

[Ruiz de Mendoza, Francisco & Mairal Usón, Ricardo. 2007. “Challenging systems of lexical representation”. *Journal of English Studies* 4; University of La Rioja; vol. in honor of Carmelo Cunchillos; in preparation]

Challenging systems of lexical representation¹

Francisco José Ruiz de Mendoza Ibáñez
University of La Rioja
francisco.ruiz@dfm.unirioja.es

Ricardo Mairal Usón
UNED, Madrid
rmairal@flog.uned.es

1. Introduction

One of the most ambitious enterprises in linguistic research is to formulate a unified theory that provides an elegant description of the anatomy of the lexicon and more particularly of the different theoretical aspects of lexical meaning construction. Since the early 1980s the design of the lexical component of grammatical description has become a central concern in various approaches to language (ranging from formal accounts to functional and cognitive models, as will be discussed below). Lexical description has further been shown to have important implications not only for linguistic theory proper (e.g. the development of a bidirectional linking algorithm from semantics to syntax and from syntax to semantics), but also for language processing, robotics, word computation, and the design of intelligent reasoning algorithms within the context of the semantic web, just to name a few.

We could affirm that the lexicon has been and still is a central concern for linguists, psychologists, neurologists, computer scientists, and in general for scholars working on cognitive science. Then, it comes as no surprise that most linguistic models, formal, functional and cognitive have endeavored to formulate proposals that concern the overall architecture of the lexicon and more particularly the type and amount of information that a lexical entry should contain². As a result, a significant amount of lexical representation theories have been posited from different scientific angles and theoretical perspectives with a view to developing applications both in linguistic theory and in cognitive science.

¹ Financial support for this research has been provided by the DGI, Spanish Ministry of Education and Science, grants HUM2004-05947-C02-01/FILO and HUM2005-02870/FILO. The research has been co-financed through FEDER funds.

² Cognitive approaches to lexical structure essentially differ from functional and formal accounts in two ways. First, the former generally formulate very rich semantic descriptions for lexical entries (e.g. in the form of ‘frames’; see below), which is not the case for the latter, which tend to code only syntactically-relevant information (e.g. Dik, 1997a; Van Valin, 2005; Levin and Rappaport, 2005). Second, unlike functional and formal accounts, cognitive approaches do not regard all morphosyntactic information as predictable from the argument structure of a predicate. The second difference has in fact been one of the hallmarks that have fragmented linguistic models into the so-called ‘projectionist’ and ‘construction-based’ approaches. We refer the reader to Levin and Rappaport (2005), González-García and Butler (2006) for a very thorough discussion of the actual scope of lexical entries in linguistic theory.

Within the context of this panlexicalist orientation, this paper aims to present an overview of some of the most relevant heuristic parameters that have been used for the organization of the lexicon in a representative sample of formal, functional and cognitive models. In so doing we additionally aim to offer the reader a critical account of the complexity involved in coming to grips with a comprehensive theory of the lexicon.

A glance at most recent research in lexical theory allows us to identify a number of crucial heuristic parameters that in fact have pervaded much of the debate on lexical theory in general. We will devote a section to each relevant parameter. Simplifying a bit, lexicologists have been forced to make decisions on three issues: (i) the nature of the metalanguage that should be used as part of a lexical representation theory (cf. section 2); (ii) the actual scope of the representation, that is, whether a lexical entry should only capture those aspects of the word that are grammatically relevant or should go beyond that and include richer semantic decompositions (cf. section 3); (iii) the type of formalism (cf. section 4) involved in the description of meaning for the design of robust technological applications (cf. section 5). Finally, in the light of the discussion of these three theoretical issues, we will present a sample model of lexical description³.

2. The nature of the semantic metalanguage

If we want to define the meaning of a predicate, we must decide what (meta) language we should be using. In connection with this, there are two clear strands of research; (i) one where the metalanguage for meaning representation is based on natural language; (ii) another where representations are constructed on the basis of an abstract semantic metalanguage. As examples of the first line of research, we find Dik's (1997a) *predicate frames* and the representations formulated in the *Natural Semantic Metalanguage* (NSM) research program conducted by Wierzbicka and her associates (cf. Goddard and Wierzbicka, 1994, 2002; Wierzbicka, 1996, 1999). The two representations make use of natural language definitional elements without any need to use operators, constants, or variables. Below is the representation of 'mother' as propounded in Wierzbicka (1996: 154-155)

X is *Y*'s mother. =

(a) at one time, before now, *X* was very small

(b) at that time, *Y* was inside *X*

(c) at that time, *Y* was like a part of *X*

(d) because of this, people can think something like this about *X*:

“*X* wants to do good things for *Y*

X doesn't want bad things to happen to *Y*”.

One of the crucial advantages of this representation, and of the NSM approach in general, is that all the units involved in the representation of the meaning of a predicate

³ The kind of description provided in section 4 below is part of a complex model of meaning construction at the levels of core grammar, pragmatics, and discourse structure. Besides providing rich semantic characterizations for lexical entries, the model is capable of producing representations that are ready for syntactic realization. For further information on this account, the *Lexical Constructional Model*, we refer the reader to Ruiz de Mendoza and Mairal (2006, 2007, in press) (see also www.lexicom.es and the references therein).

are typologically based, an essential aspect of the theory, especially if we want our representations to serve equally well cross-linguistically.

In a similar fashion, although far away from providing typologically valid representations, Dik's *predicate frames* specify the type of predicate, the quantitative and the qualitative valency, the selection restrictions imposed on each of the arguments and a meaning definition. Consider the following examples taken from Dik (1997:101):

$$\begin{aligned} & \textit{assassinate} [V] (x_1: \langle \textit{human} \rangle)_{\text{Ag}} (x_2: \langle \textit{human} \rangle)_{\text{Go}} \leftrightarrow \\ & \quad \textit{murder} [V] (x_1)_{\text{Ag}} (x_2)_{\text{Go}} (x_3: \textit{treacherous} [A])_{\text{Manner}} \\ & \textit{murder} [V] (x_1: \langle \textit{human} \rangle)_{\text{Ag}} (x_2: \langle \textit{human} \rangle)_{\text{Go}} \leftrightarrow \\ & \quad \textit{kill} [V] (x_1)_{\text{Ag}} (x_2)_{\text{Go}} (x_3: \textit{intentional} [A])_{\text{Manner}} \\ & \textit{kill} [V] (x_1)_{\text{Ag/Fo}} (x_2: \langle \textit{human} \rangle)_{\text{Go}} \leftrightarrow \\ & \quad \textit{cause} [V] (x_1)_{\text{Ag/Fo}} (e_1: [\textit{die} [V] (x_2)]_{\text{Proc}})_{\text{Go}} \end{aligned}$$

If we consider, for example, the predicate frame for *murder* we note the following: this verb takes two arguments (this is the *quantitative valency* expressed by means of the x_1 and x_2 variables), an Agent and a Goal (the *qualitative valency*); there are selection restrictions for each of the two arguments (in this case the feature 'human'); there is a resulting meaning definition that reads as follows: an agent kills a patient or goal intentionally. What is interesting about this type of approach is that it is capable of accounting for the distinguishing parameters of verbs that belong to the same lexical class. In this regard, consider the codification of verbs of *sleep* as formulated in Martín Mingorance (1995) within the context of this approach to lexical representation:

$$\begin{aligned} & \textit{be sleepy}_v (x_1: + \textit{anim} (x_1))_{\text{Proc}} \\ & \quad \text{def} = [\textit{begin}_v (x_1)_{\text{Proc}} (x_2: [\textit{fall asleep}_v (x_1)_{\text{Proc}}] (x_2))_{\text{Goal}}]_{\text{Process}} \\ & \textit{be drowsy}_v (x_1)_{\text{Proc}} \\ & \quad \text{def} = [\textit{begin}_v (x_1)_{\text{Proc}} (x_2: [\textit{fall asleep}_v (x_1)_{\text{Proc}}] (x_2))_{\text{Goal}}]_{\text{Process}} \\ & \quad (y_1: [\textit{appear}_v (x_1: \textit{calm}_{\text{Adj}} \& \textit{relaxed}_{\text{Adj}} (y_1))_0] (y_1))_{\text{Circumstance}} \end{aligned}$$

