Ontological Semantics and Lexical Templates: mowing the grass from the other side of the fence

Antonio Moreno Ortiz & Chantal Pérez Hernández
University of Málaga, Spain[1]
{amo, mph}@uma.es

1. Introduction

Within the linguistic framework known as Functional Grammar (FG), as envisaged by Dik (1997), a number of scholars have proposed important modifications and additions aimed to improve its descriptive power or adaptation to the needs of different applications.

A notable contribution to the FG paradigm is the proposal by Martin Mingorance (1998) known as the Functional-Lexematic Model (FLM). Within this model a large number of studies have been carried out that are proof of its ability to produce insightful linguistic descriptions. The work of Martin Mingorance and his colleagues has indeed given rise to a large number of studies on a wide variety of topics. The latest theoretical contribution within the FLM, which can in fact be regarded as the pinnacle of work carried out over several years, is embodied in Faber & Mairal (1999), where the cognitive axis of the model appears fully developed.

The FLM itself has also been the object of modifications, proposed within its practitioners. Of key importance to this work are the proposals put forward in Moreno Ortiz (1997), who explores the possibility of incorporating a language-neutral conceptual network, known as *ontology*, to improve the computational adequacy of the original framework. In a nutshell, it is Moreno's thesis that introducing such a resource would not

---

[1] We are extremely grateful to Chris Butler and Ricardo Mairal for their insightful comments on early versions of this article.

only overcome some of the problems for computational treatment that adherence to the FLM involves (largely derived from the model's stress on describing a language by means of the language itself), it also greatly improves the versatility and expressive power of the linguistic model itself. In §3 below, we expand on this idea and provide more details of the original proposal.

Echoing this initiative, Faber & Mairal's article (this volume) takes a radical departure from current FG postulates, and proposes to use RRG-style lexical templates that would serve to describe the meaning definitions of all the verbs in a given lexical domain. These lexical templates contain two types of variables, external and internal. External variables are those aspects of the meaning of a word realized syntactically, whereas internal variables are the semantic parameters which characterize a lexical class. Internal variables are considered to be "ontologically-driven", meaning that these would serve as the connecting elements with a language-independent ontology such as the Mikrokosmos ontology.

In this paper, we expand on this proposal taking the opposite perspective, that is, from the viewpoint of ontological semantics, using our experience in this field, and attempt to outline some of the foreseeable difficulties. In order to do this, we provide an ontological semantic description of the verbs of cutting offered by Faber & Mairal. The assumption is that by comparing the two types of description, we will be able to shed some light on the kind of impact that the incorporation of a language-independent ontology of concepts would have on a linguistic theory of the kind proposed by Van Valin and LaPolla (1997).

This paper should then be considered as complementary to Faber & Mairal's, in that the issue is looked at, as it were, from the other side of the fence. We also set out to identify, from this perspective, in what ways this sort of integration might be useful for the complex task of ontology and lexicon acquisition. Thus, we will also be describing these acquisition tasks within the current framework of ontological semantics.

2. What is ontology?

Defining the term *ontology* is becoming more and more difficult as different interpretations of it spring up every day. Within different environments, ontology is taken to mean different things, making it a delicate task to provide a fairly general, all-encompassing definition.

Gruber's well-known definition, "a specification of a conceptualization" (Gruber 1995: 908), has probably become the most widespread quotation, maybe due to the fact that it captures the essence of the concept without making any further claims as to its nature, content formalization, or applications. However, as is normally the case with such generalizing definitions, it provides very few clues as to what an ontology is.

Guarino (1998: 4) establishes a distinction between *ontology* and *Ontology*. The capitalized form refers to the field within philosophy, traditionally referred to as metaphysics, whereas the lower case version has two different readings: the philosophical community interprets the term as a particular system of categories accounting for a certain vision of the world; on the other hand, the Artificial Intelligence (AI) meaning refers to "an *engineering artifact*, constituted by a specific *vocabulary*, used to describe a certain reality, plus a set of explicit assumptions regarding the *intended meaning* of the vocabulary words" (ibid.: 4, italics in the original).

Gruber (1992) makes a distinction between representation ontologies and content ontologies. The former provide a framework but do not offer guidance about how to represent the world (or the domain), while the latter make claims about how the world should be described.

The following quotation by the creators of EuroWordNet is very useful when describing and classifying ontologies:

> Ontologies differ in their scope (general or domain specific), in the granularity of the units (just terminological labels or units owning more or less complex internal structure), in the kind of relations that can be defined between units, and in the

> more or less precise and well defined semantics of the units and relations (inheritance and other inference mechanisms).
> Vossen et al (1998: 19)

We think, in view of the variety of resources being referred to as an ontology, that it would be fruitful to attempt a characterization of such resources in terms of at least six parameters:

1. Perspective: stress on formalism vs. stress on content.
2. Scope: general or specific.
3. Applications level: general-purpose vs. task-oriented.
4. NL level: language-dependent vs. language-independent
5. AI techniques: number and quality of inference mechanisms supported: inheritance, pattern matching.
6. Semantics (of the entities described): ranging from simple labels to a rich internal structure.

Advancing the description offered in §2.2 below, we are now in a position to offer a profile of the Mikrokosmos ontology in terms of these parameters: it is a profoundly content-driven ontology; general in scope, though imposing no constraints on specificity; geared towards Natural Language Processing (NLP) tasks; it is language-independent; it has a strong dependence on inference mechanisms, especially inheritance; and, finally, it has a rich semantic component capable of producing complex knowledge structures.

2.1 Ontological semantics

Ontological semantics is a theory of computational semantics developed by Sergei Nirenburg and his colleagues, mainly at the Computing Research Laboratory (CRL), New Mexico State University. It is the result of several years of research in the semantic component of Natural Language Processing, and the logical consequence of a knowledge-based approach to this field. Given its primary concern with the development of computational applications, the theory is the set of post hoc conclusions derivable from actual empirical work. Its tenets, descriptions, methodology, implementations and position against the options offered by other related frameworks (e.g. Pustejovsky's (1995) Generative Lexicon) are presented in Nirenburg & Raskin (forthcoming).

Ontological semantics is defined by its proponents as "an integrated complex of theories, methodologies, descriptions and implementations" (Nirenburg & Raskin, forthcoming). As a theory of computational semantics, and unlike linguistic theories, its ultimate aim is not the explanation of linguistic phenomena, but the development of systems capable of capturing and processing word meaning and sentence meaning in a variety of natural languages for a variety of NLP applications. Thus, it makes very few explicit claims as to the nature of human language, other than the ones that can be inferred from the proposed formal treatment of it.

Ontological semantics is based on the assumption that it is possible to reduce any natural language utterance to a formalized, language-neutral representation, an *interlingua* in the traditional sense, which the proponents of this approach call a *Text Meaning Representation* (TMR for short). The TMR is obtained from a complex processing of natural language input in which a number of processors (i.e., software modules that deal with morphological, syntactic, semantic, and pragmatic aspects) are applied to that input utilizing a number of static knowledge sources: ontology, lexica, onomastica, and fact database.

