A Qualia-based Description of Specialized Knowledge Units in the Lexical-Constructional Model[1]

Arianne Reimerink, Pilar León Araúz and Pamela Faber
University of Granada

Arianne Reimerink és professora ajudant doctora de Traducció General i Especialitzada a la Universitat de Granada. Té un doctorat en Traducció i Interpretació i els seus principals interessos de recerca són la representació especialitzada del coneixement, la lingüística cognitiva i la traducció.

Pilar León Araúz és llicenciada per la Universitat de la Provença, la Universitat de Northúmbria i la Universitat de Granada, on actualment és professora substituta del Departament de Traducció i Interpretació. Els seus principals interessos de recerca són la terminologia i la representació del coneixement.

Pamela Faber és catedràtica de Terminologia i Lingüística Aplicada a la Universitat de Granada. És llicenciada per la Universitat de Carolina del Nord, la Universitat de París IV, i la Universitat de Granada, on dóna conferències al Departament de Traducció. És autora de diversos llibres i articles sobre la semàntica lèxica, la traducció i la terminologia.

Resum

EcoLexicon és una base de dades de coneixement sobre medi ambient basada en la idea de marcs semàntics. La informació que conté està estructurada coherentment dins de l'esdeveniment prototípic de domini, l'esdeveniment mediambiental (EE). S'hi ha definit un inventari tancat de relacions, tant a nivell intercategorial com intracategorial, que connecten els conceptes entre si. Això és la base per a una ontologia formal d'aquest àmbit que servirà per a finalitats computacionals, fer cerques o extreure informació automàticament. Les premisses teòriques de la terminologia basada en marcs, el lexicó generatiu i la gramàtica de construccions lèxiques, proporcionen un formalisme estricte que ens permet fer un pas endavant cap a l'ontologia formal.

Paraules clau: *terminologia basada en marcs; lèxic generatiu; gramàtica de construccions lèxiques; formalisme; ontologia*

Abstract

EcoLexicon is a frame-based knowledge base on the environment. The information it contains is coherently structured within a prototypical domain event, the Environmental Event (EE). At an intra- and intercategorial level, a closed inventory of relations has been defined that relates concepts to each other as well as to the EE. It will be the basis for a formal domain ontology which will serve computational purposes, enhance searches and allow for automatic information extraction. Theoretical premises from Frame-Based Terminology, the Generative Lexicon and the Lexical-Constructional Model provide a streamlined formalism that brings us one step closer to a formal ontology.

Keywords: *frame-based terminology; generative lexicon; lexical-constructional model; formalism; ontology*

Terminàlia 1 (2010): 17-25
DOI: 10.2436/20.2503.01.9 · ISSN: 2013-6692

1 Introduction

EcoLexicon is a frame-based multilingual knowledge resource on the environment. In its construction great care has been taken to develop an internally coherent system. At a macrostructural level, all knowledge extracted from a specialized domain corpus has been organized in a frame-like structure or prototypical domain event, namely, the Environmental Event (see Figure 1; Faber: 2007; León [et al.]: 2008; Reimerink and Faber: 2009).