These representations give an exact account of the differences between two semantically close verbs such as *be sleepy* and *be drowsy*; while the former just focuses on the process of falling asleep, the latter additionally specifies that while this process takes place, the participant x_1 appears calm and relaxed. These representations are elegant, since they specify non-redundant, syntactically relevant semantic parameters. However, nothing is said of how the argument structure of the predicate (the first part in the representation) interacts with the meaning definition. Thus, the specification consists of two unrelated modules which, as it were, do not speak to each other (cf. Mairal and Faber, 2002). One further problem relates to the inability of these definitions to account for many every-day uses of concepts that would require a broader definitional approach. Thus, the definition for *be sleepy* given above could hardly apply to a sentence like *Brisbane is more than just a sleepy town* (i.e. a town where there is not much activity and excitement). Interpreting this use of 'sleepy' requires an account that is sensitive to metaphorical and metonymic extension, as has been argued for by many cognitive linguists following the seminal proposals in Lakoff (1987). This is not a moot point. For example, consider the following uses of the concept of 'mother':

She mothered him well (i.e. she was her biological mother and took good care of him)

She mothered the baby as if it were her own (i.e. she was not her biological mother but nurtured the baby as a good biological mother would do).

My wife really mothers me! (i.e. she spoils the speaker).

Metaphorical interpretation is more than just a matter of deriving (deviant) meaning from a literal use. It works on the basis of rich semantic characterizations. In our examples above, ‘mother’ is used in connection to biological and emotional attributes that should be part of a meaning definition, even if some of them are “encyclopedic” in nature, i.e. they go beyond the set of so-called “necessary and sufficient” conditions for the concept (in this case the idea that a mother is a female human being that gives birth to at least one child; cf. Taylor, 2003, for detailed treatment of some of the problems of non-encyclopedic accounts of meaning). This is particularly so because these uses of mother as a verbal predicate are possible only to the extent that we can make a metaphorical extension of the concept that is based on the idea that mothers are affectionate, tender, and filled with love and the desire to take care of their children. This observation suggests that it might be desirable to see if the encyclopedic elements of the meaning characterization of a given predicate can influence its morphosyntactic realization, a question that has served as one of the motivating guidelines for what we have called *lexical templates* (see section 4), which are intended to include encyclopedic information that can be bound to elements of syntactic structure. These observations also apply to some extent to the NSM approach. Thus, although Wierzbicka’s definition of ‘mother’ may go some way to dealing with examples like those above, where the notion of ‘taking care of someone’ can somehow be related to ‘X wants to do good things for Y’ and ‘X doesn’t want bad things to happen to Y’, it still falls short of providing us with a characterization that can explain the semantic coherence of many other every-day uses of this notion. Consider, in this respect, *My mother divorced my father when I was six*. The use of ‘mother’ in this sentence is semantically coherent since in our culture it is common for one’s parents to be married, what Lakoff (1987) has aptly called the “marital” model of mother, which is not present in Wierzbicka’s definition. Or consider the metaphorical expression *Necessity is the mother of invention*. Here the meaning of ‘mother’ is metaphorically used in the sense of ‘source’ or ‘origin’, which is possible because we see mothers as a source of life. Again, this idea is absent from the standard NSM definition, which is too reductionistic.

From a different perspective, a significant number of proposals stemming from both formal and functional approaches make use of an abstract semantic metalanguage, thus following the tradition of formal semantics⁴. The new formalisms consist of an event structure representation that specifies the *Aktionsart* type (i.e. the internal temporal properties) of a predicate, a set of constants (or primitives), a set of variable

⁴ Although for space reasons we focus our attention on Rappaport and Levin’s event templates, it is fair to note that Jackendoff’s (1990) *Conceptual Structures*, Van Valin’s (2005) *Logical Structures* (cf. below), Pustejovsky’s (1995) *event representations*, and the various unification-based grammars also make use of some form of abstract semantic metalanguage as part of the meaning representation of a predicate. What these theories differ on is the actual scope and the exact inventory of operators and constants.

elements, and a set of modifiers or operators⁵. A case in point is Rappaport and Levin's (1998:108) event structure templates:

[x ACT <MANNER>]	(activity)
[x <STATE>]	(state)
[BECOME [x <STATE>]]	(achievement)
[x ACT <MANNER>] CAUSE - [BECOME [y <STATE>]]]	(accomplishment)
[x CAUSE[BECOME [y <STATE>]]]	(accomplishment)

For Rappaport and Levin, verbal meaning consists in the association of a constant with a particular lexical semantic template. Lexical semantic templates represent event structure types and constitute a closed set: activity, state, achievement, accomplishment, etc. These linguists maintain that lexical entries for verbal predicates contain an idiosyncratic element, which was originally called the 'constant' and in subsequent research was renamed the 'root' (Levin & Rappaport, 2005: 71). These elements act as modifiers of a predicate (e.g. <MANNER> in activities and accomplishments) or as arguments of predicates (e.g. <STATE> in activities and accomplishments). For example, activity predicates like *run* or *sweep* will be associated with an activity structure template and the constant slot that indicates manner will be filled in thus yielding the following representational format:

[x ACT <RUN>]
[x ACT <SWEEP> y]

One of the advantages of this approach is that we are able to codify the meaning of a meaning by drawing on a metalanguage that gives us a systematic procedure to codify meaning. However, this is done at the cost of ignoring relevant encyclopedic knowledge parameters that are also part of a speaker's lexical competence. An alternative approach would be to maintain the inventory of event structure templates together with the basic ingredients of the metalanguage and enrich this format by integrating relevant encyclopedic information. As noted above, it is this approach that has in fact inspired our notion of lexical template (cf. section 4).

3. The scope of the representation

Undoubtedly, one of the most controversial methodological parameters that have motivated an intense, and sometimes acid, debate is the amount of information a lexical entry should include. Theories differ on this particular issue and range from those that claim that only those aspects of the meaning of word that are syntactically visible should be part of a lexical entry to those that maintain that a lexical entry should also be sensitive to encyclopedic information. We have two extreme positions that move from the more syntactic to the more semantic poles, what we can provisionally term, adopting Jackendoff's terminology, the *syntactico-centric* and the *semantico-centric* perspectives. We will contend that both approaches can be reconciled.

3.1. The syntactico-centric perspective

⁵ For a very complete account of the various proposals that postulate predicate decomposition based on an abstract semantic metalanguage, see Levin and Rappaport (2005: 3.2).

