level-2 or implicational meaning rather than argument-structure meaning, which corresponds to level 1. The presence of explicit linguistic marking (whether prosodic or grammatical) of the speaker's attitude points to a stable form-meaning association, i. e. to a constructional configuration, which goes beyond level 1.

tualizes the Grammaticon, the linguistic module that stores the inventory of constructional schemata in FunGramKB. Section 4 focuses on the representation and processing of both argument (Level-1 constructions) and idiomatic constructions (Level-2, 3 and 4 constructions), while Section 5 is concerned with the syntactic representation of this typology within the framework of the layered structure of the clause. Finally, Section 6 includes a few concluding remarks and future prospects.

2 ARTEMIS and FunGramKB: a preliminary discussion

It is not the aim of this section to spell out the exact details of the whole architectures of ARTEMIS and FunGramKB, but simply to draw your attention to the fact that these two NLP resources are intended to represent the semantics of an input text by using RRG. This is a major step that should not go unnoticed, since in the emergent field of cross-linguistic information retrieval most projects are based on probabilistic, context-free grammars and follow stochastic approaches. In turn, our proposal is one of the first systems which, given an input text, employs a robust knowledge base to generate a full-fledged CLS to be used in NLP applications requiring language comprehension capabilities. Figure 2 is a simplified illustration of the architecture of FunGramKB, the source from where the ARTEMIS parser extracts all the information for the automatic construction of a semantic representation of an input text.

FunGramKB comprises three major knowledge levels (i. e. lexical, grammatical and conceptual), consisting of several independent but interrelated modules:

- a. Lexical level:
 - a.1. The Lexicon stores morphosyntactic and collocational information about lexical units. The FunGramKB lexical model is not a literal implementation of the RRG lexicon, although some of the major linguistic assumptions of RRG are still preserved.
 - a.2. The Morphicon helps our system to handle cases of inflectional morphology.
- b. Grammatical level:

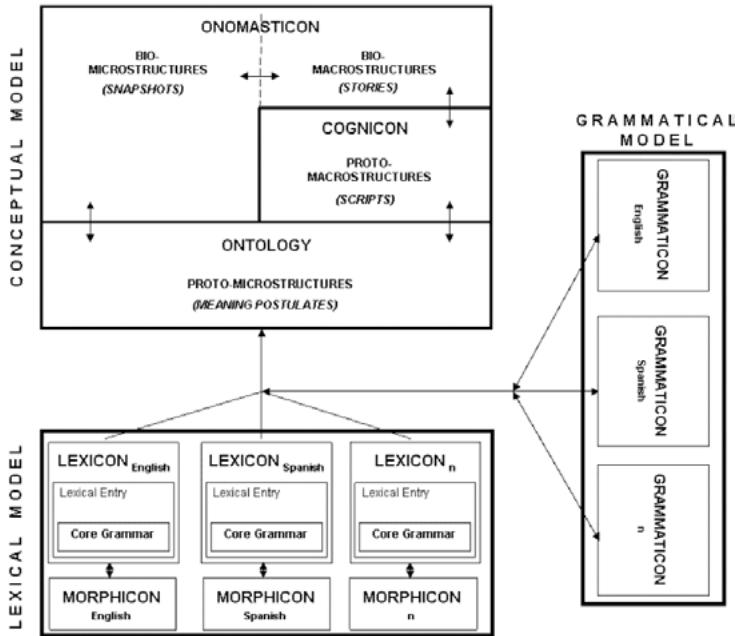


Figure 2: The FunGramKB architecture (reproduced from Perinán-Pascual and Mairal, 2012: 335).

- b.1. The Grammaticicon stores the constructional schemata which help RRG to construct the syntax-semantics linking algorithm. More particularly, the Grammaticicon is composed of several Construc-ticon modules that are inspired in the four levels of the LCM, i. e. argumental, implicational, illocutionary and discursive.
- c. Conceptual level:
 - c.1. The Ontology is presented as a hierarchical catalogue of the concepts that a person has in mind, so here is where semantic knowledge is stored in the form of meaning postulates. The Ontology consists of a general-purpose module (i. e. Core Ontology) and several domain-specific terminological modules (i. e. Satellite Ontologies).

- c.2. The Cognicon stores procedural knowledge by means of scripts, i. e. schemata in which a sequence of stereotypical actions is organised on the basis of temporal continuity.
- c.3. The Onomasticon stores information about instances of entities and events, such as Bill Gates or 9/11. This module stores two different types of schemata (i. e. snapshots and stories), since instances can be portrayed synchronically or diachronically.

In this part, we will just highlight those theoretical issues which are directly related to the central aim of this paper, that is, the Grammaticon and the computational treatment of constructional schemata. However, it is unavoidable to make at least a cursory reference to the Lexicon given the interaction between the two components.¹⁰ In the FunGramKB Lexicon, each lexical entry includes the following information (cf. Mairal & Perrián-Pascual 2009):

- Basic: headword, index, and language.
- Morphosyntax: graphical variant, abbreviation, phrase constituents, category, number, gender, countability, degree, adjectival position, verb paradigm and constraints, and pronominalization.
- Core grammar: aktionsart, lexical template (variables, macrorole assignment and thematic frame mapping) and constructions.
- Miscellaneous: dialect, style, domain, example and translation.

In the case of verbal predicates, the most important lexical component is the core grammar, which contains those attributes whose values allow the system to build the basic CLS of verbs automatically. Figure 3 is a representation of these attributes for the predicate 'break'.

At this stage of the paper, what is noteworthy is the fact that a lexical entry contains pointers to the whole repertoire of constructions a given verb can occur in. In addition to the constructions derived from the Grammaticon, every verb in the Lexicon is provided with one and only one kernel Construction, which is built on the basis of the knowledge in the core grammar, primarily the aktionsart and the lexical template (i. e. variables, thematic frame mapping and macrorole

¹⁰ In close connection with this statement, we do think that both projectionist and constructivist approaches are correct since both constructions and lexical entries are essential for constructing the propositional dimension of meaning. As a matter of fact, we postulate a complementary relationship between the two since it is often the case that it is impossible to account for the semantic representation of an input text without recurring to the Grammaticon and the information contained therein.