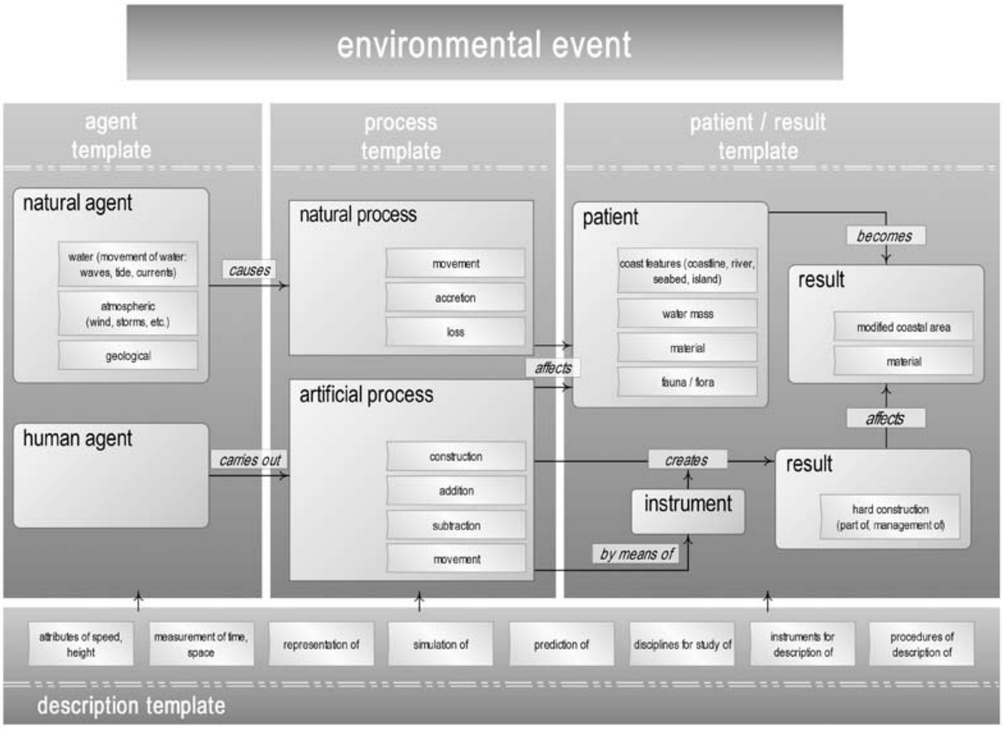

Figure 1. Environmental Event

The conceptual categories defined at this generic level are the broadest categories where all the concepts of the environmental domain can be included. The EE is conceptualized as a dynamic PROCESS that is initiated by an AGENT (either natural or human). This PROCESS affects a specific kind of PATIENT (an environmental entity), and produces a RESULT. These macro-categories (AGENT, PROCESS, PATIENT, etc.) are the concept roles characteristic of this specialized domain, which is clearly process oriented. Additionally, there are peripheral categories which include INSTRUMENTS that are typically used during the EE, as well as a category where the concepts of measurement, analysis, and description of the processes in the main event are included.

Since this knowledge base provides the foundation for an incipient linguistic ontology, the next logical step would be to convert the information in the knowledge base into a real domain ontology. This controlled knowledge structure would serve computational purposes, enhance searches and allow for automatic information extraction.

The first phase in this conversion is to find an elegant formalism capable of expressing the information in such a way that a computer can make sense of it. The formalism proposed in this paper is based on a combination of Frame-Based Terminology (FBT; Faber [et al.]: 2005; Faber [et al.]: 2007; Faber [et al.]: 2008), the Generative Lexicon (GL; Pustejovsky: 1995; Pustejovsky [et al.]: 2006), and the Lexical-Constructional Model (LCM; Ruiz de Mendoza and Mairal: 2006, 2007; Mairal and Ruiz de Mendoza: 2008).

In section 2 we explain how Pustejovsky's qualia are applied to the conceptual relations in EcoLexicon. Section 3 gives a short summary of the LCM and its application of qualia. Section 4 explains how LCM formalism could be applied to specialized knowledge units.

2 EcoLexicon and the Generative Lexicon

Pustejovsky and his colleagues define the Generative Lexicon (GL) as a theory of linguistic semantics which focuses on the distributed nature of compositionality in natural language (Pustejovsky: 1995; Lenci [et al.]: 2000; Pustejovsky [et al.]: 2006; Rumshisky [et al.]: 2006). GL describes lexical items according to their qualia structure, which constitutes the necessary modes of explanation for understanding a word or a phrase. It expresses the componential aspect of a word's meaning and is considered the meeting point of both argument and event structure. This is composed of the following roles:

1. Formal role: the basic type distinguishing the meaning of a word;
2. Constitutive role: the relation between an object and its constituent parts;
3. Telic role: the purpose or function of the object, if there is one;
4. Agentive role: the factors involved in the object's origins or "coming into being" (Pustejovsky [et al.]: 2006, 3).

GL and qualia structure have been successfully applied to the SIMPLE ontology, where an extended version of the qualia structure was developed (Lenci [et al.]: 2000) and in the creation of the Brandeis Semantic Ontology (BSO; Pustejovsky [et al.]: 2006). In GL, the information related to a lexical item consists of four levels: lexical typing structure; argument structure; event structure; and qualia structure. GL designates three major types: entity, event, and property. Each of these is in turn divided into three hierarchies: natural, artifactual, and complex:

1. Natural types: natural kind concepts with only formal and constitutive qualia roles;
2. Artifactual types: concepts with purpose, function, or origin.
3. Complex types: concepts integrating reference to a relation between types. (Pustejovsky [et al.]: 2006, 1).

In the construction of EcoLexicon, conceptual relations are associated with a particular qualia role, depending on each concept type. As a result, the macrostructure and microstructure of all concepts in the domain are represented in terms of these possible combinations (see Figure 2). The construction of the knowledge resource thus turns into a highly consistent and coherent process.

The most recurrent concepts of the domain (physical objects and processes) are the ones that can be linked to others through a greater number of relations. However, there are also certain relations exclusive of a single type, such as ATTRIBUTE_OF, for properties, and STUDY (for sciences and disciplines). For natural physical object types, apart from the relations traditionally linked to formal and constitutive roles, two non-hierarchical relations have been added: HAS_LOCATION and MADE_OF. The material that an object is made of or its location are key properties of subordinate concepts, and can even be the most essential feature. For instance, a GROYNE is not a groyne if it is not located in the sea.

The notion of *qualia* is also applied to the definitions of specialized environmental concepts in our knowledge base. Qualia make the knowledge base systematic both at the macrostructural level (the event) and the microstructural level (concept definitions).