This strand of research gained a lot impetus during the 1980s, when it was shown that the syntactic behavior of a predicate is in large part predictable from its argument structure. As a first approximation, and continuing the pioneering work by Fillmore in the two preceding decades, the semantic properties of a predicate were reduced to a set of semantic roles, a type of approach that came to be known as ‘role-centered’. However, it was soon noted that semantic role lists were not sufficient to account for the full complexity of a predicate meaning (cf. Levin and Rappaport, 2005). As a reaction to the insufficient explanatory capacity shown by thematic role lists, most linguistic models began to develop more articulated systems of lexical representations based on an abstract semantic metalanguage (cf. section 2) that was deeply imbued with the notion of *Aktionsart*. A case in point is the inventory of logical structures postulated in *Role and Reference Grammar* (RRG), as shown in the following table extracted from Van Valin (2005:45):

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), <i>or</i> 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), <i>or</i> BECOME do' (x, [predicate' (x) or (x,y)])	receive	BECOME have' (x,y)
Active accomplishment	do' (x, [predicate'₁' (x, (y))]) & BECOME predicate'₂' (z,x) or (y)	drink	do' (x,[drink' (x,y)]) & BECOME consumed' (y)
Causative accomplishment	α CAUSES β where α, β are LS of any type	kill	[do' (x, \emptyset)] CAUSE [BECOME [dead' (y)]]

Table 1: Inventory of RRG logical structures

RRG uses a decompositional system for representing the semantic and argument structure of verbs and other predicates (their *Logical Structure*, LS). 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. States and activities are primitives, whereas accomplishments, semelfactives and achievements consist of either a state or activity predicate plus a BECOME, SEML and an INGR operator. First, we have to determine the lexical class membership of a predicate by following a number of tests and then use the appropriate template. For example, a predicate like *see*, which is a state predicate, is represented by means of the template designed for states; in the same way, an activity predicate like *run* is formalized by an activity template; and so on. In our view, this type of approach is elegant and precise in its formulation as well as in the expression of the formalism itself. However, elegance is achieved at the expense of sacrificing the full gamut of semantic and pragmatic parameters that should be part of a predicate meaning's decomposition (cf. section 3.2.), a theoretical assumption that we cannot agree on.

Furthermore, as shown in Van Valin (2006), these representations have cross-linguistic validity. However, when one works in a language other than English the question that arises is the following; is the analyst supposed to use English primes to encode the meaning of, say, *sing* in Amele, Xoxoni, German, or Swahili? A possible, though not very felicitous, answer would be: it is necessary to use a language, so why not English? One has the impression that logical structures only “speak English” when it would be ideal if they could be constructed by drawing on a universal inventory of features, which would certainly provide a satisfactory answer to our question. A further controversial issue is to determine where the decompositional chain actually ends; e.g. what criteria should we be using to determine the prime element?, why do we use *sing*, *drink*, *run* or *popped* as putatively semantic primes when we know on the grounds of typological analysis that they are not? This is something that Van Valin and Wilkins (1993) already noted and made a first proposal for semantically decomposing the predicate *remember*. In subsequent work (e.g. Van Valin and LaPolla, 1997), they included an enriched semantic representation for speech act verbs, a line of research that has been continued in work by Mairal and Faber (2007) and incorporated into the *Lexical Constructional Model* (Ruiz de Mendoza and Mairal, 2006, 2007, in press).

3.2. The semantico-centric perspective

Although the elegance of the Logical Structure formalisms is beyond question, there is more in the meaning of a word than what is specified in these constructs. One of the best representatives of this line of research is Frame Semantics and the FrameNet Project⁶. As Fillmore and Atkins (1994:370) write:

Frame semantics [...] begins with the effort to discover and describe the conceptual framework underlying the meaning of a word, and ends with an explanation of the relationships between elements of the conceptual frame and their realizations within the linguistic structures that are grammatically built up around the word".

Frames have been used to encode the semantic properties of a lexical entry and have been described as "specific unified frameworks of knowledge, or coherent schematizations of experience" (Fillmore, 1985:223). Semantic frames are schematic representations of situation types (e.g. ‘buying and selling’, ‘eating’, ‘spying’, etc.) describable in terms of participants and their roles. For example, the ingestion frame consists of an “ingestor” that consumes food, drink, etc. (“ingestibles”) often with the help of an “instrument”. These are “core” elements. Other non-core elements are the “degree” or extent to which ingestibles are consumed, the “manner” of performing the action, the “place”, the “time”, the “purpose”, and the “source”. Associated with the frame there are a number of lexical units (e.g. *consume*, *devour*, *dine*, *feed*, *gobble*, *slurp*) and inheritance relations with other frames (for ingestion, “intentionally affect”, i.e. performing an intentional act that affects a patient).

A frame semantics dictionary needs to specify semantic frame elements and then look for regularities among these and their grammatical realizations. Fillmore and Atkins (1994) give an example by analyzing the 'risk-frame', which has the following elements:

Protagonist [Pr]: the central person in the frame.

⁶ For an updated account of the *FrameNet Project*, see <http://framenet.icsi.berkeley.edu>

Bad [Ba]: the possible bad outcome, or harm.
Decision [De]: the decision that could trigger this.
Goal [Go]: the desired outcome.
Setting [Se]: the situation within which the risk exists.
Possession [Po]: something or someone valued by the Protagonist and endangered in the situation.
Source [So]: something or someone which could cause the harm.

Each of these sentences, in turn, responds to each of the following three schemas:

- Schema A: a path leads to two alternative uncertain futures, one of them being bad.
- Schema B: the protagonist makes a decision which renders him or her vulnerable to some sort of harm.
- Schema C: the same as Schema B, but the protagonist has in mind a desired outcome and he or she is aware of the potentially bad outcome.

This set of analytical tools permits Fillmore and Atkins to explain some of the differences in meaning that we find in sentences like:

Newborn babies run the risk of hypothermia [Pr, Ba]
I had no idea I was risking my life [Pr, Po]
You'll have to calculate the risks involved [Pr, De]
The health risk from apples is 'minuscule' [Pr, So]
Living in San Francisco is a risk [Pr, Se]
They were willing to risk everything for their faith [Pr, Go]

An added advantage of this proposal is that it provides an elegant format to deal with polysemy and thus reduce dictionary senses, which may be described in terms of different underlying schemas or as different ways of structuring grammatical elements on the basis of a single schema. For example, in *He risked his life*, both schemas B and C can be called up (i.e. 'he risked his life but was not aware of it'; or 'he risked his life for a worthless cause'). One more advantage is found in that the proposed framework provides a way to make a difference between the two common phrases 'take a risk' and 'run a risk': only 'run a risk' fits Schema A; then, both 'run' and 'take' are acceptable with [Ba] as complement, but [De] forces the use of 'take'. Consider the following examples, some of them simplified from the ones provided by the authors:

*Newborn babies run (*take) the risk of hypothermia* [Pr, Ba] (example 1 above).
(Schema A).
He was running/taking the risk of collapsing, though he didn't know it [Pr, Ba].
(Schema B).
He chose to run/take the risk of being hit by a car as he started to cross the road.
[Pr, Ba] (Schema C).
He took the risk of jumping off the cliff [Pr, De]. (Schema C).

We can relate Fillmore and Atkin's discussion to Lakoff's (1987) claims on propositional cognitive models (comparable to Fillmore's frames), which are structured as sets of predicate-argument relationships and often take the form of clusters of concepts that converge into one single gestalt. Thus, the notion of 'mother' can be

described in terms of a cluster of models: the *birth*, *nurturance*, *marital*, *genealogical*, and *genetic* models. In this connection, Ruiz de Mendoza and Pascual (1998), and Santibáñez (1999), have refined Lakoff's proposal by introducing Langacker's (1987) notions of *profile* and *base*. Thus, the notion of 'mother' can be seen as being profiled against a number of different base domains within each member of the cluster; for example, the birth model is profiled against the NEW-LIFE domain, the HOSPITAL domain, the PHYSIOLOGY domain and the PREGNANCY domain⁷:

– The birth model (the mother gives birth to her children)

The NEW-LIFE domain:

- (a) The mother gives birth and, as a result, a baby is born.
- (b) Once born, the child begins to develop as a physiologically independent being.