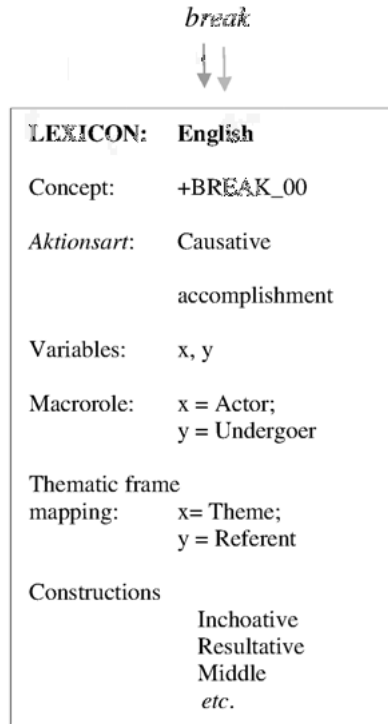


Figure 3: Simplified representation of ‘break’.

assignment). Depending on the number of variables in the lexical template, the verb will typically occur in a Kernel-1, Kernel-2 or Kernel-3 Construction. For instance, the system can directly derive the Kernel-2 Construction from the core grammar of *break*:

(5) [John broke the window]_{Kernel-2}

So, every lexical entry in the Lexicon has one basic kernel structure together with pointers to the rest of constructions, which will be stored in the Grammaticon. The Grammaticon, which is directly linked to the Lexicon in terms of what has been termed the lexical-grammatical interface, stores the inventory of constructional schemata to which words in the Lexicon are linked. As a matter of fact, a given construction can be licensed in a particular language if and only if there is at least one entry in the Lexicon which contains a pointer to that construction. In

the case of *break*, this predicate can occur with a resultative construction since there is a pointer in the Lexicon that shows this connection:

(6) [[John broke the window]_{Kernel-2} into pieces]_{Transitive-Resultative}

The FunGramKB Grammaticon, unlike the RRG constructional schemata, stores both argumental and non-argumental (or idiomatic) constructional schemata, although it is fair to say that the computational implementation of idiomatic constructions is still preliminary.

In sum, the linguistic level in FunGramKB includes a Lexicon, which is connected to the Ontology, and a Grammaticon. From this architecture, Perinián-Pascual (2013b) and Perinián-Pascual & Arcas (2014) began to elaborate ARTEMIS, where the Ontology, the Lexicon and the Grammaticon are the sources of information for the elaboration of the grammar rules through the FunGramKB Grammar Development Environment. Here is a general representation of the architecture of ARTEMIS:

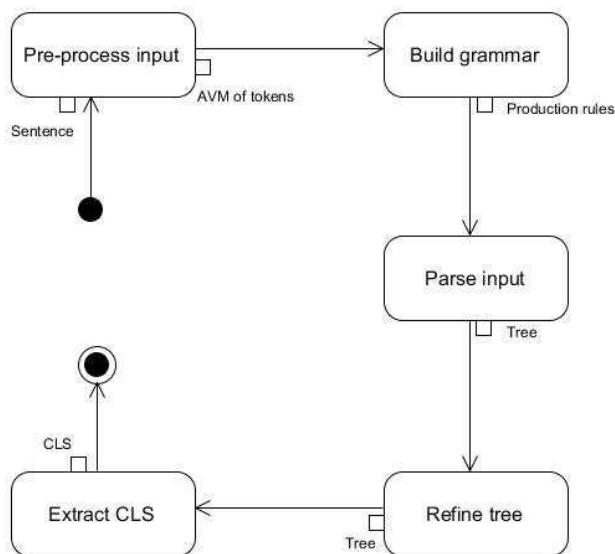


Figure 4: The architecture of ARTEMIS (Perinián-Pascual & Arcas, 2014).

For the purposes of this paper, we focus on the ‘Build grammar’ task, which implies three types of feature-based production rules, i. e. lexical, constructional

and syntactic. These rules automatically generate a parse tree, from which a fully-specified semantic representation is constructed, that is, the CLS acting as a sort of interlingua. While syntactic rules are concerned with the construction of the layered structure of the clause (LSC), constructional and lexical rules specify the properties of constructional schemata and lexical entries respectively. Unlike syntactic rules, both constructional and lexical rules are generated automatically, so we can affirm that a significant amount of grammar rules is dynamically built through ARTEMIS. More particularly, constructional rules are generated with the aid of the Lexicon and the Grammaticon (i. e. the core grammar of the verb together with all its constructional schemata), and lexical rules mainly require the Lexicon and the Ontology.

Continuing with the Grammar Builder module, this is inspired in the paradigm of constraint-based grammars, or also called unification grammars. Each grammatical unit is described in terms of an attribute value matrix (AVM), which includes a number of features that can be eventually merged by means of unification. Hence, phrase structure rules are not longer used but rather parsing will be guided by a number of satisfaction constraints, which are responsible for determining structural preference and semantic plausibility. Both lexical entries and constructional schemata are described in terms of AVMs, each including a number of descriptors and constraints (cf. below). Figure 5 is an example of the representation of the lexical entry for the predicate ‘pound’.

Rather than going into the formal expression of these rules — see Perinián-Pascual & Arcas (2014) for a detailed description, we would like to concentrate on the Grammaticon and the representation of constructional schemata, both argumental and idiomatic, which provide the material for the automatic generation of constructional rules. Additionally, we should comment on how the RRG LSC is actually enhanced.

3 The FunGramKB Grammaticon

The FunGramKB Grammaticon stores an inventory of constructional schemata, both argumental and non-argumental, which are language specific. We maintain that constructional schemata play a fundamental role in propositional meaning constructions, since it is very often the case that it is impossible to account for the semantic structure of an input text by looking only at its argument structure

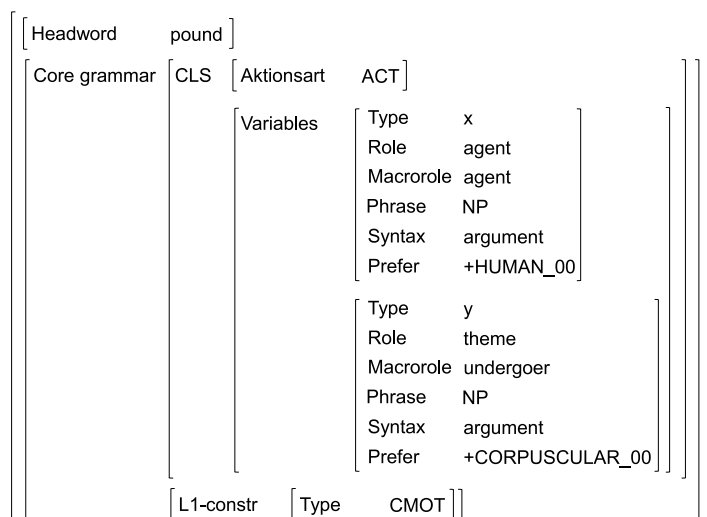


Figure 5: AVM of the lexical entry for the predicate *pound*.

in the Lexicon. This overrides any further debate on whether a theory should be projectionist or constructivist since we maintain that both perspectives are correct (cf. Mairal & González 2010, Perinián-Pascual 2013b, and footnote 10 above). The Lexicon and the Grammaticon are fully interconnected in such a way that every lexical entry in the Lexicon should have pointers to all those constructions in which this predicate participates. So, we vindicate the explanatory potential of both lexical entries and constructions.