The TMR is the result of a compositional process that relies on the meaning of words, bound morphemes, syntactic structures and constituent order in the input text (Nirenburg & Raskin, forthcoming), and it is meant to capture the meaning of a NL utterance. The question here may be raised as to what kind of meaning is in fact represented; for example, presupposition and entailment may be considered as part of the meaning of utterances. A TMR can in fact capture such non-explicit, non-literal meaning elements, but these may not be necessary at all. For example, for the purpose of MT, the output is bound to be interpreted by a human, able to deduce this kind of information from it. Conversely, a question-answer system will need a larger amount of inferencing if it is to produce satisfactory responses. The bottom line is that a computational semantic application should allow for both brevity (of descriptions) and explicitness.

The static knowledge sources on which processors rely for the construction of the TMR are of two types: *language-specific* (lexica and onomastica) and *language-independent* (ontology and fact database). It is assumed, then, that it is possible to have a separation of what belongs in the description of a language, i.e., how that language expresses communicative content, and the content itself. This subject has been a source of debate not only in linguistics, but also, and especially, in philosophy (Quine's (1960) impossibility of translation). From the point of view of linguistics, the discussion would lead us to other well known bones of contention, such as the separation of linguistic and encyclopedic or real-world knowledge, which definitely lie beyond the scope of this paper. Philosophers, at least from Aristotle's metaphysics, have always been concerned with the distinction between what exists (being qua being) and our perception of it.

We will not attempt to discuss the subject further, but it is important to point out that the motivation, though defensible on purely theoretical grounds, lies on this occasion in the engineering drive that has guided the creation of the resources we are discussing.

An important feature of ontological semantics, especially relevant in this context, is that it does not impose a particular type of morphological or syntactic analysis. Although Nirenburg & Raskin (forthcoming) do provide a specific morphosyntactic framework (an eclectic version of Lexical Functional Grammar) integrated in the CRL implementation, they also point out that many other formulations are possible, and that "ontological semantics is neutral to any such formulations and can be adapted to work with any good quality morphological and syntactic analyzer".

Undoubtedly, of all the static knowledge sources, it is the ontology that sits at the center and provides the key information for the most important processes.

2.2 The Mikrokosmos ontology

A few hints have already been given about the nature of the Mikrokosmos ontology. Its name derives from the project for which it was originally created. The aim of the Mikrokosmos project (Mahesh & Nirenburg 1995, Raskin & Nirenburg 1995) was the creation of a Knowledge-Based Machine Translation (KBMT) system to translate English and Spanish texts on the subject matter of company mergers and acquisitions. Even though the input was restricted to this specific domain, a point was made, following the spirit of former projects, such as Pangloss (Nirenburg 1994, Farwell et al. 1994), that the ontology should contain not only specialized knowledge, but that this be integrated in a general knowledge ontology. The project finished in 1999, and since then the ontology has been constantly refined to accommodate other domains, such as that of Olympic sports, and used in other major projects of a different nature, such as CAMBIO (Nirenburg 2000a) and CREST (Nirenburg 2000b).

The description of the Mikrokosmos ontology offered here is fairly condensed. A proper description would fall beyond the limits of this paper. Such a description can be found in (Mahesh & Nirenburg 1995). We hope this description will be enough for our purposes and that the examples shown in §4 below will also help to clarify it.

The Mikrokosmos ontology is best described as an intertwined hierarchy of frames[2], each frame representing a concept. There is a

---

[2] "Frame" here is used in the AI sense, i.e., a kind of knowledge representation formalism first introduced by Minsky (1975). This term has been imported into a number of linguistic and cognitive studies, such as Lehrer & Kittay (1992) or Fillmore (1982, 1985), sometimes with a fairly different meaning (generally more abstract). Of course, it is also quite different from the "frame" in "predicate frame".

concept ALL[3] that subsumes all the concepts in the ontology except itself. All concepts (except ALL) are of one of three types: objects, events or properties (see the graphic representation in Faber and Mairal, this volume). Properties are to be considered the real primitives in the ontology as they are used to describe objects and events, which are just named sets of properties (and therefore decomposable themselves). Properties are of two types: relations and attributes. Relations establish links (of different kinds, as defined by the relations) between any two concepts (and, by inheritance, their descendants), whereas attributes assign characteristics to them (which are also inherited down the hierarchy of concepts and events). The difference between these properties and other proposed systems of semantic primitives[4] is that they are not just uninterpreted labels, but functions from their domains into value sets. The domain of a property, which must be either specified explicitly (i.e., locally) or inherited (see below) for each property, is the set of concepts (objects and or events) that these properties describe. The value set of a property is called its range, and specifies the set of permitted values for that property. The domain of a property is always a set of objects or events. The range of a property is different depending on the type of property (relation or attribute). The range of a relation is always another set of objects or events, whereas the range of an attribute is either a scalar (a numerical range) or a literal value.

Properties, unlike traditional semantic primitives, cannot be applied freely to description of lexical items; their utilization, in terms of their domain and range is contained in the properties themselves –they come with their own handbook, so to speak. This handbook is in itself the definition of the property, and it is partly determined by the property's position in the hierarchy; for example, a property located under OBJECT-EVENT-RELATION defines a relation that has an object as

---

[3] We will be adopting here the fairly accepted convention of using small caps for signaling concepts in the ontology.

[4] Although this may be an unfortunate generalization, we are loosely making reference here to traditional semantic primitives systems in the sense of, for example (Wierzbicka 1996).

its domain and an event as its range (or the other way round, see the description of inverse relations below). Further, they cannot be applied directly to lexical items, only to concepts (objects or events) in the ontology. It is the assignment of lexical items to these concepts (with or without further specialization) that provides the effective mapping between lexicon and ontology.

In §4 below we discuss the details of ontology-lexicon interaction; suffice it to say for now that it is the properties that describe a concept that effectively describe the lexical item, though in an indirect manner. Properties can be further specialized or refined in the lexical entry, but this is only in those cases in which the lexical item conveys a more specific meaning than that of the concept they are initially assigned to. Because the lexical item's semantic representation depends on the concept it is assigned to, the set of primitives (properties) available for its description is restricted to those that define the concept. In other words, not all the primitives considered in the ontology are liable to describe all lexical items, only those licensed by the concept that a given lexical item is assigned to, which is in turn determined by the domain of each property.

Getting back to the ontology itself, we have already mentioned that inheritance plays a central role in the ontology, which should also be apparent from what we have just mentioned. In fact, inheritance is a common inferencing device in frame hierarchies, as well as being a powerful descriptive feature. Inheritance in the Mikrokosmos ontology is *multiple*, in the sense that a concept may have more than one parent concept, and therefore inherit properties from all of them as well as their ancestors, and *non-monotonic*, meaning that any given default-inherited property can be either overridden or negated for any given concept. Inheritance is ever-present in the ontology and it is impossible to describe a concept without making reference to its inherited properties. For example, when a relation is established between two concepts, it is in effect being established between two (possibly very large) sets of concepts: the two concepts themselves and each of their descendant concepts, as that relation will be inherited along the concept hierarchy (i.e., in both directions). This is so because every relation has an inverse relation, for example: EFFECT <> CAUSED-BY, RECEIVED-BY <> RECEIVER-OF, in such a way that, whenever a relation is specified between two concepts A and B (A - RELATION - B), the system triggers the addition of the relation in the opposite direction (B – INVERSE-RELATION - A); of course, the inverse relation must have been previously defined.