The HOSPITAL domain:

- (a) When the mother suffers from painful contractions and she realizes that she is about to give birth, the father takes her to hospital.
- (b) At hospital, the mother is taken to the maternity room, where doctors and nurses assist her in safely giving birth to the child.

The PHYSIOLOGY domain:

- (a) Before birth, the child is inside the mother's womb. When the child is about to be born, the muscles of the womb start tightening and the mother feels painful contractions; as childbirth approaches, the period of time between the contractions grows shorter and shorter. The head of the child is the first part to come out of the mother's body. The placenta is expelled after the child comes out. A nurse cuts the umbilical cord linking the baby to its mother.

The PREGNANCY domain:

- (a) The unborn child grows within the mother's body for about nine months. The child is linked to the mother through the placenta and the umbilical cord; this allows the child to get oxygen and food from the mother.
- (b) The mother often goes to the doctor's and is subjected to different tests in order to find out whether everything is all right with her child. Sometimes images of the child are taken with special equipment, so it is possible to know the sex of the baby before it is born.
- (c) During pregnancy, the mother is advised not to smoke, drink alcohol, or do other dangerous things because they may damage the unborn child.

Cognitive modeling as presented here lays out a different scenario from that encoded in logical structures. On the one hand, frames are not formal representations of those parameters that are determinant for argument realization, but simply form-meaning pairings where there is not necessarily a one-to-one correspondence between each element of the conceptual representation and syntactic realization. On the other

⁷ For a full discussion of the different models and their corresponding bases and profiles, see Ruiz de Mendoza and Pascual (1998) and Santibáñez (1999).

hand, frames provide a comprehensive account of the conceptual framework underlying the meaning of a predicate, which can account for many of its use aspects. In our view, it would be nice if we could find a way to maintain the elegance in the formalism as posited in the more syntactico-centric approaches together with the rich expressiveness in the conceptual representation as formulated from the more semantically-oriented proposals. In this regard, we have developed a new system of lexical representation that bridges the gap between the two proposals and strikes a productive balance between the two representation orientations.

4. Towards a unified account: a glimpse at the notion of lexical template

At this stage, we have presented, concisely and selectively, some of the most relevant parameters that permeate the most recent research in lexical representation. We have seen that lexicologists have to decide upon the nature of the language they will be using, and on whether they will develop enhanced semantic representations that go beyond traditional logical form postulates. The issue now is to ascertain what type of lexicological procedure best suits the real demands of a robust lexical theory. As a reaction to this caveat, Ruiz de Mendoza and Mairal (2006, in press) began to articulate a new model of language description firmly rooted in strong functional and cognitive bases. The resulting framework, the *Lexical Constructional Model*, develops a new system of lexical representation, called *lexical template*, which is part of a comprehensive theory of lexical representation.

In what follows we will outline the major features of a lexical template by contextualizing the proposal within the larger setting of both the more syntactically oriented approaches – e.g. logical structures – and the more semantically rooted – e.g. frames. In this regard, we would like to show that a lexical template is a type of structure that falls halfway between the two representational systems and manages to integrate many of their features into one unified representational format. Here are some of the most relevant methodological criteria⁸:

- A lexical template is based on an abstract semantic metalanguage, very much in accord with those event structure templates or RRG's logical structures. However, unlike these formal representations, the type of metalanguage used is typologically based *a la* NSM and stems from lexicographic work, which means that most units have been extracted from a natural language inventory (Faber and Mairal, 1999).
- Lexical templates are constructed on the basis of the *Aktionsart* distinctions proposed in RRG. This feature of lexical templates allows the analyst to introduce a large degree of regularity in his description of "inheritance" mechanisms, which enhances the explanatory adequacy of the model. Such *Aktionsart* regularities are captured by the external variables of the template (which roughly correspond to RRG's logical structures) and by a set of high-level elements of structure that resemble traditional semantic primitives but that differ from them in significant ways.

⁸ For a full account of the notion of lexical template, we refer the reader to Mairal (2004), Mairal and Faber (2002, 2007), Ruiz de Mendoza and Mairal (2006, in press) as well as all the material posted on the website: www.lexicom.es.

- The methodological procedure to construct a lexical template allows us to identify where the chain of a semantic decomposition actually ends. Thus, for every lexical class we have identified an undefinable item, which we will use as the prime in a logical representation. With this procedure we avoid identifying primes on a purely ad-hoc basis, as has been common practice in most syntactically oriented models, where the inventory of constants has never been specified. This means that if we want to represent the logical structure for, say, *embelesar* ('fascinate') in Spanish, we will use the primitive *sentir* ('feel'), which is the prime that underlies all predicates that belong to the domain of FEELING.
- Lexical templates take a firm step towards enriching lexical representations with detailed semantic descriptions thus very much in consonance with, for example, frame semantics. However, there are several differences. Instead of postulating situation types, our approach divides a lexical domain into a number of lexical subdomains which focalize different semantic and pragmatic phases: for example, 'to cause somebody to understand something', 'to think something is true' are just some of the lexical subdomains that pervade the domain of cognition (cf. Faber and Mairal, 1999, for a classification of the architecture of the English lexicon in terms of lexical domains and subdomains). The crucial issue is that frames and lexical templates are different ways of capturing the elements of semantic / pragmatic scenarios, although they share the idea that there is a core element plus a number of peripheral elements, the sum of which give expression to the meaning of a predicate.
- As shall be seen below (section 4.1), lexical templates incorporate a formalism that includes the RRG logical structures plus a number of operators, called lexical functions. These are used to codify the exact semantic and pragmatic parameters of a predicate. An added advantage of having a formalism of this type is that in terms of developing an algorithm that links semantic with syntactic information, all the representations that follow the algorithm, namely, constructional templates and lexical templates, make use of the same metalanguage. This is an issue that should not be underestimated because of its evident implications both in terms of linguistic theory and for the computational implementation of the model. In some other publications, we have fully developed this algorithm, which revolves around a conceptual operation called *subsumption*, whereby higher-level structures take in features of lower-level structures. This means lexical templates are in fact lower-level constructions that can be fused into higher-level characterizations such as the *caused-motion* (e.g. *The audience laughed the actor off the stage*), the *resultative* (e.g. *He kissed her unconscious*), and the *benefactive* (e.g. *He drank to my health*) constructions. Since the formal apparatus of lexical templates shares with higher-level constructions all elements excepting those that are specific to a lower-level class (the internal variables), absorption of a lexical template by a construction becomes a straightforward process.

Let us now move on to comment on the exact ingredients of the standard formalism for lexical templates including a more recent extension that is being developed within the context of Pustejovsky's (1995) generative lexicon.

4.1. Just a word on the formalism

When we began working on enriching logical structures with a more robust semantic component, we saw the need to include a semantic module as part of the logical structure of the predicate, thus following the pioneering proposals in Van Valin and Wilkins (1993:511) for the predicate *remember* or Van Valin and LaPolla (1997:117) for the representation of speech act verbs. As a result, we designed lexical templates with the following format (cf. Mairal, 2004; Mairal and Faber, 2002):

contact-by-impact verbs (Mairal, 2004)

[[**do'** (w, [**use.tool**.(α).**in**.(β).**manner.for**.(δ)' (w, x)] CAUSE [**do'** (x, [**move.toward'** (x, y) & INGR **be.in.contact.with'** (y, x)], $\alpha = x$).

Despite the fact that this representation is couched in more elaborate semantic decompositions, these first lexical templates were still not systematic enough in their use of activity and state primitives. Primitives such as *manner*, *tool*, and *use* appear in these representations, but again no explanation is given of how they have been obtained. Moreover, we noted that the resulting representations turned out to be too unwieldy and lacked transparency and elegance in the expression. These observations became more obvious both when we began to use lexical templates in cross-linguistic analysis and when we took the first steps towards the computational implementation of the model.