Constructional schemata within ARTEMIS are conceived as machine-tractable representations of constructions. It is important to emphasize the fact that constructional schemata have a semantic load, i.e. they are meaning-bearing devices, regardless of whether their semantics is coded in the CLS or in the COREL scheme¹¹. Unlike RRG, and we would venture to claim most construction-based models, FunGramKB, following work within the LCM, works with a four-level typology of schemata, thus aspiring to provide an explanatory scheme for both the propositional and the non-propositional dimension of meaning. Following the work of Ruiz de Mendoza & Mairal (2008) and Mairal & Ruiz de Mendoza

¹¹ As discussed later in sections 4.1. and 4.2., each constructional schema consists of two types of descriptors, a CLS, a syntactically-oriented notational formalism which serves to build a bridge between the linguistic realization of an input text and the conceptual realm, and the COREL scheme, which is a language-independent conceptual representation of the semantics of the text.

(2009), here is a brief summary of the description of the four levels of analysis, which are thought to account for the way meaning constructions processes take place at all descriptive levels, including those meaning dimensions that have a long tradition in pragmatics and discourse:

- Level-1, or argumental level, is concerned with the type of argument constructions distinguished in standard constructional grammar approaches (e.g. Goldberg 1995, 2006), although ARTEMIS differs substantially from these approaches. For example, the ditransitive, the resultative or the caused motion construction belong to this level-1 Constructicon.
- Level-2, or implicational level, accounts for the way meaning can be obtained on the basis of a combination of degrees of pragmatically guided and linguistically guided situation-based low-level inferencing. The former has been termed implicature while the latter - the linguistically guided inferencing - has been called presupposition¹². Recall the instances above: the What is X doing Y construction (e.g. *What is the child doing in the kitchen?*, cf. Kay & Fillmore 1999), which suggests that the situation the speaker is asking about is wrong; Don't you X me!, (e.g. *Don't you honey me!*), which is used to indicate annoyance at the addressee unawareness that (s)he has done something wrong; Who's Been VP Y? (e.g. *Who's been messing up the*

¹² The best way to understand presuppositions is to think of them as covert assertions that naturally follow from the constructional properties of an utterance. For example, if someone says "I'm sorry that your cat died", we are certain that the speaker thinks that the hearer's cat has died on the basis of the constructional properties of "Be Sorry That X". Some scholars have argued that "be sorry" is a factive predicate, but we should bear in mind that there are uses of "(be) sorry" that do not convey any clear presupposition that something is or has been the case:

- She will never be sorry about anything.
- Better safe than sorry.
- How could I be sorry?

The predicate "(be) sorry" has to be part of the right constructional configuration to become a factive predicate:

- I am sorry that you didn't pass your exam [the speaker thinks the hearer didn't pass the exam]
- I am sorry to be so harsh [the speaker thinks he's being harsh]
- She is sorry about the incident [the speaker thinks that there was an incident]
- You will be sorry that you retire [the speaker thinks the hearer is going to retire]

These examples suggest that "(be) sorry" develops factive predicate properties if it has a specific (rather than a generic) complement, which is a constructional issue.

bulletin board?) which carries the meaning implication that someone has done something that irritates and bothers the speaker.¹³

- Level-3, or illocutionary level, deals with the inferential and constructional mechanisms involved in the derivation of speech act meaning. The semantic component is made up of high-level situational cognitive models usually corresponding to traditional speech act categories (e. g. requesting, offering or apologizing) (cf. Ruiz de Mendoza & Baicchi 2007). Consider the following examples: You shall have X (e. g. *You shall have a bicycle*), which is used to make promises; Would you mind if I X? (e. g. *Would you mind if I sat next to you?*), which is used to ask for permission; I won't X (e. g. *I won't give up*), which involves a refusal.
- Level-4, or discourse level which deals with cohesion and coherence phenomena from the point of view of the activity of discourse constructions based on high-level non-situational cognitive models like reason-result, cause-effect, condition-consequence. For example, X let alone Y (Fillmore, Kay & O'Connor 1988) (e. g. *I won't eat that garbage, let alone pay for it*), which is used to refer to two unlikely states of affairs where the second one is less likely to be the case than the first; Just because X doesn't mean Y: *Just because we don't talk doesn't mean I don't think about you*) (cf. Bender & Kathol 2001), which sets up an evidence conclusion relationship according to which Y does not necessarily follow from X¹⁴.

While Level-1 deals with the propositional dimension of meaning, levels 2, 3 and 4 provide the analytical tools to account for the non-propositional dimension of meaning¹⁵. In relation to the format, both argumental and idiomatic constructions are expressed in terms of AVMs, which are used to express the descriptors

¹³ Then, presupposition arises from level 1. Implicature from level 2. Implicature is an inferred assumption that follows from the application of a reasoning schema of the premise-conclusion kind and it is context-dependent. Presupposition is a context-independent covert assumption that is derived from the properties of a construction as a necessary implication of what is said. For this reason, it remains constant under negation ("I [don't] regret stepping on your toe" presupposes that the speaker thinks that he stepped on the hearer's toe), although it is cancellable by making explicit further information: "I don't regret stepping on your toe because I didn't step on your toe".

¹⁴ All the examples plus the accompanying comments have been taken from Mairal & Ruiz de Mendoza (2009) and Ruiz de Mendoza (2013).

¹⁵ Following Ruiz de Mendoza (2013), idiomatic constructions in the LCM are classified according to the following two parameters: (i) their degree of fixity, i. e. whether they are fully fixed or they contain some variable elements; (ii) their meaning function, which is essential to determine the level of description that they belong to, that is, whether they belong to levels 2, 3 or 4.

and their corresponding constraints, where the latter license compositionality with other constructs or constructions by means of unification.

4 The representation of constructional schemata

While the initial phase of the FunGramKB Grammaticon has been devoted to the analysis and representation of argument constructions, in the last few months the second phase of the project has focused on the representation of idiomatic constructions (Levels 2, 3 and 4). This has been indeed a major challenge given the complexity of providing a machine tractable framework to codify the pervasive nature of non-propositional meaning. While Section 4.1 summarizes previous work on argument constructional schemata, Sections 4.2 and 5 offer a first approximation to the inventory of idiomatic constructions.

4.1 Argument (Level-1) constructional schemata

Argument constructions have been the first focus of the FunGramKB Grammaticon. As a matter of fact, Perinián-Pascual (2013b) and Perinián-Pascual & Arcas (2014) evaluate ARTEMIS within the framework of various constructional schemata, i.e. the caused motion and the resultative. Moreover, Van Valin & Mairal (in press) compare the RRG formalism and the FunGramKB formalism and contend that both are compatible to the extent that FunGramKB schemata can enrich the insufficient semantic description of RRG schemata. All in all, ARTEMIS, which retrieves information from the Grammaticon to generate the construction rules that form part of the Grammatical Development Environment, seems to function fairly well within a conceptual framework like that provided in FunGramKB.