The constitution of the ontology, its fundamental ontological commitments, restrictions, etc. should be explained in terms of its underlying assumptions or axioms. Nirenburg & Raskin (forthcoming) provide a set of 34 axioms that formally condense the properties of the ontology, the most important of which we have just glossed. In the next section we exemplify all that has been explained here by means of the ontological representation of the *manner-of-cutting* verbs.

3. The predicate frame revisited: FG, RRG, and ontological semantics

Faber & Mairal (this volume) point out some of the problems of using predicate frames in lexical descriptions, derived from the fact that lexical entries in an FG lexicon are represented by a predicate frame and a meaning definition, but without an indication of how these two descriptive elements interact. An additional problem, recognized in Moreno Ortiz (1997), is that lexical entries formalized in terms of predicate frames use natural language phrases for lexical representation. This poses several problems, especially in terms of the model's computational adequacy. The limited set of allowed case roles, or *semantic functions* in FG terminology, may also not be enough to cover the semantic richness of natural languages. Faber & Mairal's article makes it clear that the FG inventory of semantic functions cannot account, for instance, for the *result* component codified in *manner-of-cutting* verbs.

(1) **lop** [V] (x_1: animate)$_{Ag}$ (x_2: object)$_{Go}$

df = cut [V] $(x_1)_{Ag}$ $(x_2)_{Go}$ $(x_3\text{: object})_{Source}$ $(\sigma_1$: stroke [N]: quick [A]& strong [A]$)_{Manner}$

(2) **slice** [V] $(x_1\text{: animate})_{Ag}$ $(x_2\text{: object})_{Go}$
df = cut [V] $(x_1)_{Ag}$ $(x_2)_{Go}$ $(x_3$: pieces [N] :thin[A] : flat [A]$)_{Result?}$

This problem was addressed by Moreno Ortiz (1997): The solution offered in this work consists of mapping the predicate, as well as each of its arguments directly to a concept in the ontology. This mapping was meant to be specified for each lexical item individually, very much as it is carried out within the current ontological semantics framework. As we show below, however, ontological semantics does this by shifting the whole semantic description to the ontological level, by mapping only the predicate and optionally refining this by means of a *constrained mapping*. We believe that this approach is more powerful than the one offered in Moreno Ortiz (1997), since the kind of semantic description that it allows can be as detailed as desired.

Faber & Mairal's approach to overcome the limitations of FG predicate frames proposes the characterization of lexical classes by means of lexical templates, a further development of the *predicate schema* proposed in Faber & Mairal (1999), from which the meaning definitions for all of the verbs in the same lexical class can be derived. Programmatically, they propose the following general lexical template for the *manner-of-cutting* verbs analyzed in their paper:

(3) [[**do´** (w, [**use.sharp-edged.tool**(α)**in**(β)**manner´** (w, x) $\wedge$ [BECOME **be-at´** (y, x)]] CAUSE [[**do´** (x, [**make.cut.on´** (x, y)])] CAUSE [BECOME **pred´** (y, (z))]], α = x.

Verbs within this lexical class differ in how they lexicalize a number of semantic parameters: (i) the way in which the activity is performed; (ii) the instrument used; (iii) the entity affected by the action of cutting; and (iv) the resulting state of that entity. Faber & Mairal's lexical template treats these semantic characteristics as *internal* *variables*, named **instrument**, **manner**, **affected object** and **result**. The characterization as internal is due to the assumption that “they are semantic parameters which characterize an entire lexical class” (p.??). The lexical template in (3) also includes *external variables* which, in contrast to the former, are those aspects of the meaning of a word with syntactic realization: **effector**, **patient**, **instrument** and **result** are considered external. We then conclude that some variables can be both internal and external. In the example, **instrument, affected object** (which in fact coincides with **patient**) and **result** are regarded both as internal and external variables, whereas **manner** is only internal and **effector** is only external.

Noticeably, only two of the internal variables, those which are supposed to have ontological status, in this case **instrument** and **manner,** are explicitly marked in the lexical template, referenced by the Greek letters α and β respectively. This in fact implies that some variables, those which are characterized as external only, would not require ontological status. In (3) above, the **effector** would fall into this category, thus being left without an ontological representation. Furthermore, because internal variables are said to characterize an entire lexical class, the number of such variables is kept to a minimum[5], leaving out the possibility of ontologically representing other aspects of lexical meaning that, even if not applicable to a whole lexical class, may still be relevant.

Let us now look at semantic descriptions from “the other side of the fence”, that is, from the point of view of ontological semantics. Within such framework, the ontology is, as we have mentioned, central to all and each of the lexical representations and processors. This means that a strict separation is made between the conceptual and the lexical level. In effect, this translates into the construction of conceptual categories that do not take into account language-specific syntactic

---

[5] In relation to this, we should add that Mairal (personal communication) has stressed the necessity of studying the complete semantic architecture of the lexicon to arrive at a complete and definitive inventory of internal variables.

realizations. Whether a given predicate realizes a given argument syntactically is represented in its lexical representation.

The semantic description of a lexical entry is performed in terms of *ontological concepts* and their *properties* (see § 2.2). The first task for the lexicon acquirer is to find in the ontology the concept that more accurately represents the sense of the word. At this stage, the acquirer must be fully aware that the ontology has been built to be *language-independent* and names for concepts are just labels attached as a kind of mnemonic device. The grammatical class of nouns in English, for instance, can map onto any class of concepts. A noun sense can correspond to an OBJECT (e.g. *river*), an EVENT, when it implies causation (e.g. *flood*) or agentivity (e.g. *punishment*), or may even map onto a PROPERTY (e.g. *cost, weight*). Most verbs map onto EVENT concepts, although some may correspond to a RELATION (e.g. the verb *describe*, maps onto the relation DESCRIBES).

The ontology is not a finished product, as modifications are made to it by acquirers as needs come up while acquiring lexical information[6]. The general policy, however, is that addition of concepts to the ontology should be avoided as much as possible. It is possible to specialize concepts when the semantics of the lexical item being defined do not match exactly the referent of the concept in the ontology. This specialization or refining is called *constrained mapping*.

In the particular case of *manner-of-cutting* verbs, the concept in the ontology that best maps onto the sense of the verbs is CUT and its subclasses. This concept is located in the hierarchy as a subclass of APPLY-FORCE, which in turn is a subclass of PHYSICAL EVENT. Table 1 below contains the set of properties that describes this concept, and that can be used to refine the meaning of lexical items assigned to it.

---

[6] Within the CRL framework, an "acquirer" may be a lexicographer or an ontology constructor. Even though these roles are differentiated and a controlled, systematic protocol is established for interaction between the different acquiring roles, they can in fact be merged in one person.