Consequently, we decided to simplify the system by postulating two different modules both of which were based on a universal abstract semantic metalanguage. The resulting templates have two parts: (i) the semantic module, and (ii) the logical representation or *Aktionsart* module, each of which is encoded differently. Here is the basic representational format for a lexical template:

predicate: [SEMANTIC MODULE<lexical functions>] [AKTIONSART MODULE <semantic primes>]

The rightmost 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 putatively candidates for semantic primes, or else, these cannot be further decomposed. 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 *red* 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.

The semantic and pragmatic properties of the semantic module are formalized by making use of lexical functions such as those used in Mel'cuk's Explanatory and Combinatorial Lexicology (ELC) (cf. Mel'cuk, 1989; Mel'cuk et al., 1995; Mel'cuk and Wanner, 1996; Alonso Ramos, 2002)⁹. These lexical functions have also been shown to

⁹ According to Mel'cuk *et al* (1995: 126-127), a lexical function (LF) is written as: $f(x) = y$, where f represents the function, x , the argument, and y , the value expressed by the function when applied to a given argument. The meaning associated with an LF is abstract and general and can produce a relatively

have a universal status (cf. Mel'cuk, 1989), something which is in keeping with our aim of providing typologically valid representations. Unlike what is the case in Mel'cuk's work and the complete literature on the *Explanatory Combinatorial Dictionary*, in our approach lexical functions are essentially paradigmatic and capture those pragmatic and semantic parameters that are idiosyncratic to the meaning of a word, which allows us to distinguish one word off from others within the same lexical hierarchy. For example, if we want to account for the semantic differences between *mandar* ('command'), *ordenar* ('order'), *decretar* ('decree'), *preceptuar* ('set up a precept'), *preinscribir* ('preregister') from the lexical domain of speech acts or *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. Then, we have devised a semantic module that 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 grammatical impact.

Thus, far from having two independent modules that do not speak to each other, the two representations here do have a direct correlation since external variables as encoded in the *Aktionsart* module are co-indexed with the numeral subscripts used in the semantic module, which has strong computational implications. Here is a sample of three lexical templates, although we refer the reader to the Faber and Mairal (2007) and Ruiz de Mendoza and Mairal (in press) for abundant exemplification:

fathom: [MAGNOBSTR & CULM_{12[ALL]}] **know'** (x, y)
order: <MAGN_{1[PERM]23},PURP₂ (do)₃> [**do'** (x, [**say'** (x,y))]] CAUSE [**do'** (y,z)]
command: LOC^{soc†}₍₁₎ (PLACE_TYPE: political/military) [order]

In the case of *fathom*, 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. Additionally, the semantic module includes two lexical functions: MAGNOBSTR, which indicates the great difficulty on the part of x in understanding y, although this process of acquiring knowledge is successfully culminated as encoded in the lexical function CULM_{12[ALL]}, where [all] refers to the propositional content of the object of apprehension. In the domain of speech act verbs, one of the subdomains is that of *ask* verbs (*to say something to cause somebody to do it*), which encode a rich set of pragmatic factors dealing with social status, the power differential between speaker and receiver etc¹⁰. A case in point is the predicate *order*, and *command*. The *Aktionsart* module designates a causative accomplishment where there is a causing activity – x says something to y –

high number of values; e.g. **Magn** expresses intensification and can be applied to different lexical units thus yielding a high set of values:

Magn	(Engl. smoker) =	heavy
Magn	(Engl. bachelor) =	confirmed
Magn	(Sp. error) =	craso
Magn	(Sp. llorar) =	llorar como una Ma

¹⁰ See Faber and Mairal (fc) for a full discussion of 'ask' verbs and their corresponding formalization in terms of lexical templates.

that causes that *y* does *z*. This event structure is modified by two semantic parameters that specify a rather forceful way of asking [MAGN], because the speaker is trying to get the addressee to do something, and a second lexical function PERM that is co-indexed with the first argument and signifies that the speaker has power over the addressee and is thus licensed to ask him/her to do things. The predicate *command* is a hyponym of *order* and it inherits all its semantic properties. The *Aktionsart* module remains the same while the semantic module specifically includes the powerful social status and the speaker's very high social position, a facet that is encoded in the superscript (^{soc}) in conjunction with the function LOC, which in this case refers to social location. The resulting function, LOC^{soc}, is followed by arrows that indicate whether the speakers' social status is high ([↑]) or low ([↓]). The parenthesis (PLACE_TYPE: political/military) refers to context or (social) place type (cf. Faber and Mairal, fc).

In an attempt to simplify the formalism in order to avoid the sometimes *ad hoc* adscription of a lexical function to a semantic parameter, Mairal and Cortés (in prep.) have recently initiated a reconversion of the inventory of lexical functions by looking at Pustejovsky's (1995) generative lexicon¹¹ and more in particular to the set of *qualiae*, which we reproduce here for convenience (Pustejovsky 1995: 76, 85-86):

- CONSTITUTIVE (Q_C): the relation between an object and its constituent parts

- i. material
- ii. weight
- iii. parts and component elements

- FORMAL (Q_F): that which distinguishes it within a larger domain

- i. orientation
- ii. magnitude
- iii. shape
- iv. dimensionality
- v. color
- vi. position

- TELIC (Q_T): its purpose and function

- i. purpose that an agent has in performing an act
- ii. built-in function or aim which specifies certain activities

- AGENTIVE (Q_A): factors involved in its origin or 'bringing it about'

- i. creator
- ii. artifact

¹¹ Pustejovsky's (1995: chapter 5) generative lexicon includes four levels of representation: (i) argument structure; (ii) event structure; (iii) *qualia* structure and (iv) lexical inheritance structure, together with a complete set of generative devices (e.g. type coercion, selective binding, co-composition) that connect up the four levels. In the present paper, we focus on how *qualiae* serve the same purpose as the lexical functions in the semantic module. Unfortunately work on the notion of *qualiae* has, to the best of our knowledge, been discontinued. We believe that the inventory of *qualiae*, as it stands, is far from exhaustive and a fined-grained description would certainly be welcome.

- iii. natural kind
- iv. causal chain

The following are examples of lexical representations based on this system (cf. (Pustejovsky 1995: 82, 101), although we have slightly changed some of Pustejovsky's notational devices and have adapted them to a system that is closer to ours:

book
 ARGSTR = [ARG1 = **x: information**
 [ARG2 = **y: phys_obj**]

 QUALIA = **information·phys_obj_lcp**
 FORMAL = **hold (y,x)**
 TELIC = **read (e,w,x·y)**
 AGENT = **write (e', v, x·y)**

This representation specifies that the nominal predicate *book* belongs to the lexical conceptual paradigm (*lcs*) of physical objects and accounts for the telic and agentive interpretations that make reference to the dotted arguments (*x* and *y*), which are in turn featured as 'information' and 'physical object'. Now, consider a more complex representation :

build
 EVENTSTR = [E₁ = **e₁: process**
 E₂ = **e₂: state**
 RESTR = <_α
 HEAD = **e₁**

 ARGSTR = [ARG1 = **x: animate_ind**
 FORMAL = **phys_obj**]

 [ARG2 = **y: artefact**
 CONST = **z**
 FORMAL = **phys_obj**]

 [D-ARG = **z: material**
 FORMAL = **mass**]

 QUALIA = **create-lcp**
 FORMAL = **exist (e₂, y)**
 AGENTIVE = **build_act (e₁, x, z)**

This representation specifies the event structure, the argument structure and the *qualiae* of the predicate *build*. The event structure indicates that the verb *build* is an accomplishment verb that involves a process and a result state ordered by the relation "exhaustive ordered part of", <_α. The initial event has been headed, which means that the action that brings about the state is focused upon or fore-grounded. In relation to the argument structure, there are two true arguments (i.e. those that are syntactically realized) and a default argument (parameters that are relevant for the *qualiae* but are not

syntactically realized). In the *qualia* structure, the lexical conceptual paradigm is also noted, i.e. build is a creation verb, as well as the two processes involved: the agentive, that involves both argument1 and the default argument, which is related to the object by the constitutive relation of argument 2. The formal role indicates the final result state (cf. Pustejovsky, 1995:63: 71-73; 82)

Both representations include an event structure description –which, details aside, coincides to a large extent with the *Aktionsart* module– and a *qualia* structure, whose function is to specify the specific semantic properties of each of the arguments involved in the event. Interestingly enough, a brief mention to the lexical class is also included, which happens to be one of the hallmarks in our approach.