Summarizing a bit, argument constructions, like lexical entries in the Lexicon, are represented in FunGramKB by means of AVMs. Let us use the format of the intransitive resultative construction for illustration purposes. The FunGramKB Grammaticon provides the interface shown in Figure 6. This interface is divided into two blocks of information:

- a) The CLS which includes the following items:
 - a. Type of aktionsart: accomplishment.
 - b. Number and type of variables: *y* and *w*.

a) CLS:
 Aktionartsart: ACC [help] Variables: y,w
 Variable: w
 Thematic role: Result
 Macrorole:
 Phrases: ADJP, PP [help]
 Syntax: Nucleus
 Prepositions: into
 Preferences:
 Assign Clear
 Variable W:
 Result
 ADJP, PP
 Nucleus
 into
 b) COREL scheme:
 + (e1: <EVENT> (f1: (e2: +BECOME_00 (x1: y) Theme (x2: w) Attribute)) Result)
 Save Done

Figure 6: Interface of the English intransitive resultative construction.

- c. For each new variable, that is, for that variable that is contributed by the construction, the following information should also be provided:
 - Thematic role.
 - Macrorole status, if any.
 - Phrases: morphosyntactic realization.
 - Syntax: the status of this new constituent in the LSC, i. e. argument or a nucleus.
 - Preposition, if any.
 - Selectional preferences, if any.
- b) The COREL scheme includes a language-independent semantic description: in this case, there is an event and as a result the *y* participant comes to get a new state.

Note that in the representation in Figure 6, the *y* variable is inherited from the information in the Lexicon, so there is no need to specify this information again. In contrast, the *w* is the argument that is contributed by the resultative construction. Therefore, the different properties of this variable are spelled out. The following AVM includes the information of the FunGramKB editor:

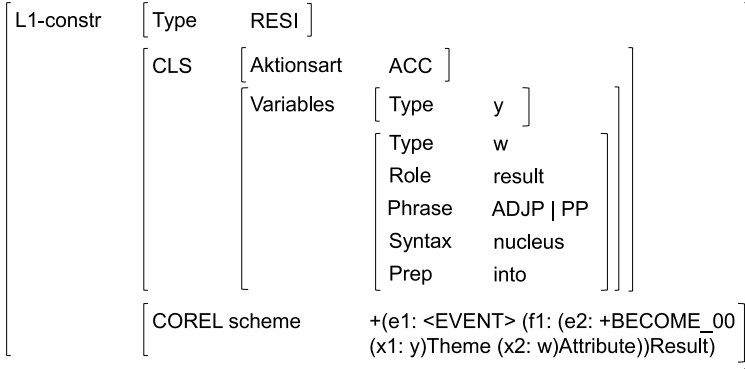


Figure 7: AVM of the English intransitive resultative construction.

From this information FunGramKB is able to automatically generate the following CLS as in (7c) for an input text such as (7a):

- (7) a. The milk froze black in the basement.
 b. <IF DEC <TNS PAST < be-in' (basement, [[do' (milk, [freeze' (milk))]
 CAUSE [BECOME **black'** (milk)]]>>>
 c. (<IF DECL <Tense past <CONSTR-L1 RESI <CONSTR-L1 INCH <AKT ACC
 [+FREEZE_00 (+MILK_00-Referent, +BLACK_00-Result)]
 (+BASEMENT_00-Location) >>>>

What is noteworthy is the fact that two of the operators in (7c), AKT and Constr-L1, are not part of the RRG inventory. Why are they used? What is the motivation behind this proposal? Following Perinián-Pascual (2013a: 219):

(...) the RRG decompositional system turns out to be excessively noisy from a computational view, since the semantic burden of the sentence is not actually carried by the CLS but by its corresponding COREL scheme.

The RRG logical structure in (7b) includes a number of operators, i. e. CAUSE and BECOME, which can be ignored without any loss of information providing that we explicitly state the aktionsart together with the argument pattern headed by the event. Therefore, the CLS in (7c) can be deprived of the RRG skeleton, resulting in a full-fledged formalism which can be effectively employed by computer applications. Moreover, a constructional operator is incorporated (i. e. CONSTR-

L1),¹⁶ which plays a prominent role in the syntax-semantics linkage, as will be discussed in Section 5.

4.2 Idiomatic constructions: the non-propositional dimension of meaning

Idiomatic constructions are a cover term that includes those constructions ranging from implicational to discourse constructions. Unlike argument constructions (i. e. Level-1), idiomatic constructions consist of fixed and variable elements that can be parametrized in different degrees. Let us consider the following schemata:

Level-2: What is X doing Y?

Double Be

Level-3: I won't X

I shall X

I wonder if you could X

Level-4: Just because X doesn't mean Y

X on condition that Y

(*You can have the day off tomorrow on condition that you work on Saturday*)

In the case of idiomatic constructions, we shall distinguish two types of units: (i) fixed (or non-parametrizable) elements and (ii) parametrizable (or variable) elements. In our application of this distinction to Fillmore and Kay's treatment of *What's X Doing Y?*, the lexical unit *doing* would be distinguished as a non-parametrizable element, while the X and Y elements are highly parametrizable, since they admit a large amount of variability. A similar example is the so-called *Double Be* (or *copula doubling*) construction (e. g. *The thing is, is that he didn't tell the truth*). McConwell (1988), Tuggy (1996), and Massam (1999) have studied the details of this construction, which serves to call the hearer's attention to a given situation while asserting its truthfulness or relevance. It usually takes the configuration *X is, is Y*, where X, which is marked by a high tone, is the topic and Y, which takes a low tone, is the focus. While Y is a relatively unconstrained element (it can be realized by any that-clause), there is a fairly limited range of

¹⁶ Indeed, every argumental construction is embodied in a constructional operator whose scope is the core.

options for X, normally *the thing, the problem, the question, what I mean*, and *what happens*. The same applies to the level-3 constructions, where the *I won't* element is fixed while the X is subject to parametrization. Finally, discourse constructions are subject to the same pattern: in the case of X on condition that Y, X and Y can represent any clause while the element – *on condition that* – is fixed (see Ruiz de Mendoza & Mairal 2008 for an extensive discussion of fixed and variable elements in the different types of constructions).

Unlike argument constructions, idiomatic constructions only serve to embed some constructional operator into the CLS; in other words, they do not alter the CLS of the text but only extend their COREL scheme. As shown in the following interfaces, only the COREL scheme is relevant, while the box containing the different realizations serves for ARTEMIS to identify through pattern matching the type of constructional type. Here is the interface of the illocutionary construction *Requesting*, an example of Level-3:

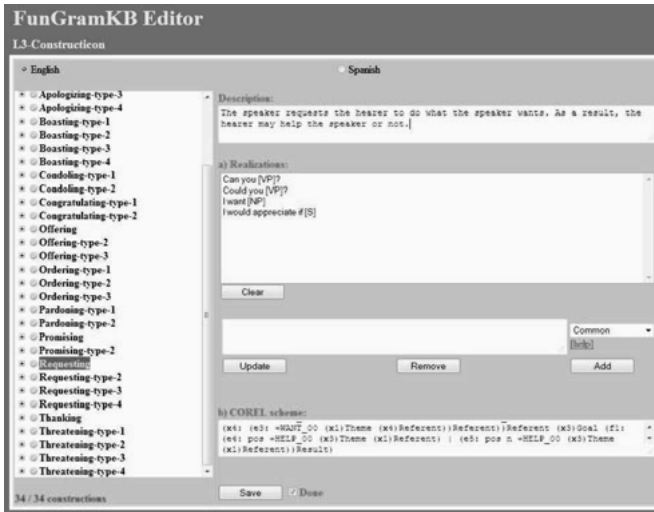


Figure 8: Interface of the Level-3 construction *Requesting*.