CONCEPT: CUT
DEFINITION: TO PENETRATE WITH A SHARP INSTRUMENT.
[**IS-A**: APPLY-FORCE]
[**SUBCLASSES**: CARVE, DISMEMBER, PIERCE, SAW-EVENT]

[**ACCOMPANIER**: ANIMAL]
[**ADEQUACY**: (<> 0 1)]
*[**AGENT**: PHYSICAL-OBJECT]
[**BENEFICIARY**: PHYSICAL-OBJECT]
[**CAUSED-BY**: EVENT]
[**CERTAINTY-ATTRIBUTE**: (<> 0 1)]
[**COMPONENT-OF**: EVENT]
[**CONTINGENCY**: EVENT, OBJECT]
[**CONTRASTIVE-RELATION**: EVENT, OBJECT]
[**CONTROLLED-BY**: PHYSICAL-EVENT, PHYSICAL-OBJECT, SOCIAL-EVENT, SOCIAL-ROLE]
[**CONTROLS**: ACTUALIZE, ARTIFACT, NATURAL-OBJECT, SOCIAL-ROLE]
[**DESCRIBES**: REPRESENTATIONAL-OBJECT]
[**DESTINATION**: EVENT, OBJECT]
[**DESTINATION-OF**: EVENT]
[**DIFFICULTY-ATTRIBUTE**: (<> 0 1)]
[**DIRECTION-OF-CHANGE**: NEGATIVE, POSITIVE]
[**DIRECTION-OF-MOTION**: BACKWARD, CLOCKWISE, COUNTERCLOCKWISE, DOWNWARD, EASTWARD, FORWARD, FROM-FORCE, INWARD, NORTHEASTWARD, NORTHWESTWARD, NOTHWARD, OUTWARD, SIDEWARD, SOUTHEASTWARD, SOUTHWARD, SOUTHWESTWARD, TOWARD-FORCE, UPWARD, WESTWARD]
[**DOMAIN-OF**: PROPERTY]
[**EFFECT**: EVENT, OBJECT]
[**ELABORATION-SET-MEMBER**: EVENT, OBJECT]
[**ENABLEMENT**: EVENT, OBJECT]
[**EXCLUSIVITY-ATTRIBUTE**: NO, YES]
[**EXPERIENCER**: ANIMAL]
[**FAIRNESS**: (<> 0 1)]
[**HARSHNESS**: (<> 0 1)]
[**HAS-COMPONENT**: EVENT]
[**HAS-INTENTION**: EVENT, INFORMATION]
[**HAS-PARTICIPANT**: HUMAN]
[**HAS-WORK-EQUIPMENT**: ARTIFACT, DOCUMENT]
*[**INSTRUMENT**: CUTTING-IMPLEMENT]
[**INSTRUMENT-OF**: EVENT]
[**INTENSITY**: (<> 0 1)]
[**LOCATION**: PLACE]
[**PATH**: PHYSICAL-OBJECT, PLACE]
[**PURPOSE**: EVENT]
[**PURPOSE-OF**: EVENT]
[**RANDOMNESS-ATTRIBUTE**: (<> 0 1)]
[**RAPIDITY**: (<> 0 1)]
[**RATIONALITY-ATTRIBUTE**: (<> 0 1)]
[**SOURCE**: EVENT, OBJECT]
[**SOURCE-OF**: EVENT]
[**SUCCESS-ATTRIBUTE**: (<> 0 1)]
[**SUSTAINABILITY**: (<> 0 1)]
[**TEDIUM**: (<> 0 1)]
*[**THEME**: PHYSICAL-OBJECT]
[**THEME-OF**: EVENT]
[**UTILITY-ATTRIBUTE**: (<> 0 1)]
[**VIOLENCE-ATTRIBUTE**: (<> 0 1)]

Table 1: Ontological description of the CUT event.

It is extremely important to understand a number of aspects at this stage. Firstly, most properties (in bold type above) are not assigned to the concept CUT locally, but inherited from higher-level concepts in the ontology. Only some of them have been assigned locally: in this case the slots[7] IS-A and SUBCLASSES, as it is the case for every concept, and the slots AGENT, THEME, and INSTRUMENT (marked with an asterisk in the chart) to override the inherited fillers and make them more specific[8]. Secondly, the values defined in the range of each property are of two types: attributes can have as a value either a scalar, e.g. [INTENSITY: (<> 0 1)], or a set of literals, e.g. [DIRECTION-OF-CHANGE: NEGATIVE, POSITIVE]; relations have another object or event as value, e.g. [THEME: PHYSICAL-OBJECT]. This means that the latter are liable to be further refined, just like the base concept, whereas the former do not allow any further constrained mapping of that kind.

For example, the INTENSITY attribute mentioned above could have (> 0.7) as a value, which could not be further specified; the THEME mentioned above is said to be any PHYSICAL-OBJECT, that is, any descendant of the concept PHYSICAL-OBJECT, which, in turn could be further specialized.

Finally, it is perhaps important to note that case roles are treated by the ontology as relations with a somewhat special status. All of them are subclasses of the concept CASE-ROLE, a subclass of RELATION. Their special status is just due to the fact that only this set of relations is available for linking the syntactic arguments of the predicates with their corresponding semantic roles. As with every relation, however, an inverse case role exists for every case role; for example AGENT <> AGENT-OF, THEME <> THEME-OF, and so on. Thus, case roles are used within constrained mappings along with other relations without any special constraint on their usage other than the ones defined by their domains and ranges, just like any other relation.

---

[7] In the terminology of frame-based knowledge representation every frame has a name (the concept's name) and a number of *slots*, which in turn have *facets*, which can be filled with (references to) other concepts or (literal or scalar) values. The facet determines the type of *filler* that the slot can take; for example, a SEM facet allows a reference to other concepts, a VALUE facet requires a literal or scalar, a DEFAULT facet specifies a typical value and is not inherited down the hierarchy. See Rich & Knight (1991) for an introduction to frame-based knowledge representation.

[8] The inherited value of the slot AGENT, for instance, would have been OBJECT, which, apart from PHYSICAL-OBJECTs, also includes MENTAL or INTANGIBLE OBJECTs.

4. Mapping *manner-of-cutting* verbs onto the ontology: the concept CUT and its “constrained mappings”

In section 3 above we mentioned that all the verbs in the *manner-of-cutting* lexical class map onto the base concept CUT or one its subclasses. The semantic distinguishing features of each of these verbs are described in the process of lexicon acquisition by means of constrained mappings, in which the acquirer has the chance to further specifying the values of the properties already assigned to the base concept. Table 2 below summarizes the constrained properties specified for each of the *manner-of-cutting* verbs.

As this table shows, constrained mappings make it possible to include in the semantic description of each verb not only the semantic features mentioned in Faber & Mairal’s work, but also other distinguishing features encoded in the natural language definitions given in their paper but not explicitly encoded in the proposed internal and external variables of their lexical template.

In constrained mappings, the semantic specification of the way the event of cutting is carried out (the internal variable **manner** in Faber & Mairal’s paper) is made by adding scalar or literal event attributes (i.e., attributes whose domain is an event) to the base concept CUT. In some cases, a relation that describes a sub-event implicit in the action has also been added[9]. This is formalized by the relation HAS-COMPONENT, used to relate the event CUT to some particular accompanying event, such as motion or violent contact.

---

[9] The fact that most cutting verbs consist of two sub-events is also pointed out in Mairal & Faber, but only in relation to the cases in which these verbs encode both the activity and the resulting state of the affected object.