Since both *qualiae* and lexical functions are used to impose the semantic constraints that are operative in a lexical entry, there should not be a lot of difficulties in reorganizing and rephrasing lexical templates following a *qualia* format. For example, if we look back at the representation for the predicate *fathom* above, we could rewrite the semantic module in the template as follows:

fathom:

[{Q_F: MagnObstr **think'** (x, y) / Q_T: Culm **know'** (x,y)}] **know'** (x, y <ALL>)

This new format is expressed in terms of two *qualiae*: the *formal* and the *telic*. The formal *qualia* describes the great difficulty involved in carrying out the process of thinking, while 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. At this stage, the question that arises is the following: what are the potential advantages of this new formalism? As explained in Mairal and Cortés (in prep), both modules – the *Aktionsart* and the semantic module – are closely intertwined: semantic restrictions of the kind expressed in *qualia* structures show the often complex ways in which subevents are interrelated. As Pustejovsky (1995: 101-104) has pointed out, individual *qualiae* compete for projection, and there are mechanisms such as foregrounding or ‘focalising’ a single *quale* of the verbal semantic representation. For example, consider the lexical template for a causative change of state verb like *break* as illustrated in Mairal and Cortés (in prep):

break:

[{Q_F: *broken'* (y) / Q_A: **do'**(x, *break_act'*)}] **do'** (x, Ø) CAUSE [BECOME/ INGR *broken'* (y)]

As is commonly known, change-of-state verbs typically describe an initial activity followed by a resulting state. These two phases in the causative structure map onto the agentive and the formal *qualia* respectively. Depending on which *quale* is foregrounded (‘headed’ in Pustejovsky’s terminology) the verb can be constructed in a transitive (causative) or an intransitive (anticausative) structure.

5. Applications

So far, we have been arguing that the notion of lexical template is a serious alternative to both syntactico- and semantico-centric approaches. Lexical templates provide richer descriptions than other alternatives (e.g. the NSM, FG, and RRG

approaches) and at the same time bind each semantic element to logical variables that can be projected syntactically. It is thus sensitive both to frame semantics criteria and to meaning extensions through cognitive operations such as metaphor and metonymy but does not multiply meaning components *ad infinitum*. Each lexical template is part of a complex lexematic network that contains all the meaning ingredients that are necessary for common pragmatic and discourse implications. Lack of space prevents us from dealing with this issue, but the reader may refer to Faber and Mairal (1999) and Ruiz de Mendoza and Mairal (in press) for details on the nature of lexematic networks and their relation to our proposal on lexical templates.

Lexical templates have further shown to be applicable in computation and lexicography. In the computational context, Guest and Mairal (2007) have taken the first steps towards the implementation of these structures within a complex ontological framework called *Universal Lexical Metalanguage* (ULM). Briefly put, this framework rests upon an ontology, which consists of two subontologies: one that account for predicates (a *predicate ontology*) and a second module that is concerned with the semantic properties of objects (an *object ontology*) plus the set of mechanisms that specify the different ways the two ontologies interact. The overall aim is to develop a system that allows the formulation of intelligent reasoning algorithms within the context of the semantic web from a system that is based on a rich description of the meaning properties of the predicates in the lexicon. One of the innovations in this project is the reformulation of lexical functions in terms of ‘intervals’, a mathematical notion that is used within a fuzzy logic context. Briefly put, intervals are used to define semantic space and specify a continuous range such that words can map onto a number of intervals. Intervals specify ranges that have to do with the physical world in which we live, our (common) thought and emotional processes, and the results of actions in the world. Various operators can be defined on them (cf. Guest and Mairal, 2007: 3.1.):

SUP: top end
MID: middle
INF: bottom end

PLUS: move up the interval
MINUS: move down the interval

An interval is attached to a prime and describes the range of meanings each prime can have. Moreover, those predicates that are derived from more than one prime can inherit all the intervals from all of its parent primes. So, for example, **understand** can be defined as **know** (the prime), but where the depth of knowledge of X about Y is at the upper end of the Depth interval, as shown in the following representation:

SUP (Depth) **know** (X,Y)

In addition to intervals that are used to partition semantic space, we need a precise semantic definition: the *semantic structure*. The semantic structure provides a readable definition for a given predicate, which is done by means of combining primes and intervals together. Needless to say, since primes are regarded as undefinable units, these do not have a semantic structure as part of their representation. Following Guest and Mairal (2007:209), here is a representation of some of the functions that operate at this level of analysis:

ACTION	Describes any actions involved in the verb.
SEQUENCE	Describes a sequence of actions that occur consecutively
RESULT	Describes the results of an action or sequence
BEFORE	Describes the situation before the start of the ACTION or SEQUENCE
CAUSE	Anything directly caused by an ACTION or SEQUENCE. May also occur within other headings.
PURP	The purpose of an ACTION or SEQUENCE
REASON	The reason why an ACTION or SEQUENCE is carried out
SOCIAL OPINION	Describes any social/cultural background that is key to the meaning.
MEANS	Gives a list of predicates that could describe how the RESULT is achieved
ASSUMPTIONS	List of assumptions behind the predicate

The lexical representation for the predicate *peep* is represented by a logical structure with two arguments (x = agent and y = thing or scene) where three different primes and their corresponding intervals are at work: **see**, **do** and **want**. The semantic structure module specifies the action of seeing whereby x takes some time (MID(length)) and expends from a fair amount to a lot of effort (MID&SUP(Effort)) in order to see y . The action of seeing occurs because x wanted to see y from moderately to very badly (MID&SUP(WantIntens)) and achieved his aim for a while (MID(SeeTime)). Finally, there is a social opinion keyword in the representation that accounts for the fact that what x did to y is not socially accepted, i.e. what x did is bad.

peep(X,Y): see, do, want

Intervals: SeeTime, Process, Length, Effort, WantIntens

Participants:

Actor X: {human}

Patient Y: {thing, scene}

Semantic Structure

ACTION

MID(Length) MID&SUP(Effort) **do**(X, MID(SeeTime) **see**(X,Y))

REASON

MID&SUP(WantIntens) **want**(X, **see**(X,Y))

SOCIAL OPINION

be(X,bad)

Syntactic Template: peep at

The predicate **stare** takes in two arguments x and y and is defined in terms of two primes –see and do– and their corresponding intervals. The definition reads as follows: x sees y and in this action x makes use of a fair amount of time and effort: MID(Length) MID&SUP(Effort) **do**(X , MID(SeeTime)). Besides, there are two reasons to explain why x did the action: (i) x is surprised at what x is seeing or else x wants to make y feel fear.