Unlike argument (or Level-1) constructions, Level-2, Level-3 and Level-4 constructions do not include a CLS. Only the possible realizations are indicated in terms of fixed and variable elements, usually signaled by means of X, Y and Z. These variables will be filled in by items in the Lexicon or by a Level-1 construction (cf. below). Idiomatic constructions add a non-propositional dimension of

meaning, something which is represented in terms of a COREL scheme, that is, a semantic description of this dimension of meaning:

- (8) $+(e1: +REQUEST_01 (x1: <SPEAKER>)Theme (x2: (e2: +DO_00 (x3: <HEARER>)Theme (x4: (e3: +WANT_00 (x1)Theme (x4)Referent)) Referent)) Referent (x3)Goal (f1: (e4: pos +HELP_00 (x3)Theme (x1)Referent) | (e5: pos n +HELP_00 (x3)Theme (x1)Referent)))Result)$

The speaker requests the hearer to do what the speaker wants. As a result, the hearer may help the speaker or not.

The different realizations that are included in the box ‘Realizations’ carry this semantic load written in COREL (see footnote 5), which can be translated in natural language as follows: a speaker says something to a hearer with the purpose that the hearer gives something to the speaker because the speaker needs whatever is requested. A similar example of a Level-3 construction is that which includes those constructions which express the illocutionary act of ‘promising’. As noted in Figure 9, the format will be the same: first, the different possible realizations are included and second a semantic description in terms of COREL is provided. In this case, the speaker says something (x2) to a hearer and whatever the speaker says refers to something that the speaker will do in the future.

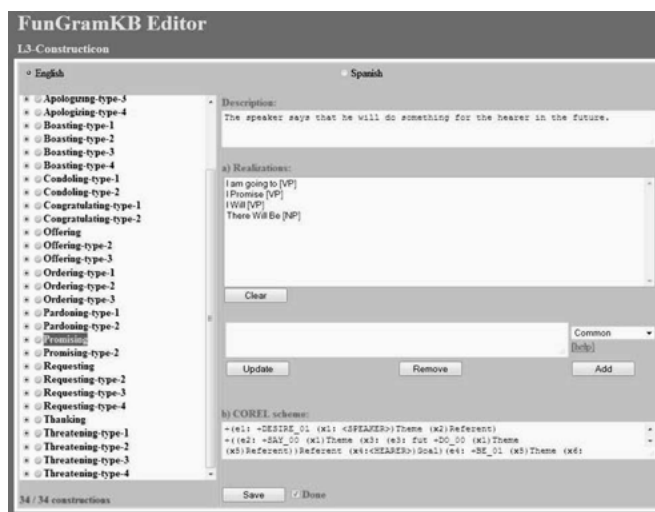


Figure 9: Interface of the Level-3 construction *Promising*.

Finally, we shall like to include the analysis of the *What is X doing Y construction?* given that this is a good example to show that it is possible to arrive at a fine-grained semantic analysis via the COREL scheme. Here are a few examples:

- (9) a. What's the child doing?
 b. What's the child doing in the kitchen?
 c. What's the child doing in the kitchen with the carving knife?

As discussed in Mairal and Ruiz de Mendoza (2009) and noted above, this construction and its variants (e.g. *Who's -ing X?*) seem to convey the idea that there is something wrong about the situation described; this value readily cues for a complaint reading. However, what is noteworthy at this state is the fact that the variable element Y is subject to different degrees of elaboration, that is, this can range from a simple prepositional phrase (e.g. *in the kitchen* as in 9a) to two prepositional phrases (9b) or even more (e.g. *what is the child doing in the kitchen with a carving knife hitting the new tiles on the wall?*). Interestingly enough, this arrangement is not accidental but semantically motivated by the fact that the greater the elaboration of the Y element, the greater the idea that something is wrong, and consequently, the complaint interpretation becomes greater. It seems that the speaker, in being able to supply so much information about the propositional content, already knows the answer to his/her question. Undoubtedly, this implicated meaning seems to be crucial for a comprehensive semantic representation of this input text and therefore any NLP application should be sensitive to it. If we compare the representation of this input text in terms of a standard logical structure (10b), this proves itself insufficient to capture this implicated meaning, while this is not the case with the representation in (10c):

- (10) a. What's Tom doing in the kitchen with the carving knife?
 b. <IF INT <TNS PRES <ASP PROGR <be-in' (kitchen, [[do' (Tom, [do' (Tom, what) ^ use' (Tom, carving knife)]])]]
 c. +(e1: pro pres +DO_00 (x1: %TOM_00) Theme (x2: ?) Referent (f1: +KITCHEN_00) Location (f2: \$CARVING_KNIFE_00) Instrument) +(e2: +FEEL_00 (x1: e1) Agent (x2: <SPEAKER>) Theme (x3: +ANGRY_00) Attribute)

The speaker wants to know what Tom is doing in the kitchen with the carving knife; Tom's action makes the speaker feel angry as well.

When the utterance matches the syntactic pattern *What is NP doing Y?* or any of its related variants, then the FunGramKB Grammaticon will extend the COREL schema, as can be seen in (10c), since this Level-2 *Call for Redressive Action* construction contributes with the following meaning:

- (11) +(e1: +FEEL_00 (x1: <EVENT>) Agent (x2: <SPEAKER>) Theme (x3: +ANGRY_00)Attribute)
The event makes the speaker feel angry.

Here, the conceptual metavariable <SPEAKER> should be replaced by an instance from the Onomasticon representing the speaker involved in the utterance, e. g. %TOM_00, %MARY_00, etc., and the metavariable <EVENT> stands for the eventive causer that makes the speaker feel angry about this state of affairs.