	AFFECTED OBJECT	AGENT	ASPECT (ITER.)	COMPONENT EVENT	EFFECT	FORMALITY (HIGH)	INSTRUMENT	MANNER	PURPOSE
behead	✓	✓			✓				✓
carve	✓	✓			✓				✓
chisel	✓	✓					✓		✓
chop	✓	✓	✓	✓	✓				
clip		✓					✓		✓
cut		✓							
decapitate	✓	✓			✓	✓			
gash	✓	✓			✓				
hack	✓				✓			✓	
hew	✓	✓						✓	
lop	✓			✓	✓			✓	
mow	✓	✓							
nick		✓			✓			✓	
prune	✓	✓							✓
saw	✓	✓							
shave	✓	✓					✓		✓
shear	✓	✓					✓		
slash				✓				✓	
slice	✓	✓			✓				
snip		✓	✓	✓			✓	✓	
whittle	✓	✓		✓					✓

Table 2: Semantic features grid for “manner-of-cutting” verbs.

Let us turn now to see some practical examples of how constrained mappings are carried out in the process of lexicon acquisition[10]. In (1) we show how in the case of the verb *hew,* a DIFFICULTY and a VIOLENCE-ATTRIBUTE (both SCALAR-EVENT-ATTRIBUTES) have been specified with a high value to reflect the fact that the action is carried out with difficulty and in a rough way[11]. In (2) one of the distinguishing features of the verb *lop* is that the action of cutting is performed with a quick, strong stroke. To account for this, the attributes RAPIDITY and VIOLENCE are implicit in the concept HIT which is why they have not been included in the mapping.

(1) ***hew***	(2) ***lop***
CUT PROPERTY: AGENT FACET: SEM FILLER: ANIMATE PROPERTY: THEME FACET: SEM FILLER: SOLID PROPERTY: FRAGILITY FACET: VALUE FILLER: 0 PROPERTY: SIZE FACET: VALUE FILLER: >0.7 PROPERTY: **DIFFICULTY-ATTRIBUTE** FACET: VALUE FILLER: **>0.7** PROPERTY: **VIOLENCE-ATTRIBUTE** FACET: **VALUE** FILLER: **>0.7**	**CUT** PROPERTY: EFFECT FACET: SEM FILLER: DIVIDE PROPERTY: **HAS-COMPONENT** FACET: SEM FILLER: **HIT** PROPERTY: THEME FACET: SEM FILLER: PHYSICAL-OBJECT PROPERTY: THEME-OF FACET: SEM FILLER: DIVIDE

---

[10] The senses of *manner-of-cutting* verbs described by means of constrained mappings in this section are based on the definitions given in Faber & Mairal (this volume).

[11] The indenting in the tables below account for the possibility of further modifying and specifying any filler already included in a constrained mapping. Properties that appear indented apply, then, directly to the concept immediately above and not to the base concept CUT.

The main semantic characteristic of the verb *nick* (3) is that the action of cutting is usually made accidentally[12], whereas in *slash* (4) the cutting is made with a strong and swinging movement. These features have been encoded, respectively, with a RANDOMNESS-ATTRIBUTE, and sub-event component CHANGE-LOCATION that describes the accompanying movement, modified in turn with a VIOLENCE- and a RAPIDITY-ATTRIBUTE.

(3) ***slash***	(4) ***nick***
CUT PROPERTY: **HAS-COMPONENT** FACET: SEM FILLER: **CHANGE-LOCATION** PROPERTY: **VIOLENCE-ATTRIBUTE** FACET: VALUE FILLER:**>0.5** PROPERTY: **RAPIDITY** FACET: VALUE FILLER:**>0.5**	**CUT** PROPERTY: AGENT FACET: SEM FILLER: ANIMATE PROPERTY: **RANDOMNESS-ATTRIBUTE** FACET: **DEFAULT** FILLER:**<0.5** PROPERTY: EFFECT FACET: SEM FILLER: INJURE PROPERTY: INTENSITY FACET: SEM FILLER:>0.5

The verb *hack* (5) is defined as “to cut something into uneven pieces in a rough, violent way”. The attributes VIOLENCE-ATTRIBUTE and HARSHNESS are used describe this particular manner of cutting. In the case of the verb *snip* (6), the action is carried out with quick, short movements: the sub-event CHANGE-LOCATION, modified with a RAPIDITY attribute describes this semantic feature.

(5) ***hack***	(6) ***snip***
CUT PROPERTY: **VIOLENCE-ATTRIBUTE** FACET: VALUE FILLER:>**0.5** PROPERTY: **HARSHNESS** FACET: VALUE	**CUT** **ASPECT:** ITERATION: **MULTIPLE** PROPERTY: AGENT FACET: SEM FILLER: HUMAN

---

[12] This semantic feature, that could be considered as manner, is not included in Faber & Mairal’s table for the internal variable of manner. It is, however, implicit in the definition of the verb: “to cut a nick of sth., usu. accidentally”.

FILLER: **>0.5** PROPERTY: EFFECT FACET: SEM FILLER: SEGMENT PROPERTY: THEME FACET: SEM FILLER: PHYSICAL-OBJECT PROPERTY: THEME-OF FACET: SEM FILLER: SEGMENT	PROPERTY: INSTRUMENT FACET: SEM FILLER: SCISSORS PROPERTY: **HAS-COMPONENT** FACET: SEM FILLER: **CHANGE-LOCATION** PROPERTY: **RAPIDITY** FACET: VALUE FILLER: **>0.5**

As we have already mentioned, our description of **manner** is enhanced in our constrained mappings with the possibility of adding sub-events to the action and further modifying these sub-events. This is also the case in (7): the verb *chop* contains a component of violent contact described as [HAS-COMPONENT: HIT][13]. In the case of the verb *whittle* (8), the action of cutting implies the removal of small and thin pieces of wood. To describe this the affected object (termed THEME in our list of semantic roles) of the sub-event REMOVE has been specified as WOOD, and further modified with attributes pertaining to its particular size and thickness. Aspect has also been marked for the first of these two verbs, *chop*, as well as for *snip* (6) above, to indicate that the action is performed iteratively.

(7) ***chop***	(8) ***whittle***
CUT **ASPECT:** ITERATION: **MULTIPLE** PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: **HAS-COMPONENT** FACET: SEM FILLER: **HIT** PROPERTY: EFFECT FACET: SEM FILLER: SEGMENT PROPERTY: THEME	**CUT** PROPERTY: PURPOSE FACET: SEM FILLER: SHRINK PROPERTY: **HAS-COMPONENT** FACET: SEM FILLER: **REMOVE** PROPERTY: THEME FACET: SEM FILLER: WOOD PROPERTY: SIZE FACET: VALUE FILLER: <0.1

---

[13] It is not necessary in this case to specify a VIOLENCE-ATTRIBUTE, as we did in (3) and (5), because the concept HIT is already described in the ontology by means of this attribute.

FACET: SEM FILLER: PHYSICAL-OBJECT PROPERTY: THEME-OF FACET: SEM FILLER: SEGMENT	PROPERTY: THICKNESS FACET: VALUE FILLER: <0.1 PROPERTY: THEME FACET: SEM FILLER: PHYSICAL-OBJECT PROPERTY: MADE-OF FACET: SEM FILLER: WOOD PROPERTY: AGENT FACET: SEM FILLER: HUMAN

The entity affected by the action of cutting (named a**ffected object** as internal variable and **patient** as external variable in Faber & Mairal's study) is specified by means of the filler assigned to the relation THEME. This filler, as can be seen in (1), (5), (7), and (8), can be further qualified with the assignment of properties that apply directly to it. In the case of *hew*, for instance, the THEME is described as a SOLID, a subclass of the concept MATERIAL that makes reference to several types of hard materials. This SOLID concept has been modified with a very low fragility value and a high size value to account for the fact that someone hews large pieces of hard material such as rock or stones.