In this respect, all definitions in EcoLexicon are based on a series of general templates for the description of generic concepts. For example, even though a PROCESS can activate all the relations shown in Figure 2, the prototypical definitional structure is constrained. A NATURAL PROCESS only activates the formal role, since this is the minimum information needed for description (see Figure 3). In contrast, an ARTIFICIAL PROCESS activates both the formal quale (the action itself) and the constitutive quale since artificial processes are generally composed of several steps or actions (see Figure 4). Furthermore, an artificial process always has a purpose (telic quale) and in certain engineering operations, an instrument may be used, which would also add the agentive role. All the information contained in these templates was extracted from a specialized domain corpus created for EcoLexicon (Faber [et al.]: 2006).

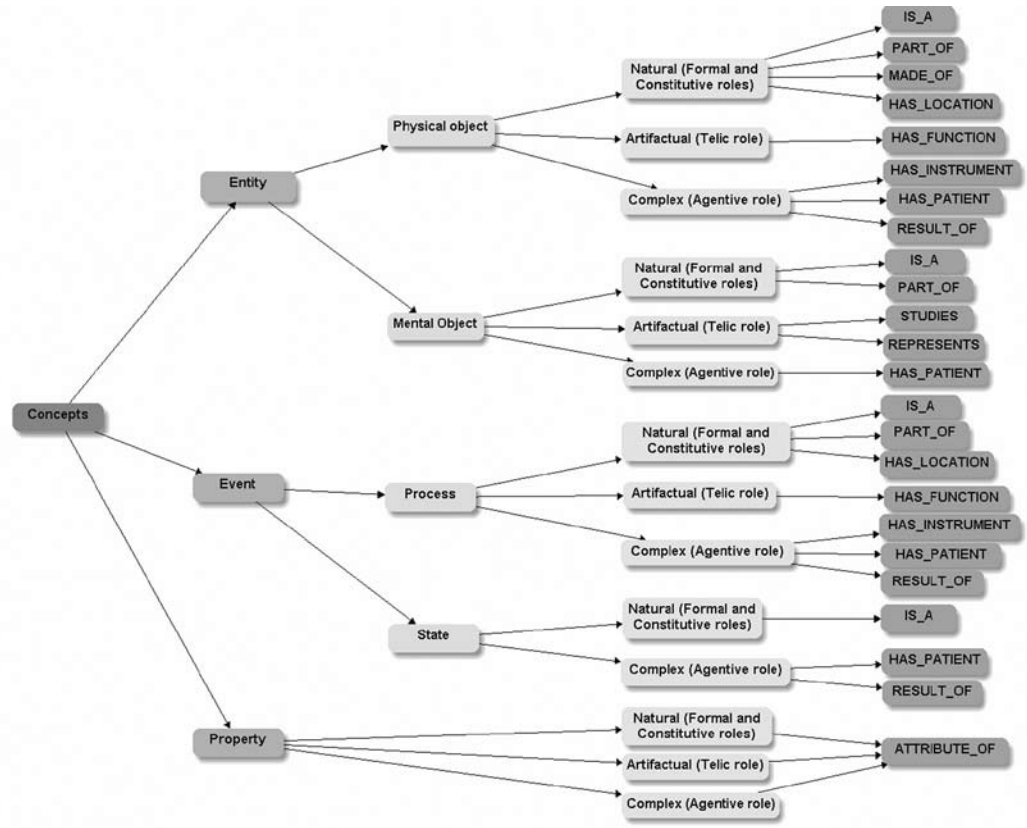

FIGURE 2. Combination of the concept typology and conceptual relations with Pustejovsky's qualia roles

NATURAL PROCESS: A succession of actions that happen or take place	
	■ FORMAL ROLE

FIGURE 3. Definitional template of NATURAL PROCESS

ARTIFICIAL PROCESS: A succession of actions and steps carried out for a specific purpose	
	■ FORMAL ROLE ■ CONSTITUTIVE ROLE ■ TELIC ROLE

FIGURE 4. Definitional template of ARTIFICIAL PROCESS

To explain how qualia structure is used to describe specific environmental processes, we will analyze the examples of EROSION and DREDGING. The definitions of EROSION and DREDGING can be segmented in terms of their qualia structure, and are derived from the general process template, although new qualia can be activated depending on their specificity. For example, a natural process may be initiated by an agent in the form of a natural force.

In the definitional template in Figure 5, EROSION is described as a natural process by which material is worn away from the earth's surface. Like all natural processes, EROSION does not have a function and therefore the telic quale is not part of its template. This is what differentiates natural and artificial processes. The template shows all the possible agents of erosion. This does not mean that all these agents have to be present in the process; the process involves at least one of them and can involve several.

EROSION		
FORMAL	[IS_A]	Natural process of reduction
AGENTIVE	[HAS_AGENT]	Gravity Water ▸ River ▸ Stream ▸ Rain Ice ▸ Glacier Wind Animals
	[HAS_PATIENT]	Earth's surface ▸ Beaches ▸ Mountains ▸ Soil ▸ ...