5 Revisiting the parser

At this stage, we have seen that a CLS is sensitive to the four-level distinctions in the Grammaticon. As a matter of fact, as discussed in Section 4.1, a CLS includes two new operators, AKT and CONSTR-L1. Let us retake the CLS for the input text *the milk froze black in the basement*:

- (12) (<IF DECL <Tense past <CONSTR-L1 RESI <CONSTR-L1 INCH <AKT ACC
 [+FREEZE_00 (+MILK_00-Referent, +BLACK_00-Result)]
 (+BASEMENT_00-Location) >>>>>

In this case, the lexical entry for *freeze* in the Lexicon includes a structure with two arguments that designate a causative accomplishment (e. g. *Peter froze the milk*), i. e. a Kernel-2 structure¹⁷. However, this predicate now occurs in the context of an inchoative construction and a resultative construction. So, it seems as if these two constructions are modeling the output lexical entry for *freeze*. In other words, ARTEMIS will have to go to the Grammaticon and retrieve that informa-

¹⁷ As noted in Perinián-Pascual & Arcas (2014), kernel constructions are the only type of constructions which are not formalised in the Grammaticon, but they are modeled within the lexical entry of the verb: Kernel-1, Kernel-2 and Kernel-3 are distinguished depending on the number of arguments of the verb in the Lexicon. On the contrary, the L1-constructions, e. g. the inchoative and the resultative, come from the Grammaticon. Moreover, note as a methodological decision we understand that the causative use is regarded as basic although we are aware that this should be backed up with some empirical data, something which is out of the scope of this paper.

tion which is needed to generate the correct representation. This means that the AVM for the inchoative construction will unify with the lexical entry, the output of which will then unify with the AVM for the resultative construction¹⁸. Hence, ARTEMIS needs a label to identify each of these constructs that are functional in the semantic representation of an input text. Moreover, we could affirm that the very same notion of constructional meaning seems to be a universal distinction regardless of the fact that the constructional inventory is language specific. If this is so, and in line with Van Valin's (2005: 3) corollary that "*a theory of clause structure should capture all of the universal features of clauses*", the construction as a universal category should be part of the LSC. Therefore, the clause is configured now as one or more argumental constructions (L1-CONSTRUCTION) which are recursively arranged:

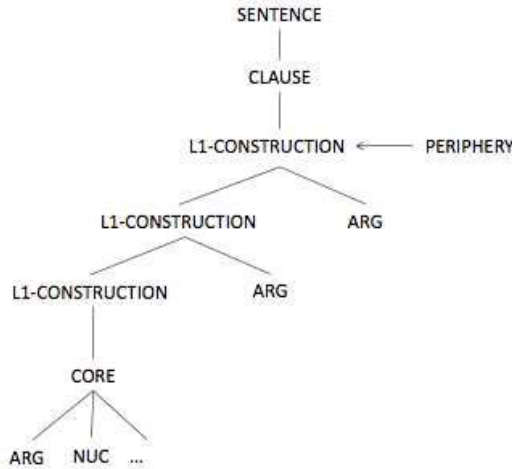


Figure 10: A new look at the LSC (Perrián-Pascual, 2013b).

The innermost construction introduces the core, which can be modeled by other L1-constructions adding a new argument. Unlike idiomatic constructions, Level-1 constructions can occur more than once within the same clause. For example, the inchoative, the resultative and the caused motion constructions are present in the following instance:

¹⁸ This process of unification is similar to that of coercion as stated in most construction grammar approaches (cf. Michaelis 2003): constructional meaning (i.e. a constructional AVM) always wins over lexical meaning (i.e. a lexical AVM).

- (13) a. John kicked the ball flat out of the stadium.
 b. [[[John kicked the ball]_{Kernel-2} flat]_{Transitive-Resultative}
 out of the stadium]_{Caused-Motion}

So, the lexical entry for *kick* is a Kernel-2 structure, which is further modeled by the presence of two argument (Level-1) constructions that are retrieved from the Grammaticon.

However, ARTEMIS is also sensitive to non-propositional meaning as encoded in Level-2, Level-3 and Level-4 constructional schemata. In much the same way as was the case for Level-1 argument constructions, the LSC also contains these new distinctions. Recall that in the case of idiomatic constructions, each constructional node (i. e. L2, L3 and L4) consists of a fixed and a variable element, which is subject to parametrization: the fixed element will be represented under a node provisionally termed CLM (Constructional Level Marker), while the variable elements will be broken down into the predicate and its potential arguments under the clause node. Here is the enhanced representation of the LSC:

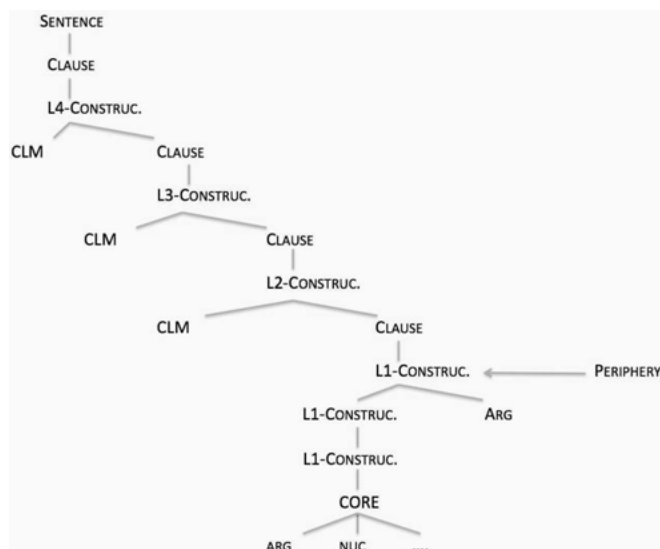


Figure 11: Enhanced model of LSC (refined tree).

This new format of the LSC identifies that constructional schema which is most salient, most prominent, whether be these a Level-2, Level-3 or Level-4. Once

identified, the process will go to Level-1 and/or to the Lexicon to saturate those elements which are not fixed. For example, let us consider the following instances, whose LSC representations are presented in Figures 12 and 13:

- (14) a. I wonder if you could get a copy:
 [Level-3 Constructicon \rightarrow Lexicon]
 b. I won't eat that garbage let alone pay for it:
 [Level-4 Constructicon \rightarrow Level-3 Constructicon \rightarrow Lexicon]

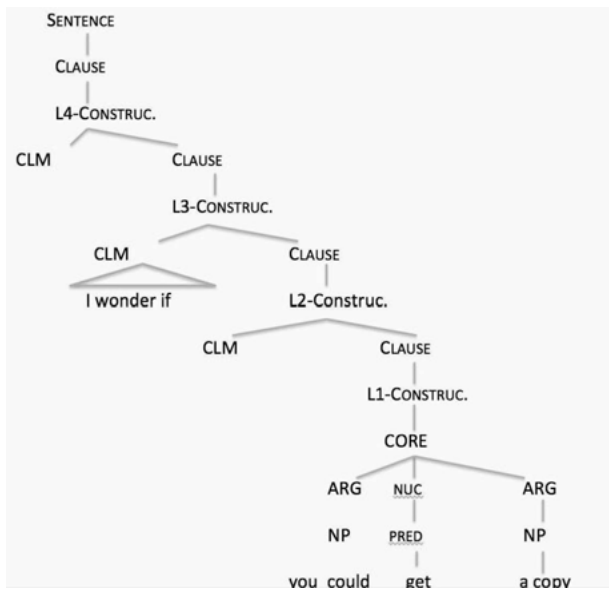


Figure 12: Enhanced LSC of 'I wonder if you could get a copy'.