<p>stare(X,Y): see, do</p> <p>Intervals: SeeTime, Process, Length, Effort</p> <p>Participants: Actor X: {animate} Patient Y: {thing, scene}</p> <p>Semantic Structure</p> <p>ACTION MID(Length) MID&SUP(Effort) do(X, MID(SeeTime) see(X,Y))</p> <p>REASON feel(X, surprise) OR want(Y, feel(fear)) (Y = animate)</p> <p>Syntactic Template: stare at</p>
--

The description of these two lexical entries is an oversimplification of a full complex project that manages to formally code meaning within an ontological framework. For the purposes of this paper, suffice to note that if we posit semantically enriched lexical entries using the right formal metalanguage, it will be easier to develop and retrieve contextual information by means of intelligent reasoning algorithms.

The second application that has emerged from our treatment of lexical templates falls within the area of lexicography. In connection with this, a group of researchers have begun working on a syntactic dictionary based on semantic principles provisionally termed *DISSE* (*Diccionario Sintáctico y Semántico del Español Actual* ‘A Semantic and Syntactic Dictionary of Present-Day Spanish’)¹². This dictionary, organized into coherent semantic classes, aims to provide a finer description of the set of morphosyntactic properties and configurations of a lexical entry. One of the central corollaries is that the morphosyntactic properties cannot only be described exhaustively but can also be explained exhaustively in semantic terms. Most of the existing syntactic dictionaries only describe the syntactic properties of a given lexical entry without any attempt to explain the semantic motivation that underlies the different complement strategies. This is then the leading thesis behind this project, i.e. the search for the set of

¹² In what follows we include a very brief discussion of the lexicographic dimension of the preceding lexical representations. This work was done within the framework of the *Lexicom* research group. The fundamental lexicological guidelines together with the analysis of the format of a lexical entry are developed in Ruiz de Mendoza et al. (2007).

semantic regularities that motivate syntactic occurrences. Let us discuss the range of phenomena that this dictionary can provide an answer for. Firstly, a syntactic dictionary based on semantic principles accounts for the contrasts between structures like *considero que + object clause* and *te considero + object NP XCOMP*. In the former the subject/speaker's judgment can be based on indirect or second-hand evidence although this judgment does not necessarily have to coincide with his/her personal opinion because he/she might be speaking on behalf of someone else. This contrasts with the latter configuration where the subject/speaker's judgment comes from his/her own perceptions and by virtue of a direct first-hand evidence as shown by the following entailments:

Considero que eres un fontanero eficiente
 considerPRES1sg that bePRES2sg a plumber efficient
 'I consider John an efficient plumber'

(a) *o al menos eso es lo que me dicen todos los vecinos del bloque* ('or at least that is what all the neighbors in the block tell me')

(b) *aunque yo personalmente creo que algunos aspectos de la mano de obra podrían mejorarse* ('although I personally believe that some aspects of his work can be improved')

Te considero un fontanero eficiente
 you considerPRES1sg a plumber efficient
 'I consider you an efficient plumber'

(a) # *pero la verdad es que nunca has trabajado para mí ni te he visto trabajar para otra persona* ('but the truth is that you never worked for me nor have I seen you work for someone else')

(b) # *aunque yo personalmente no creo que seas eficiente* ('although I personally don't think that you are efficient')

Extending this semantic principle somewhat further also explains the ill-formedness of this structure with the reflexive since the reflexive in Spanish entails that the judgment expressed in the proposition has to be necessarily based on direct first-hand evidence. This observation accounts for the fact that if *considerar* selects a reflexive pronoun as its object, this pronoun cannot be combined with a clause introduced by *que* ('that'):

* *Pedro se considera que es un fontanero eficiente*
 Peter_i himself_i considerPRES3 sg that bePRES3 sg a plumber efficient
 '*Peter considers himself that he is an efficient plumber'

Pedro se considera un fontanero eficiente
 Peter_i himself_i considerPRES3 sg a plumber efficient
 'Peter considers himself an efficient plumber'

It is not surprising that verbs such as *comprender* and *entender* (both of them meaning 'understand'), unlike *considerar*, block out this construction since these verbs involve the subject/speaker's acceptance of a judgment or opinion from a different person or source. This semantic interpretation clashes with that of the secondary

predication, which, as advanced above, requires the judgment or opinion encoded in the proposition to exclusively come from the subject/ speaker's universe of perception:

**Juan comprende/entiende a Pedro un fontanero eficiente*
 John understandPRES3 sg to-Peter a plumber efficient
 ‘*John understands Peter an efficient plumber’

Juan considera a Pedro un fontanero eficiente.
 John consider PRES3sg to-Peter a plumber efficient
 ‘John considers Peter an efficient plumber’

Each lexical entry thus consists of a clear-cut delineation of the different senses involved together with their corresponding syntactic patterns in such a way that the user can ascertain the semantic principles that explain contrasts like the following:

* *Estoy considerando que voy a ir a la fiesta*
 be PRES1sg considering that goPRES1sg to-go to-the-party
 ‘*I am considering that I go to the party’

Estoy considerando si voy a la fiesta
 be PRES1sg considering if goPRES1sg to-the-party
 ‘I am considering if I will go to the party’

Estoy considerando ir a la fiesta
 be PRES1sg considering to-go to-the-party
 ‘I am considering going to the party’

The dictionary also allows the user to be aware of cases of constructional coercion (cf. Michaelis, 2003; Goldberg, 2006) where grammatical form and function overrides the default properties of a lexical item. A case in point the use of the imperative form with state predicates which, far from designating a property, encode some sort of invitation, advice, suggestion or request from the speaker to the hearer (cf. González-García, 2007)

‘*Considérense ustedes en su propia casa*’ --empezó diciendo el padre prior— ‘y sírvanse disculpar los modales de nuestro portero’ (ADESSE)
 ‘Consider yourselves at home --began to say the father prior-- and please excuse our doorman’s manners’

In essence, following the spirit of work on lexical templates, where semantic description motivates syntactic projection, *DISSE* is a syntactic dictionary based on semantic principles such that syntagmatic properties are not only described but are also explained.

6. Conclusions

After discussing two of the most relevant parameters in lexical design, that is, the nature of the metalanguage and the scope of the representation, this paper has laid out the foundations for an alternative form of lexical representation called lexical template.

Lexical templates draw insights both from models with a stronger syntactic orientation (e.g. RRG's logical structures) and from accounts where semantic description is more important (e.g. Frame Semantics). Moreover, lexical templates make use of a semantic metalanguage obtained through factorization of common meaning components of items belonging to the same lexical class. The metalanguage thus consists of a number of semantic primes which largely coincide with those proposed on the basis of extensive typological analysis by scholars like Mel'cuk and Wierzbicka.

Worthy of note is also the fact that some linguists claim that argument realization is not strictly lexical but rather constructional, a cardinal methodological underpinning within constructionist-based approaches (e.g. Goldberg, 1995, 2005). While we do think that constructions are influential in determining argument realization, and in fact the *Lexical Constructional Model* includes a very rich inventory of constructions that operate from the hard core grammar level to the discourse level of language (cf. Ruiz de Mendoza and Mairal, in press), it is our belief that stressing a non-lexical position as far as to regard verbal semantics as not particularly different from constructional semantics is a too radical move that we cannot agree on.

As a concluding remark, we would like to assert that the anatomy of a lexicon and more particular the design of lexical representations still face a number of difficult problems that a serious theory of language, regardless its methodological orientation (functional, formal or cognitive), has to circumvent. We refer to the following issues: (i) the definition of an accurate metalanguage that gives a precise expression to the conceptual ontology, that is, the explicitation of a complete catalogue of what both formal and functional lexicologists have called 'constants'; (ii) the expression of the internal make-up of the conceptual ontology and the way it interacts with the lexicon; (iii) the identification of the real determinants of argument structure – whether these are lexical, extra-lexical or both; (iv) the formulation of exact mechanisms that deal with polysemy. These are just a few challenges that, hopefully, will serve to gradually approximate functional, formal, and cognitive paradigms.