FIGURE 5. Qualia roles and definitional template of EROSION

In the subtypes of EROSION, such as SHEET EROSION, SPLASH EROSION, MASS WASTING, SLUMPING, etc., the specific agent involved is specified. All these subtypes follow the same template mapping back to the same formal quale although with different values. The process generally is of long duration, and consists of iterative sub-events. For example, the wind has to blow for a very long time and on repeated occasions in order to erode a cliff face. Since the process affects the entire surface of the Earth, Patient and Location coincide. Nonetheless, certain contexts refer to a specific Patient that is part of a bigger area, which can thus be considered the Location. This Location, however, is not specified in the definition, since the Patient dimension is more relevant.

As shown in Figure 6, the definitional template of the artificial process of DREDGING includes information regarding the action carried out, its phases as well as the instrument used, and its purpose.

The formal role includes two conceptual relations: IS_A and HAS_LOCATION. The IS_A relation expresses category membership and the HAS_LOCATION relation where the process takes place. DREDGING takes place underwater, but more specifically, it can occur under the water of rivers, canals, harbours, or offshore. These concepts are thus subordinate to underwater.

DREDGING		
FORMAL	[IS_A]	Artificial process of subtraction: removal
	[HAS_LOCATION]	Underwater ▸ Rivers ▸ Canals ▸ Harbours
CONSTITUTIVE	[HAS_PART]	Pumping Excavation Piping Material placement ▸ Sand placement
TELIC	[HAS_FUNCTION]	Construction Maintenance of water depths Beach nourishment
AGENTIVE	[HAS_PATIENT]	Solid material ▸ Sand
	[HAS_INSTRUMENT]	Dredger

FIGURE 6. Qualia roles and definitional template of DREDGING

The constitutive role reflects the phases of the dredging process. The last step, material placement, has a subordinate concept, SAND PLACEMENT, which restricts information as the context becomes more focalized. The same is true for the relation HAS_PATIENT in the agentive role. For example, in beach nourishment contexts, the material dredged can only be sand. The agentive role also includes the HAS_INSTRUMENT relation, since the dredger is one of the participants in the event, and in fact is the one that makes the process possible. Finally, the telic role expresses the three possible functions of DREDGING in three contexts with different degrees of specificity. In a general engineering context, dredging is used for construction purposes. Dredging is also used to maintain the navigability of channels and rivers. When the dredging process involves sand placement, the context is restricted to the very specific purpose of beach nourishment.

3 Lexical-Constructional Model and Qualia

According to Ruiz de Mendoza and Mairal (2007) and Mairal and Ruiz de Mendoza (2008), the Lexical-Constructional Model (LCM) provides a comprehensive description of the full inventory of parameters involved in meaning construction (idem: 2008, 137). This means that it is intended to be operational at all levels of linguistic description, including pragmatics and discourse. The authors provide a four level catalogue of construction types:

1. Constructions producing core grammar characterizations.
2. Constructions accounting for heavily conventionalized situation-based lower-level meaning implications.
3. Constructions that account for conventionalized illocutionary meaning.
4. Constructions based on very schematic discourse structures (Mairal and Ruiz de Mendoza: 2008, 138).

Level 1, called the argument module, is the result of the interaction between a lexical template and a constructional template. The lexical templates consist of three components:

1. A semantic component, which provides a set of primes (i.e. a set of basic terms or primitives that can be used to define the subordinate concepts in the same category).
2. A syntactic component, which consists of a series of lexical functions based on Mel'cuk's Explanatory and Combinatorial Lexicology (Mel'cuk [et al.]: 1995) that describe how the primes combine and define the whole set of predicates that converge within a lexical class (Ruiz de Mendoza and Mairal: 2007, 34).
3. A formalism to represent the combination of the semantic and syntactic components based on the logical structures of Role and Reference Grammar (Van Valin and LaPolla: 1997; Van Valin: 2005), enriched with the semantic component.

Constructional templates use part of the same metalanguage as lexical templates because constructions are an abstraction of what is common to a number of lexical items. Level 2 accounts for aspects of linguistic communication. Level 3 deals with the traditional illocutionary force. Finally, level 4 describes the discourse aspects of the LCM.

Recently, the LCM has incorporated Pustejovsky's qualia in their lexical templates to streamline the lexical description for future computational applications of the LCM (Mairal and Ruiz de Mendoza: 2008). The LCM basic representational format of a lexical template is based on a more systematic representation of the *Aktionsart* distinctions proposed in Vendler (1967), and the decompositional system is a variant of the one proposed in Dowty (1979):

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

Specifically, the lexical template of change of state verbs is the following:

predicate: [**do'** $(x, e_1)]_{E1}$ CAUSE [BECOME/INGR
pred' $(y)]_{E2}$

However, after reconverting the inventory of lexical functions by incorporating Pustejovsky's qualia, the lexical template of change of state verbs looks like this:

predicate:
EVENTSTR: [**do'** $(x, e_1)]_{E1}$ CAUSE [BECOME/INGR
pred'$(y)]_{E2}$
QUALIASTR: {QF: **MANNER pred'** (y); QA: e1: **Oper**
x, z <Instr>}