In (14a), ARTEMIS identifies, through pattern matching, an instance of a Level-3 constructional schema (i. e. *I wonder if you could X*), from which the engine will employ information from the Lexicon to saturate the rest of the elements. In (14b), ARTEMIS, using the same pattern matching technique, will identify an instance of a Level-4 construction (i. e. *X let alone Y*), from which the processor will go to each of the two clauses; in the first case, *I won't eat that garbage*, ARTEMIS will find the realization *I won't X* in the Level-3 Grammaticon, while the remaining part will be saturated in the Lexicon.

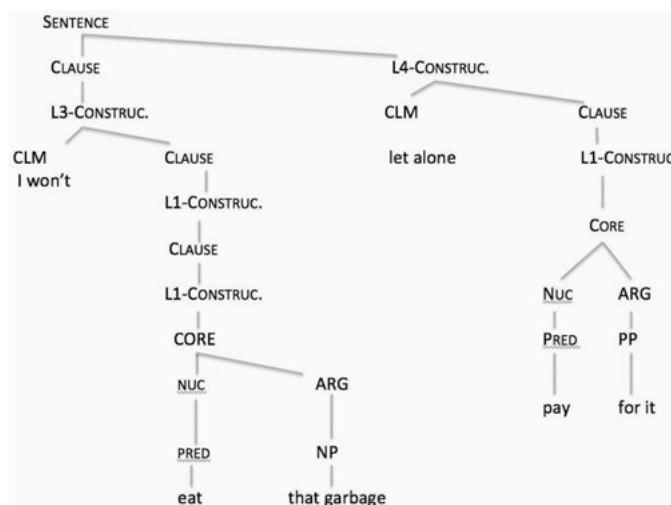


Figure 13: Enhanced LSC of 'I won't eat that garbage let alone pay for it'.

This type of processing imposes a number of restrictions:

- Idiomatic constructions (Level-2, Level-3 and Level-4) have their own unique features which are not shared by other level constructions: a given lexical item cannot be shared by other idiomatic constructions in the same sentence. For example, a lexical item cannot activate both a Level-2 construction and a Level-3 construction in the same sentence.
- Every sentence must have at least an argumental construction, that is, AR-TEMIS will always visit Level-1. In other words, there must be at least one instantiation of a Level-1 construction.
- There can be only one instantiation of the same non-propositional constructional level, i.e. if there is a level-2 instantiation, there cannot be another instantiation of the same level. The same applied to Levels 3 and 4. However, as noted above, this is not the case with Level-1, where more than one instantiation of different argumental constructions can occur within the same text.

Finally, let us consider the following instance and its representation within the framework of the new format of the LSC:

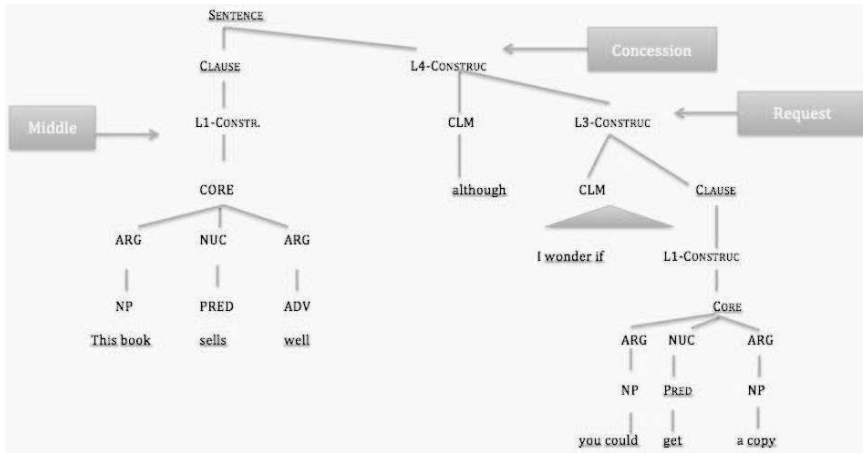


Figure 14: Enhanced LSC of 'This book sells well although I wonder if you could get a copy'.

Firstly, ARTEMIS identifies that this is an instance of a Level-4 construction, given that there is a realization with the format *X although Y*. From there, the processor analyzes each of the two clauses: in the first case, there are no idiomatic constructions but just the presence of a Level-1 construction, i. e. the middle. This is part of the Level-1 Grammaticon and thus it is saturated there with the help of the Lexicon. As for the second clause, the processor finds an exact match with the Level-3 constructional schema *I wonder if you could X*, which is a realization of the illocutionary construction Requesting. The fixed element goes under the node CLM, while the rest is saturated in the Lexicon as a kernel structure. This is an incomplete representation since nodes in the parse tree are represented by means of feature structures and not purely syntactic features.

6 Conclusions

This paper is concerned with the representation of constructional schemata within the framework of ARTEMIS, one of the first systems which employs a robust knowledge base (FunGramKB) to generate a full-fledged semantic representation of an input text. We show that any computer application designed with the aim of understanding the meaning of a text cannot be silent about the non-propositional dimension of meaning, a facet RRG does not include as part of the theory. In this regard, this paper discusses the way both argument and idiomatic constructions

are formalized and represented within the Grammar Development Environment in ARTEMIS. The research done in the LCM is used for the description and identification of idiomatic constructions, L2, L3 and L4 constructions. Both lexical entries and constructional schemata are represented in terms of AVMs, describing features which can be merged through the unification operation, the output of which is a CLS. Moreover, we discuss the repercussions of these four-level constructional schemata in the parser, that is, in the LSC. A constructional node marking the different constructional levels is part of the enhanced LSC. Unlike argument constructions, idiomatic constructions consist of a fixed element (CLM) and a variable element. Finally, a preliminary discussion is offered in terms of how the parser actually functions. This is an issue, which is related to the psychological adequacy of the model, an aspect that needs further work in the future.

References

- Bender, E. & A. Kathol. 2001. Constructional effects of just because ... doesn't mean ... In *BLS 27*: 13–25. Berkeley: University of California.
- Butler, Ch. S. 2009. The lexical constructional model: genesis, strengths and challenges. In Ch. S. Butler & J. M. Arista (eds.), *Deconstructing constructions*, 117–151. Amsterdam: John Benjamins.
- Butler, Ch. S. 2013. Constructions in the lexical constructional model. In B. Nolan & E. Diedrichsen (eds.), *Linking constructions into functional linguistics: The role of constructions in grammar*, 271–294. Amsterdam: John Benjamins.
- Diedrichsen, E. 2011. The theoretical importance of constructional schemas in RRG. In W. Nakamura (ed.), *New perspectives in role and reference grammar*, 168–197. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Fillmore, Ch. J., P. Kay & M. C. O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64: 501–538.
- Goldberg, A. E. 1995. *A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. New York: Oxford University Press.
- Kay, P. & Ch. J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The 'What's X doing Y' construction. *Language* 75, 1–33.