The THEME of the constrained mapping for the verb *prune* in (10) is constrained to TREE-BRANCH. In the verb *slice* (9), the physical object affected by the action of cutting is segmented into thin, flat pieces. To describe this in the constrained mapping, we have introduced the relation EFFECT, with the event concept SEGMENT as its filler. In turn, SEGMENT has been modified to indicate that the THEME-OF this event is a PHYSICAL-OBJECT with a very low value in the THICKNESS-ATTRIBUTE that modifies it.

(9) ***slice***	(10) ***prune***
CUT PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: **EFFECT** FACET: SEM FILLER: **SEGMENT** PROPERTY: **THEME**	**CUT** PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: AGENT-OF FACET: SEM FILLER: GARDEN-ACTIVITY PROPERTY: **THEME**

FACET: SEM	FACET: SEM
FILLER: **PHYSICAL-OBJECT**	FILLER: **TREE-BRANCH**
PROPERTY: **THEME-OF**	PROPERTY: PURPOSE
FACET: SEM	FACET: SEM
FILLER: **SEGMENT**	FILLER: GROW-ANIMATE
PROPERTY: **THICKNESS**	PROPERTY: THEME
FACET: VALUE	FACET: SEM
FILLER:<0.5	FILLER: PLANT
	PROPERTY: HEALTH-ATTRIBUTE
	FACET: VALUE
	FILLER:>0.7

Similarly, the objects affected by the action of *chisel* (12) have been described as physical objects MADE-OF wood, metal or stone, whereas the ones affected by *carve* (11) are characterized as MADE-OF wood or stone.

(11) ***carve***	(12) ***chisel***
CUT	**CUT**
PROPERTY: PURPOSE	PROPERTY: **THEME**
FACET: SEM	FACET: SEM
FILLER: TRANSFORM	FILLER: **INANIMATE**
PROPERTY: EFFECT	PROPERTY: THEME-OF
FACET: SEM	FACET: SEM
FILLER: INANIMATE	FILLER: TRANSFORM
PROPERTY: AESTHETIC-ATTRIBUTE	PROPERTY: **MADE-OF**
FACET: SEM	FACET: SEM
FILLER:>0.5	FILLER: **WOOD**
PROPERTY: **THEME**	PROPERTY: **MADE-OF**
FACET: SEM	FACET: SEM
FILLER: **INANIMATE**	FILLER: **METAL**
PROPERTY: **THEME-OF**	PROPERTY: **MADE-OF**
FACET: SEM	FACET: SEM
FILLER: **TRANSFORM**	FILLER: **ROCK**
PROPERTY: **MADE-OF**	PROPERTY: INSTRUMENT
FACET: SEM	FACET: SEM
FILLER: **WOOD**	FILLER: CHISEL
PROPERTY: **MADE-OF**	PROPERTY: PURPOSE
FACET: SEM	FACET: SEM
FILLER: **ROCK**	FILLER: TRANSFORM
PROPERTY: AGENT	PROPERTY: DIRECTION-OF-CHANGE
FACET: SEM	FACET: SEM
FILLER: HUMAN	FILLER: POSITIVE
	PROPERTY: AGENT
	FACET: SEM
	FILLER: HUMAN

Together with the specification of the object affected by the action, it is also important to highlight the way in which constrained mappings allow the acquirer to encode two other semantic features very relevant for the description and differentiation among the *manner-of-cutting* verbs. The first feature is that of **result**, already mentioned above in reference to the verb *slice* (9). The verbs that contain a component of **result** are described in the constrained mappings by adding a relation named EFFECT, with an event that encodes the resulting state as filler. In the case of the verbs *gash* and *nick* (4), the filler of the EFFECT relation is INJURE*,* in the verb *lop* (2) is DIVIDE*,* in the verbs *slice* (9)*, chop* (7) and *hack* (5) is SEGMENT. Sadly enough, in the cases of the verbs *behead* and *decapitate* the filler for the effect relation has to be specified as KILL.

The second of these semantic features is the specification of the PURPOSE that the agent wants to achieve by performing the action. If we recall the constrained mappings of the verbs shown above, it is possible to see that in *prune* (10) the action of cutting has been specified as having the purpose GROW-ANIMATE, with the concept PLANT as THEME, and a further modification of the concept PLANT with a high value in its HEALTH-ATTRIBUTE. In the case of the verb *whittle* (8)*,* the affected object is cut into a smaller size [PURPOSE: SHRINK]*,* whereas with *clip* (15), one cuts pieces from something to make the object shorter [PURPOSE: DECREASE]. Quite sadly again, the PURPOSE of *behead* has to be described as PUNISH. Both *chisel* and *carve* have as purpose TRANSFORM the affected object. However, only the latter has an implicit aesthetic intention, specified in the mapping by the inclusion of an AESTHETIC-ATTRIBUTE with a fairly high value as filler. This aesthetic attribute has also been included in the mapping of the verb *shave* (see below)*,* whose purpose is described as GROW*,* with HAIR as THEME and a very low VISIBILITY-ATTRIBUTE (one shaves to make hair not visible to the human eye).

The purpose semantic feature was not included in Faber and Mairal's internal variables, it might be, however, important to take it into consideration as it has a direct impact on the possible fillers of the

AGENT: if the action is performed with a PURPOSE, it has to be performed by a volitional, and very probably human, AGENT.

(13) ***shave***

```
CUT
    PROPERTY: AGENT
    FACET: SEM
    FILLER: HUMAN
    PROPERTY: THEME
    FACET: SEM
    FILLER: HAIR
    PROPERTY: INSTRUMENT
    FACET: DEFAULT
    FILLER: RAZOR
    PROPERTY: INSTRUMENT
    FACET: DEFAULT
    FILLER: SHAVING-CREAM
        PROPERTY: OPTIONALITY-ATTRIBUTE
        FACET: SEM
        FILLER: YES
    PROPERTY: PURPOSE
    FACET: SEM
    FILLER: GROW
        PROPERTY: THEME
        FACET: SEM
        FILLER: HAIR
            PROPERTY: VISIBILITY
            FACET: VALUE
            FILLER:<0.1
    PROPERTY: AESTHETIC-ATTRIBUTE
    FACET: SEM
    FILLER:>0.7
```

The frame of the verb *shave* above may also help to illustrate the way specifications of the **instrument** used to perform the actions of cutting are included in constrained mappings using the relation INSTRUMENT. In the case of *shave*, the instrument specified is the concept RAZOR, together with SHAVE-CREAM, although this last one has been modified with an OPTIONALITY-ATTRIBUTE, as one may choose not to use any cream when shaving. Also, the facet of these two slots has been specified as DEFAULT.