According to Ruiz de Mendoza and Mairal (2008, 367), change of state verbs (e.g. *break*, *smash*, *shatter*) are causative telic predicates; their event structure involves an activity and a final resulting state modified by a telic operator (BECOME or INGR). The state predicate is part of the formal qualia characterization of all change of state verbs. The semantic specificities of each predicate within the lexical class are expressed with the specific values ascribed to the semantic function MANNER. The causing activity event maps onto the agentive quale, as it expresses what is done by the Agent (x) in order to cause the Patient (y) to end up in the resulting state. The subevent e1 in the Agent quale describes the use of an instrument (z) by the Agent (x). The lexical function **Oper** is a semantically empty verb that will have different values depending on its arguments. Finally, the lexical template of *break* is as follows:

break:
EVENTSTR: **do'** (x, ∅)] CAUSE [BECOME/INGR *broken'* (y)]
QUALIASTR: {QF: *broken'* (y); QA: **do'** (x, *break_act'*)}

4 EcoLexicon, LCM, and Specialized Units

So far, the LCM has only dealt with verbs, whose templates are based on formalisms developed for several categories such as EXISTENCE, COGNITION, CHANGE OF STATE, CAUSED-MOTION, etc. Ruiz de Mendoza and Mairal (2007, 34), however, are aware of the fact that they have to expand their research to other grammatical categories. In the following section, we explore how the LCM can be applied to verbs as well as nouns in the specialized domain of the environment.

As explained in section 2, our definitions are based on templates. However, for ontology construction these templates must be converted into something more restricted such as the formalism proposed in the LCM.

Since the LCM has focused on verb meaning, our first attempt is to create a formalism for the verbs *dredge* and *erode*, two examples of caused-motion and change of state verbs, respectively, which are the most recurrent categories in the environmental domain. Then, we try to apply the LCM to nouns (*dredging* and *erosion*), both of which denote processes and involve the same entailments expressed by the verbs.

4.1 Caused-motion: the case of DREDGING, *dredge*, and *dredging*

Dredge is a clear example of a caused-motion verb, as it implies the movement of material (usually sand) from one place to another. Actually it is the change of location phenomenon what characterizes this construction. Ruiz de Mendoza and Mairal (2007, 38) give the following lexical template for caused-motion verbs:

predicate: do' (x, [**pred'** (x, y)]) CAUSE [BECOME NOT **be-in'** (y, z)]

This means that an Agent (x) causes an object (y) not to be in a place (z). The following sentence, extracted from our corpus, illustrates this basic template:

a. Many of the sediments (y) in tidal inlets (z) are dredged by hopper dredgers (x).

In (a) the argument (x) is filled with the instrument used in dredging operations. However, that argument is ultimately a human being, which is not necessarily mentioned in real texts. This is why in our corpus the argument structure is often restricted to Patient and Location, which is the core meaning of the verb (see Figure 7).

```
 A settling basin/trap may be dredged at the entrance of the ch
ourishment using relict sands (dredged from the seafloor at a
lts and carbonates, have been dredged from the walls and floors
 further treated - usually by dredging up offshore sand and pla
ch. Some nourishment sand was dredged from the entrance channel
use these  sand resources are dredged from waters owned by the
age sediment thickness of 3 m was dredged from the central port
\par  the beach. The material was dredged from New York Harbor
```

Figure 7. Patient, Location and *dredge*

Sometimes either the Patient or the Location is not explicitly mentioned in the text. The following examples only activate one argument (y or z):

```
 1904-1905 a channel was dredged through the pass but was
The Southern Channel was dredged in 1995, but  the present
amount of material to be dredged, relatively shallow water
hore  sediment source is dredged. In addition to concerns
ons (the study area was  dredged in 1995), by comparing th
ume of sediment that was dredged for the initial placement
e beach nourishment  was dredged and placed on the beach b
```

Figure 8. Patient or Location and *dredge*

A combination of the above information with the qualia structure and the template of caused-motion verbs can be designed as shown in Figure 9, where the formal role of *dredge* maps onto the template of its hyponym, the more basic motion verb, *remove*. The agentive role, apart from expressing the change of location notion, includes the instrument used through a lexical function (INSTR). In addition, the verb *dredge* implies the accomplishment of several phases expressed by the verbs *excavate*, *pump*, *pipe* and *place*. These phases take place at different times and are conveyed by verbs belonging to different paradigms. Faber and Mairal (2005, 29) provide a list of lexical functions of which INVOLV seems to be the most applicable to include them.