7. References

- Alonso Ramos, Margarita. 2002. "Colocaciones y contorno en la definición lexicográfica", *Lingüística Española Actual* 24, 1: 63-96.
- Dik, Simon C. 1997. *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Berlin /New Yourk: Mouton de Gruyter.
- Dik, Simon C. 1997. *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Edited by Kees Hengeveld. Berlin /New Yourk: Mouton de Gruyter.
- Dowty, David 1979. *Word Meaning and Montague Grammar*. Dordrecht: Riedel.
- Faber, Pamela and Ricardo Mairal (fc) 'Lexical Functions and Semantic Metalanguage'; ms. in preparation.
- Faber, Pamela and Ricardo Mairal. 1999. *Constructing a lexicon of English verbs*. Berlin / New York: Mouton de Gruyter.
- Fillmore, Charles and B. S. Atkins. 1994. Starting where the dictionaries stop: The challenge for computational lexicography. In B. S. Atkins and A. Zampolli (eds.), *Computational approaches to the lexicon*. Oxford: Oxford University Press. 349-393.
- Fillmore, Charles and B. S. Atkins. 1998. "FrameNet and lexicographic relevance". In A. Rubio, N. Gallardo, R. Castro, A. Tejada (eds.) *Proceedings of the First International Conference on Language Resources and Evaluation*. Vol. 1. 417-423.

- Fillmore, Charles. 1985. Frames and the semantics of understanding. *Quaderni di semantica* VI: 222-254.
- Goddard, Cliff and Anna Wierzbicka (eds.) 1994. *Semantic and Lexical Universals. Theory and Empirical Findings*. Amsterdam: John Benjamins.
- Goddard, Cliff and Anna Wierzbicka (eds.) 2002. *Meaning and Universal Grammar*. Amsterdam: John Benjamins.
- Goldberg, Adele. 1995. *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, Adele. 2005. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- González-García, Francisco and Christopher S. Butler 2006. Mapping functional-cognitive space. *Annual Review of Cognitive Linguistics* 5: 39–96.
- González-García, Francisco. 2007. ‘Constructional polysemy meets coercion: The case of the subjective-transitive construction in English and Spanish’ *Language Sciences*; in press.
- Guest, Elizabeth and Ricardo Mairal Usón. 2007. “Building a computational lexicon for automatic translation: a preliminary discussion”. In Fuertes, Pedro (ed.). *Problemas lingüísticos de la traducción especializada*. Valladolid: Universidad de Valladolid, 197-226.
- Jackendoff, Ray. 1990. *Semantic Structures*. Cambridge, Mass.: The MIT Press.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Langacker, Ronald. 1987 *Foundations of Cognitive Grammar I*. Stanford: Stanford University Press.
- Levin, Beth and Malka Rappaport. 2005. *Argument Realization*. Cambridge: Cambridge University Press.
- Mairal, Ricardo and Francisco Cortés (in preparation). “Rethinking logical representations in Role and Reference Grammar: lexical and morphological templates”
- Mairal, Ricardo and Pamela Faber. 2002. “Functional Grammar and lexical templates.” In Mairal, Ricardo and María Jesús Pérez Quintero (eds.) *New perspectives on argument structure in Functional Grammar*. Berlin /New York: Mouton de Gruyter. 41-98.
- Mairal, Ricardo and Pamela Faber. 2007. “Lexical templates within a functional cognitive theory of meaning”. *Annual Review of Cognitive Linguistics* 5: 137-172.
- Mairal, Ricardo. 2004. “Reconsidering lexical representations in Role and Reference Grammar.” *Proceedings of the XXVII International AEDEAN Conference*. Salamanca: University of Salamanca.
- Martín Mingorance, Leocadio. 1995. “Lexical logic and structural semantics: methodological underpinnings in the structuring of a lexical database for natural language processing.” *Panorama der Lexikalischen Semantik*. Ed. Ulrich Hoinkes. Tübingen: Gunter Narr. 461-474.
- Mel’cuk Igor, A. Clas and A. Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve (Belgique), Duculot / Aupelf - UREF.
- Mel’cuk, Igor and Leo Wanner. 1996. “Lexical functions and lexical inheritance for Emotion lexemes in German.” *Recent trends in Meaning-Text Theory*. Ed. Leo Wanner. Amsterdam: John Benjamins. 209-227.
- Mel’cuk, Igor. 1989. “Semantic primitives from the viewpoint of the Meaning-Text Linguistic Theory.” *Quaderni di Semantica* 10 (1): 65-102.
- Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, Mass.: MIT Press.

- Rappaport Hovav, Malka and Beth Levin. 1998. "Building verb meanings". In Butt, Miriam and Wilhelm Geuder (eds.) *The Projection of Arguments: Lexical and Compositional Factors*, CSLI Publications, Stanford, CA, 97-134.
- Ruiz de Mendoza, Francisco and Cristina Pascual Aransaez. 1998. "Conceptual schemas and propositional idealized cognitive models: in search of a unified framework for the analysis of knowledge organization", *Cuadernos Investigación Filológica* XXIII-XXIV: 257-270.
- Ruiz de Mendoza, Francisco and Ricardo Mairal (in press) "Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model". In Mario Brdar and Milena Zic Fuchs (eds.) *Converging and Diverging Tendencies in Cognitive Linguistics*. (Human Cognitive Processing). Amsterdam /Philadelphia: John Benjamins; volume in preparation.
- Ruiz de Mendoza, Francisco and Ricardo Mairal 2006. "Levels of semantic representation: where lexicon and grammar meet." *Interlingüística* 17: 26-47.
- Ruiz de Mendoza, Francisco and Ricardo Mairal. 2007. "High-level metaphor and metonymy in meaning construction". In Günter Radden, Klaus-Michael Köpcke, Thomas Berg, and Peter Siemund. (eds.) *Aspects of Meaning Construction in Lexicon and Grammar*. Amsterdam/Philadelphia: John Benjamins, 33-49.
- Ruiz de Mendoza, Francisco, Ricardo Mairal, Francisco González-García, María Beatriz Pérez Cabello de Alba, and Rocío Jiménez Briones. 2007. 'Diccionario sintáctico y semántico del español actual (DISSE)' Unpublished manuscript.
- Santibáñez, Francisco. 1999. "Semantic structure, relational networks, and domains of experience". *Journal of English Studies* 1: 271-288.
- Taylor, John R. 2003. *Linguistic Categorization*, 3rd edition. Oxford: Oxford University Press.
- Van Valin, Robert D. Jr. 2005. *The Syntax-Semantics-Pragmatics Interface: An Introduction to Role and Reference Grammar*. Cambridge: Cambridge University Press.
- Van Valin, Robert D. Jr. 2006. 'Some universal of verb semantics' In Mairal, Ricardo and Juana Gil (eds.) 2006. *Linguistic Universals*. Cambridge: Cambridge University Press, 155-178.
- Van Valin, Robert D. Jr. and David P. Wilkins. 1993. "Predicting syntax from semantics". In Robert D. Van Valin (ed.). *Advances in Role and Reference Grammar*. Amsterdam/ Philadelphia: John Benjamins. 499-534.
- Van Valin, Robert D. Jr. and Randy LaPolla 1997. *Syntax, Structure, Meaning and Function*. Cambridge: Cambridge University Press.
- Vendler, Zeno 1967. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.
- Wierzbicka, Anna. 1996. *Semantics: Primes and Universals*. Oxford: Oxford University Press.
- Wierzbicka, Anna. 1999. *Emotions across Languages and Cultures. Diversity and Universals*. Cambridge: Cambridge University Press.