The table below shows three more verbs in which the specification of INSTRUMENT was carried out differently. In the first case, the verb *saw* (14), the acquirer does not need to include in the constrained mapping any further information about the cutting instrument used to perform the action. This verb maps onto the base concept SAW-EVENT, already included in the ontology as a subclass of the concept CUT. As SAW-EVENT includes in its frame SAW as the INSTRUMENT used to perform the action of cutting, it was not necessary to include again this information in the mapping of the lexical entry. In the cases of *clip* (15) and *shear* (16) it is possible to see how the fillers for the INSTRUMENT relation have been given again the facet DEFAULT (instead of SEM). This was done to indicate that SCISSORS[14] is the most typical instrument but not the only possible one.

(14) ***saw***	(15) ***clip***	(16) ***shear***
SAW-EVENT PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: THEME FACET: SEM FILLER: INANIMATE PROPERTY: MADE-OF FACET: SEM FILLER: WOOD	**CUT** PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: INSTRUMENT FACET: DEFAULT FILLER: SCISSORS PROPERTY: PURPOSE FACET: SEM FILLER: DECREASE PROPERTY: THEME FACET: SEM FILLER: INANIMATE PROPERTY: THEME-OF FACET: SEM FILLER: DECREASE	**CUT** PROPERTY: AGENT FACET: SEM FILLER: HUMAN PROPERTY: THEME FACET: SEM FILLER: HAIR PROPERTY: INSTRUMENT FACET: DEFAULT FILLER: SCISSORS

5. Problems in ontological semantics

5.1 Ontology acquisition

Because the ontology plays such a central role in the linguistic description system in an ontology-based semantic representation framework such as ontological semantics, it is clear that the success and performance of the overall computational system will to a large extent depend on the consistency and soundness of the ontology. We now aim to discuss some of the main problems involved in ontology and lexicon acquisition in such an environment.

---

[14] We should remember again that names for concepts are just mnemonics. The concept SCISSORS in the ontology covers any cutting instrument with two opposing blades pivoted together so that they can work against each other, including thus other more specific types of scissors, such as shears.

A common criticism to the Mikrokosmos ontology and, in fact, to most ontologies, is the linguistic bias of its constructors. The more so in this case since the Mikrokosmos ontology is linguistically motivated, i.e., it is designed for describing and manipulating linguistic objects in several languages. This problem has been recognized by the Mikrokosmos team themselves. Mahesh (1996) warns that it is not easy for the ontologist to avoid a number of pitfalls derived from the knowledge of his or her native language and the particular conceptualization that that language imposes.

The proposed solution (Mahesh 1996) is, first, the adoption of a well defined ontology acquisition protocol, by means of which lexicon acquirers are considered "customers" who issue "requests" for concepts to ontologists. Second, the establishment of a set of guidelines for ontologists to help them decide:

- what to add as a concept; what not to add;
- where to place the concept;
- how to name a concept;
- how to rearrange a part of the ontology; and
- possible actions for lexicon request.

These guidelines are said to collectively define the methodology for ontology acquisition, and keep the ontology development process as uniform and consistent as possible. In general, they are ad hoc rules of thumb based on experience rather than on a priori studies of semantic universals. Butler (this volume) suggests that perhaps such a study would be necessary to overcome the language bias problem and points towards the work of Wierzbicka (1988, 1992, 1996) on universal human concepts. Our position is that, whilst such an initiative may well prove fruitful, it remains to be seen whether a flat collection of such semantic universals can be successfully integrated in a fully structured ontology such as the one we have described. Also since this list is, just like the Mikrokosmos ontology, also being modified in successive work in the light of new evidence, we cannot really see what practical advantage it would have.

In connection with the linguistic bias problem, we also encounter problems when we try to employ techniques borrowed from purely linguistic or lexicographic methodologies. We will try to illustrate the type of issue involved here with reference to the analysis of the verb *catch* that Butler (this volume) presents. After offering a detailed description of the various senses found in this lexical item through both exhaustive context analysis and informants' observations, he attempts to confront these different, but related, senses with a candidate lexical template for this English verb. He then concludes that this template would fail to capture all the uses of *catch* on the basis that non-physical, but statistically relevant uses of the verb, such as 'catch a disease' or 'catch a means of transport', would fall outside the descriptive power of the template.

Let us now face the question from the point of view of the ontologist, whose task is generalizing abstract meanings, independently of how these are lexicalized in different languages. From this perspective, these senses of the English verb *catch* do in fact correspond to totally different concepts, which just happen to have been agglutinated in one lexical item in English.

This type of linguistic analysis is in fact a recognized pitfall for the ontologist. Mahesh (1996: 55) provides similar examples when discussing the semantics of the concepts SWIM and FLOAT and the related lexical items in English, Spanish and Russian, and provides examples such as "The meat is swimming in gravy" or "John floated the raft out of the dock" as the kind of evidence that the linguist or lexicographer seeks in order to come up with the various senses of words. The consequence of employing these strictly linguistic techniques when constructing an ontology is that the resulting ontology will be specific to the language in question. From the ontologist's viewpoint, it is totally irrelevant that the English verb *catch* displays a fairly complex set of finely related senses that cannot be captured in one concept (or one lexical template, for that matter), because the lexical descriptions of these different senses will be carried out in the lexicon in as many senses as necessary, mapping each to different concepts and specializing them by constrained mappings as needed. For example, the sense ‘catch a disease’ can be ontologically defined as a DISEASE-EVENT with a BEGIN phase modifier.

It may be argued that such an approach fails to capture generalizations and common alternations found in many languages. However, the aim of ontological semantics, as a theory of computational semantics, is not the explanation of linguistic phenomena, nor does it claim typological adequacy of any kind; its aim is to be able to account for and manipulate linguistic objects in NLP applications. Within the field of computational semantics, enumerative lexicon approaches such as this have also been criticized on the grounds that it is possible (and more appropriate) to generate related senses rather than enumerate them (Pustejovsky 1995). Nirenburg & Raskin (forthcoming) argue that, while it is possible to generate such senses at run time, the result in terms of performance is no better than the one rendered by a high quality enumerative lexicon, the latter being free from the computing overhead derived from a generative approach.

5.2 Lexicon acquisition

A different kind of problem comes up in the lexicon acquisition process. Unlike ontology acquisition, which departs from abstract meaning, lexicon acquirers are faced with an unstructured list of lexical items in some language and no clue as to what concepts should be used for mapping the different senses of each lexical item. In general, the approach is the following:

1. Find a concept in the ontology as close in meaning as possible as the sense denoted by the lexical item. The meaning of this concept may be more general than that of the lexical item, i.e., it may subsume it, but it cannot be more specific.
2. Assign the lexical item to that concept.
3. If necessary, further specify that concept by means of constrained mapping, the base concept.

From the point of view of the lexicon acquirer, therefore, the two most important problems are, first, the identification of the appropriate base concept to assign lexical items to, and, second, the selection of those properties necessary for the constrained mapping of the chosen base concept from the set of available properties. During each of these stages a number of guidelines are also to be followed, but these are few and loose, and the semantic acquisition of each sense tends to become a challenge resulting too often in requests for, possibly arbitrary and language-motivated, ontology additions or modifications.