At the same time, in order to contextualize lexical templates in our specialized domain, arguments x, y and z are all filled with specialized concepts. In this way, their argument structure is also a part of the lexical meaning of specialized terms.

dredge:
EVENTSTR: **do'** [x, (**pred'** (x, y))]$_{E1}$ CAUSE [BECOME NOT **be-in'** (y, z)]$_{E2}$
QUALIASTR: {QF: REMOVE **dredged** (y)
QC: INVOLVE excavate, pump, pipe, place (y)
QA: BECOME NOT **be-in** (z), INSTR (x)}
x = dredger, human being
y = material, sand
z = underwater, offshore, river, tidal inlet, harbour, channels

Figure 9. Lexical template of *dredge*

However, in the noun *dredging*, collocates show new information that matches some of the definitional dimensions of Figure 6 (see Figure 10).

```
 to harbour facilities by dredging nourishment materials in
etter).  Main reasons for dredging include:    Increasing /
 remote sources. Offshore dredging can provide a good sourc
g  purposes. In Delaware, dredging is also used for obtaini
n as a coastal structure, dredging of sediment for   naviga
```

FIGURE 10. *Dredging*

As a specialized process in an engineering domain, the telic role found in its argument structure must also be included in the formalism. Consequently, a third event (E3, its purpose) has been added as a change of state construction, since it involves the improvement of a beach, channel, harbour, etc. (the same patients as those in the verb form, *dredge*). Apart from this third event, the formalism must clarify that the grammatical category of *dredging* is noun. A possible way of doing this is adding a grammatical category tag (GRAMTAG).

> **dredging**:
> GRAMTAG: noun
> EVENTSTR: **do'** [x, (**pred'** (x, y))]$_{E1}$ CAUSE [BECOME NOT **be-in'** (y, z)]$_{E2}$ CAUSE [BECOME (y)]$_{E3}$
> QUALIASTR: {QF: REMOVE **dredged** (y)
> QC: INVOLV excavate, pump, pipe, place (y)
> QT: PURP BECOME (y)
> QA: BECOME NOT **be-in** (z), INSTR (x)}
> x = dredger, human being
> y = material, sand
> z = underwater, offshore, river, tidal inlet, harbour, channels

FIGURE 11. Lexical template of *dredging*

4.2 Change of state: the case of erosion, erode, and erosion

In the EcoLexicon corpus, the concept EROSION is lexicalized in different grammatical categories: the verb *erode*, the noun *erosion*, the adjective *erosionable*, etc. The concordances extracted from the corpus in combination with the definitional template of the concept show that *erode* is a change of state verb. As previously mentioned, a change of state verb is composed of two events. In the first event (E1) an Agent carries out an action which causes a second event (E2). As a result of this second event, a Patient undergoes a change. One of the characteristics of change of state verbs is that they allow for the causative/inchoative alternation:

a. We broke the window
b. The window broke
c. The window breaks easily (taken from Ruiz de Mendoza and Mairal 2006: 29).

The corpus shows that same alternation for the verb *erode*. In Figure 12, the basic grammatical structure in which an Agent erodes a Patient matches the first example (a).

```
ar  Although a beach may be temporarily eroded by storm waves and later partly o
ocity controls the stream's capacity to erode and transport sediment through its
in shallow lakes as the thermo cline is eroded by strong winds. One might improv
s the action of tractive currents which eroded and reworked marginal lacustrine
  Bethany Beach. Longshore currents are eroding almost the  entire ocean coastli
ls are easily attacked, decomposed, and eroded by various chemical and physical
tration of sediment-deficient    water erodes the surface of the sheetflood dep
ing erosional terraces. Radial fans are eroded mainly by tides; there are no sed
d July 1988, cold-front-generated waves eroded the entire beach face and mused a
  rivers draining the plateau have also eroded gorge sections through the   esca
lled   longshore drift. Longshore drift erodes and deposits sand continuously al
tly connected to the Mediterranean sea, eroded part of the southern late Neogene
ome dry climate areas, persistent winds erode all sediments the size of sand and
aster (13-15 per minute) plunging waves erode the beach during storms. The most
differences in atmospheric pressure can erode surface material when velocities a
ot only on the capacity of the river to erode river banks, river bed   and to tr
, boulevards and revetments, the beach  eroded as a result of wave reflection. T
```

FIGURE 12. Agent, Patient and *erode*

In Figure 13, the alternation where a Patient erodes coincides with (b), where it can include an adverb as in (c):

```
tion.  Dunes are dynamic features; they erode during periods of high waves and a
tream beaver dams. These sediments were eroded when the dams were breached by sp
ng-term cyclic patterns, where they may erode for several years and then accrete
either delivered from upstream areas or eroded from the river bed. The imbalance
ting at  an extremely fast rate as sand eroded from the Atlantic  coastline is d
tween 1.02 and 6.85. Clearly, clays are eroded  from topsoils within the upstrea
res were evident, and the shoreline was eroding at the landward ends. The cover
ng spits and land tongues are currently eroding.  Coastal protection works are t
  further to the east, so they began to erode. The less than ideal solution was
 40-50 years. Not all of the bluffs are eroding at any given time. If the timing
ws  on  Fig.  17,  left).  The coast is eroded and much material is transported
 sand in   1995. As beaches continue to erode and retreat inland, beachrock ledg
 the beach, even the  foreshore started eroding. The erosion has increased at se
```

FIGURE 13. Patient and *erode*

As shown in the above concordances, *erode* can definitely be included in the category of change of state verbs. *Erode* is a causative telic predicate whose event structure involves an Action (movement of air, water or ice) caused by an Agent (waves, rain, wind, glacier) which causes a second event, resulting in a change experienced by a Patient (mountain, rock, alluvial fan). In this case, the change involves a reduction in size. Based on the LCM and the qualia description of change of state verbs (section 3), the new formalism is expressed in Figure 14.