- Mairal, R. 2012. La arquitectura de una base de conocimiento léxico conceptual: implicaciones lingüísticas. In M. Giammatteo, L. Ferrari & H. Albano (eds.), *Léxico y Sintaxis. Volumen temático de la serie editada po la Sociedad Argentina de Lingüística*, 183–210. Editorial FFyL, UNCuyo: Mendoza.
- Mairal, R. & F. González. 2010. Verbos y construcciones en el espacio cognitivo-funcional del siglo XXI. In J. F. Val Álvaro & M. C. Horno Chéliz (eds.), *La gramática del sentido: Léxico y sintaxis en la encrucijada*, 123–152. Zaragoza: Prensas Universitarias de Zaragoza.
- Mairal, R. & C. Perinán-Pascual. 2009. The anatomy of the lexicon component within the framework of a conceptual knowledge base. *Revista Española de Lingüística Aplicada* 22: 217–244.
- Mairal, R., C. Perinán-Pascual & M. B. Pérez Cabello de Alba. 2012. La representación léxica. Hacia un enfoque ontológico. In R. Mairal, L. Guerrero & C. González (eds.), *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones*, 85–102. Madrid: Akal.
- Mairal, R. & J. F. Ruiz de Mendoza. 2009. Levels of description and explanation in meaning construction. In Ch. S. Butler & J. Martín Arista (eds.), *Deconstructing constructions*, 153–198. Amsterdam: John Benjamins.
- Massam, D. 1999. Thing is constructions: The thing is, is what's the right analysis? *English Language and Linguistics* 3: 335–352.
- McConvell, P. 1988. To be or double be: Current change in the English copula. *Australian Journal of Linguistics* 8: 287–305.
- Michaelis, L. A. 2003. Word meaning, sentence meaning, and syntactic meaning. In H. Cuykens, R. Dirven & J. R. Taylor (eds.), *Cognitive approaches to lexical semantics*, 93–122. Berlin-New York: Mouton de Gruyter.
- Nolan, B. 2011. Meaning construction and grammatical inflection in the layered structure of the Irish word: An RRG account of morphological constructions. In W. Nakamura (ed.), *New perspectives in role and reference grammar*, 64–101. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Perinán-Pascual, C. 2012. The situated common-sense knowledge in FunGramKB. *Review of Cognitive Linguistics* 10 (1): 184–214.
- Perinán-Pascual, C. 2013a. A knowledge-engineering approach to the cognitive categorization of lexical meaning. *VIAL: Vigo International Journal of Applied Linguistics* 10: 85–104.

References

- Periñán-Pascual, C. 2013b. Towards a model of constructional meaning for natural language understanding. In B. Nolan & E. Diedrichsen (eds.), *Linking constructions into functional linguistics: The role of constructions in a functional grammar*, 205–230. Amsterdam: John Benjamins.
- Periñán-Pascual, C. & F. Arcas. 2004. Meaning postulates in a lexico-conceptual knowledge base. In *15th International Workshop on Databases and Expert Systems Applications*, 38–42. Los Alamitos: IEEE.
- Periñán-Pascual, C. & F. Arcas. 2007. Cognitive modules of an NLP knowledge base for language understanding. *Procesamiento del Lenguaje Natural* 39: 197–204.
- Periñán-Pascual, C. & F. Arcas. 2010. Ontological commitments in FunGramKB. *Procesamiento del Lenguaje Natural* 44: 27–34.
- Periñán-Pascual, C. & F. Arcas. 2014. The implementation of the FunGramKB CLS Constructor in ARTEMIS. In B. Nolan & C. Periñán-Pascual (eds.), *Language processing and grammars: The role of functionally oriented computational models*, 165–196. Amsterdam: John Benjamins.
- Periñán-Pascual, C. & R. Mairal. 2009. Bringing role and reference grammar to natural language understanding. *Procesamiento del Lenguaje Natural* 43: 265–273.
- Periñán-Pascual, C. & R. Mairal. 2010a. La Gramática de COREL: un lenguaje de representación conceptual. *Onomazein* 21: 11–45.
- Periñán-Pascual, C. & R. Mairal. 2010b. Enhancing UniArab with FunGramKB. *Procesamiento del Lenguaje Natural* 44: 19–26.
- Periñán-Pascual, C. & R. Mairal. 2011. The coherent methodology in FunGramKB. *Onomazein* 24: 13–33.
- Periñán-Pascual, C. & R. Mairal. 2012. La dimensión computacional de la Gramática del Papel y la Referencia: la estructura lógica conceptual y su aplicación en el procesamiento del lenguaje natural. In R. Mairal, L. Guerrero & C. González (eds.), *El funcionalismo en la teoría lingüística. La Gramática del Papel y la Referencia. Introducción, avances y aplicaciones*, 333–348. Madrid: Akal.
- Ruiz de Mendoza, F. J. 2013. Meaning construction, meaning interpretation and formal expression in the lexical constructional model. In B. Nolan & E. Diedrichsen (eds.), *Linking constructions into functional linguistics: The role of constructions in a functional grammar*, 231–270. Amsterdam: John Benjamins.

- Ruiz de Mendoza, F. J. & A. Baicchi. 2007. Illocutionary constructions: cognitive motivation and linguistic realization. In I. Kecskes & L. Horn (eds.), *Explorations in pragmatics: linguistic, cognitive, and intercultural aspects*, 95–128. Berlin-New York: Mouton de Gruyter.
- Ruiz de Mendoza, F. J. & R. Mairal. 2008. Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model. *Folia Linguistica* 42 (2): 355–400.
- Tuggy, D. 1996. The thing is is that people talk that way. The question is is why? In E. H. Casad (ed.), *Cognitive linguistics in the Redwoods: The expansion of a new paradigm in linguistics*, 713–752. Berlin-New York: Mouton de Gruyter.
- Van Valin, R. D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Van Valin, R. D., Jr. 2013. Lexical representation, co-composition, and linking syntax and semantics. In J. Pustejovsky, P. Bouillon, H. Isahara, K. Kanzaki, & C. Lee (eds.), *Advances in Generative Lexicon theory*, 67–107. Dordrecht: Springer.
- Van Valin, R. D., Jr. & R. J. LaPolla. 1997. *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Van Valin, R., D., Jr. & R. Mairal. in press. Interfacing the lexicon and an ontology in a linking algorithm. In M. Ángeles Gómez, F. Ruiz de Mendoza & F. González-García (eds.), *Form and function in language: functional, cognitive and applied perspectives. Essays in honour of Christopher S. Butler*. Amsterdam: John Benjamins.

Authors

Ricardo Mairal

UNED

rmairal@flog.uned.es

Carlos Perrián-Pascual

Universidad Politécnica de Valencia

jopepas3@upv.es