Finally, another aspect of lexicon acquisition that calls for attention is the syntactic description of lexical items. Due to the enumerative nature of the approach and the strict separation between language-independent and specific knowledge, individual syntactic descriptions must be entered for each sense of each lexical item. The type of generalization contained in lexical templates, would then, in principle be of no use within ontological semantics. However, the acquirer's task would be very simplified if they had such a template at hand. This does not necessarily imply having to acquire the lexicon from an onomasiological point of view, since the list of lexical items described by lexical templates could be indexed in several ways for access during acquisition.

To summarize what we have discussed in this section, by separating, in terms of representation, language-independent and specific semantics, we obtain an extremely powerful representational framework. On the other hand, this separation is largely responsible for many difficulties in the acquisition process. The current methodology is to consider the acquisition process as a whole, that is, as an interactive process between ontology and lexicon acquirers: a heuristic process that should lead towards a full specification of both ontology and lexicon acquisition guidelines.

The generalizations, both syntactic and semantic, contained in lexical templates seem to us suitable for easing out these lexicon acquisition problems because they define an entire class of semantically related lexical items and, by iconicity, the relevant set of syntactic alternations.

6. Conclusions

The adoption of an ontology involves the adoption of a number of assumptions and implied methodologies, as well as the rejection of others. Incorporating a language-neutral ontology such as Mikrokosmos will inevitable have an important impact on the linguistic model as a whole. This is both because of the content of the ontology as well as the its form. The Mikrokosmos ontology is content- rather than formalism-driven, but even so, its axioms impose a number of important assumptions. These ontological assumptions determine not only *what* can be said about a particular state of affairs (represented by a given predicate of a natural language), but also *how* it must be said. It is naïve to assume that, because of its engineering-driven approach, a self-contained resource such as this will lend itself to any linguistic model regardless of the semantic assumptions made by that model. It also seems clear that some of these axioms impose a number of important restrictions on the semantic treatment of language.

If this integration is to take place effectively, it appears then necessary to take steps towards establishing the precise impact that the integration of such a resource will have on the linguistic theory, even if the resource is to be used as a formalization of certain elements of that theory. We do believe that this integration is both possible and fruitful. Ontological semantics has been described here as a theory of meaning built on the methodologies and assumptions imposed by the use of an ontology, but it is not the only possible descriptive framework of natural languages that can use that ontology.

From the other side of the fence, RRG-style lexical templates, as proposed by Faber & Mairal (this volume), can be used to characterize the meaning definitions of all the verbs in the same lexical class. From the point of ontology and lexicon acquisition, these lexical templates offer very useful information in at least the following ways:

- they may help to see what semantic features must be inherited in the hierarchy;
- they may assist in the process of assigning several lexical entries to the same concept in the ontology;
- they help the ontology acquirer to decide when a new concept should be added to the ontology; and,
- they offer guidance in the lexicon acquisition process.

Therefore, both acquisition tasks can benefit enormously from the kind of lexical templates proposed by Faber & Mairal (this volume), which, understood in ontological terms, are indeed quite appropriate for the systematic description of the lexicon of a language, while providing a stable platform from which to construct the resources targeted by ontological semantics.

References

Dik, S. C.
1997 The Theory of Functional Grammar, Part 1: The structure of the clause. (ed. K. Hengeveld) Berlin & New York: Mouton de Gruyter.
Faber, P.—R. Mairal Usón
1999 *Constructing a Lexicon of English Verbs*. Berlin & New York: Mouton de Gruyter.
Farwell, D.—S. Helmreich—W-Y. Jin—M. Casper—J. Hargrave—H. Molina-Salgado—F-L. Weng
1994 PANGLYZER: Spanish Language Analysis System. Proceedings of the 1st Conference of the Association for Machine Translation in the Americas, 56-64.
Gruber, T.R.

1992 Ontolingua: a Mechanism to Support Portable Ontologies. Report KSL 91-66. Stanford University.

Gruber, T.R.
1995 Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal of Human and Computer Studies*, 43(5/6): 907-928.

Guarino, N.
1998 Formal Ontology in Information Systems. In N. Guarino (ed.) *Formal Ontology in Information Systems*. Amsterdam: IOS Press

Fillmore, C. J.
1982 Frame semantics; in Linguistics in the Morning Calm, Hanshin Publishing Co., Seoul, South Korea, 111-137

Fillmore, C. J.
1985 Frames and the semantics of understanding; in Quaderni di Semantica, Vol 6, No. 2: 222-254.

Jiménez Hurtado, C.
2001 *Léxico y Pragmática*. Frankfurt: Peterlang.

Lehrer, A.—E. F. Kittay
1992 *Frames, Fields and Contrasts*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Mahesh, K.
1996 Ontology Development for Machine Translation: Ideology and Methodology. Memoranda in Computer and Cognitive Science MCCS-96-292, New Mexico State University: Computing Research Laboratory.

Mahesh, K.—S. Nirenburg
1995 A situated ontology for practical NLP. A Paper Presented at the IJCAI '95 Workshop on Basic Ontological Issues in Knowledge Sharing, August 19-21, Montreal, Canada.

Martín Mingorance, L.

1998 *El Modelo Lexemático-Funcional*. (The Functional-Lexematic Model.) (Ed. A. Marín Rubiales). Granada: University of Granada.

Minsky, M.
1975 A Framework for Representing Knowledge, in P. Winston (ed.) *The Psychology of Computer Vision*. New York: McGraw-Hill. 211-277.

Moreno Ortiz, A.
1997 *Diseño e implementación de un lexión computational para lexicografía y traducción automática*. PhD thesis. Published in *Estudios de Lingüística Española*. Vol. 9. http://elies.rediris.es/elies9.

Nirenburg, S. (ed.)
1994 *The Pangloss Mark III Machine Translation System*. A Joint Technical Report by NMSU CRL, USC ISI, and CMU CMT.

Nirenburg, S.
2000a CAMBIO: Progress Report. Working Paper, NMSU CRL.

Nirenburg, S.
2000b CREST: Progress Report. Working Paper, NMSU CRL. Presented at the DARPA TIDES PI Meeting, Chicago, October.

Nirenburg, S.— V. Raskin
Forthcoming *Ontological Semantics*.

Pustejovsky, J.
1995 *The Generative Lexicon.* Cambridge, MA: MIT Press.

Raskin, V.—S. Nirenburg
1995. Lexical Semantics of Adjectives: A Microtheory of Adjectival Semantics. Memoranda in Computer and Cognitive Science MCCS-95-288. New Mexico State University: Computing Research Laboratory.

Rich, E. —K. Knight
1991 *Artificial Intelligence*. London: McGraw-Hill

Quine, W. O.
1960 *Word and Object*. Cambridge, Mass.: The MIT Press.

Van Valin, R. D., Jr.—R. LaPolla
1997 *Syntax: Structure, meaning and function*. Cambridge: Cambridge University Press.

Vossen, P. et al.
1998 The EuroWordNet Top-Level Ontology. *EuroWordNet*, LE2-4003 Working Paper. Available in the EuroWordNet distribution.

Wierzbicka, A.
(1988) *The Semantics of Grammar*. Studies in Language Companion Series 18. Amsterdam & Philadelphia: John Benjamins.

Wierzbicka, A.
(1992) *Semantics, Culture, and Cognition: Universal human concepts in culture-specific configurations*. New York & Oxford: Oxford University Press.

Wierzbicka, A.
1996 *Semantic Primes and Universals*. New York and Oxford: Oxford University Press.