> **erode**:
> EVENTSTR: [**do'** (x, Ø)]$_{E1}$ CAUSE [BECOME **reduced** ´(y)]$_{E2}$
> QUALIASTR: {QF: DEGRAD **reduced'** (y)
> QA: CONT E1}
> x = wind, water, ice, gravity, animals
> y = Earth's surface

FIGURE 14. Lexical template of erode

According to the LCM, change of state verbs permit two possible lexical functions in E2, BECOME or INGR, because something can change instantly (INGR) or little by little (BECOME). It is obvious that in the case of *erode*, as well as in caused-motion verbs like *dredge*, BECOME is the lexical function that must be applied. In *erode*, the basic idea is that something changes by becoming reduced (based on Faber and Mairal: 1999). *Reduce* is a higher level change of state verb that does not have the domain specific constraints.

The qualia structure of the verb *erode* should express the characteristics that differentiate the verb from other change of state verbs such as *break* or *reduce*. Firstly, the formal quale must convey the MANNER in which the change in the Patient comes about. From the list of lexical functions provided by Faber and Mairal (2005, 29), the following seem applicable to the case of *erode*: CONT, continuity/duration and DEGRAD, to get worse.

In *erode*, the manipulation subevent (e1), does not apply because *erode* is a natural process. Metaphorically speaking, we could say for example that the RAIN (Agent) uses GRAVITY (Instrument) to bring about a change in the EARTH'S SURFACE (Patient), but this kind of manipulation event seems to be more applicable to artificial processes, such as *dredge* and *dredging*. What is important in the first event (E1) is that the action implies a long time and a continuous process. Rain must fall on a rock for a long time for it to erode. In E2, the state of the Patient (y) changes in a specific way, namely, it diminishes or the Agent degrades the affected entity.

Another thing that must be taken into account is that the argument fillers x and y (Agent and Patient) cannot just be anything. The specialized domain in which the process EROSION, and therefore the verb *erode*, is included restricts the possibilities. The possible Agents for erode are: WIND, WATER, ICE, GRAVITY and ANIMALS, and all their subordinates. The Patient of *erode* is the EARTH'S SURFACE and all its subordinates.

As in *dredge*, this additional information should be included in the formalism. On the other hand, based on the fact that many of the possible Agents and Patients will also be applicable to other verbs of the Environmental domain, a list of possible Agents and Patients could be linked to the basic template of change of state verbs to avoid redundancy. This means that specialized terms should fill different arguments at the higher level of abstraction where they can occur. As a result, all verbs belonging to the same paradigm are able to activate the same arguments or their subtypes.

As for the application of the LCM to nouns, it must be highlighted that the semantic information contained in *erode* and *erosion* is the same for both lexical items, the concept EROSION. The possible Agents and Patients involved in its argument structure are the same as well, but only conceptually speaking. In the case of this procedural noun, the only thing that has to be done is to clarify in the formalism that it is not the expression of a verb, but of a noun (see Figure 15).

erosion:
GRAMTAG: noun
EVENTSTR: [**do'** (x, Ø)]$_{E1}$ CAUSE [BECOME **reduced'**(y)]$_{E2}$
QUALIASTR: {QF: DEGRAD **reduced'** (y)
QA: CONT E1}
x = wind, water, ice, gravity, animals
y = Earth's surface

FIGURE 15. Lexical template of *erosion*

However, although arguments (x, y) are the same from a semantic perspective, they do not have the same syntactic behaviour. For example, in the case of the verb, Agents will only occur in the form of a subject. However, in the case of the noun, Agents and even Patients can be codified in different ways, as in *aeolic erosion* or *beach erosion*.

On the other hand, *sheet erosion* is a type of erosion where raindrops detach soil particles of the Earth's surface. The formalism of sheet erosion would therefore contain the specification of the Agent (see Figure 16).

sheet erosion:
GRAMTAG: noun
EVENTSTR: [**do'** (x, Ø)]$_{E1}$ CAUSE [BECOME **eroded'**(y)]$_{E2}$
QUALIASTR: {QF: DEGRAD **eroded'** (y)
QA: CONT E1}
x = rain
y = Earth's surface

FIGURE 16. Lexical template of *sheet erosion*

5 Conclusions

The combination of Frame-Based Terminology, Generative Lexicon and the premises of the Lexical-Constructional Model can bring us closer to the construction of a formal domain ontology. The coherence and consistency of the information contained in EcoLexicon provides a sound basis for the development of a formalism. Pustejovsky's qualia have proved to be very useful for streamlining the information in our domain-specific knowledge base and for the lexical templates of the LCM. We have shown a possible way to apply both qualia and LCM formalisms to the description of specialized knowledge.

For now, we have analyzed some verbs and nouns that denote processes, which is the most important category in our domain. We are aware, however, that a lot remains to be done. Further research will be necessary to find out if the formalism can be applied to all the verbs and nouns that denote processes and to other conceptual and grammatical categories.

References

Dowty, D.R. *Word Meaning and Montague Grammar*. Dordrecht: Riedel, 1979.

Faber, P; Márquez Linares, C; Vega Expósito, M. "Framing Terminology: A Process-Oriented Approach". In *META*, 50, 4, 2005, CD-ROM.

Faber, P; Montero, S; Castro, R; Senso, J; Prieto, J.A; León, P; Márquez, C; Vega, M. "Process-oriented terminology management in the domain of Coastal Engineering". In *Terminology*, 12, 2, 2006, 189-213.

Faber, P; León Araúz, P; Prieto Velasco, J.A; Reimerink, A. "Linking images and words: the description of specialized concepts". In *International Journal of Lexicography*, 20, 2007, 39-65.

Faber, P; León, P; Prieto; J.A. "Semantic relations, dynamicity, and terminological knowledge bases". In *Proceedings of the XVIII FIT World Congress*.Shanghai: FIT, 2008, CD-ROM.

Faber, P; Mairal Usón, R. *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter, 1999.

Faber, P; Mairal Usón, R. "Decomposing Semantic Decomposition: Towards a Semantic Metalanguage in RRG". In *Proceedings of the 2005 International Conference on Role and Reference Grammar*. Taiwan: 26-30 June 2005.

Lenci, A; Bel, N; Busa, F; Calzolari, N; Gola, E; Monachini, M; Ogonowski, A; Peters, I; Peters, W; Ruimy, N; Villegas, M; Zampollo, A. "SIMPLE: A General Framework for the Development of Multilingual Lexicons". In *International Journal of Lexicography*, 13, 4, 2000, 249-263, doi: 10.1093/ijl/13.4.249.

León Araúz, P; Reimerink, A; Faber, P. "PuertoTerm & MarcoCosta: a Frame- Based Knowledge Base of Coastal Engineering". In *Proceedings of the XVIII FIT World Congress*. Shanghai: FIT, 2008, CD-ROM.

Mel'cuk, I.A; Clas, A; Polguère, A. *Introduction à la lexicologie explicative et combinatoire*. Louvainla-Neuve, Belgium: Duculot, 1995.

Mairal Usón, R; Ruiz de Mendoza Ibañez, F.J. "New challenges for lexical representation within the lexical-constructional model (LCM)". In *Revista Canaria de Estudios Ingleses*, 57, 2008, 137-158.

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

Pustejovsky, J; Havasi, C; Littman, J; Rumshisky, A; Verhagen, M. "Towards a Generative Lexical Resource: The Brandeis Semantic Ontology". In *Proceedings of LREC 2006*. Genoa, Italy: 2006.

Reimerink, A; Faber, P. "EcoLexicon: A Frame-Based Knowledge Base for the Environment". In *Proceedings of the International Conference "Towards eEnvironment"*, 25-27. Prague: March 2009.

Rumshisky, A; Hanks, P; Havasi, C; Pustejovsky, J. "Constructing a Corpus-based Ontology using Model Bias". In *Proceedings of FLAIRS 2006*. Melbourne Beach, Florida: 2006.

Ruiz de Mendoza, F.J. "On the nature of blending as a cognitive phenomenon". In *Journal of Pragmatics*, 30, 1998, 259-374.

Ruiz de Mendoza Ibañez, F.J; Marial Usón, R. "Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model". In *Folia Lingüística: Acta Societatis Linguisticae Europapeae*, 42, 2, 2008, 355-400.

Ruiz de Mendoza Ibañez, F.J; Marial Usón, R. "Levels of semantic representation: where lexicon and grammar meet". In *Interlingüística*, 17, 2006, 26-47.

Ruiz de Mendoza, F.J; Mairal, R. "High-level metaphor and metonymy in meaning construction". In Radden, G; Köpcke, K.M; Berg, T; Siemund, P (ed.). *Aspects of Meaning Construction in Lexicon and Grammar*. Amsterdam/Philadelphia: John Benjamins, 2007, 33-49.

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

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

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

Notes

1. This research is part of the project ECOSISTEMA: *Espacio úniCO de SIStemas de información ontológica y TEsaurus sobre el Medio Ambiente* (FFI2008-06080-C03-01/FILO), funded by the Spanish Ministry of Science and Innovation.