FrameNet and FunGramKB: A comparison of two computational resources for semantic knowledge representation[1]

Alba Luzondo Oyón (Universidad de La Rioja)

Rocío Jiménez Briones (Universidad Autónoma de Madrid)

Abstract

This chapter presents a comparison between FunGramKB, a multipurpose lexico-conceptual base for Natural Language Processing (NLP) systems, and FrameNet, a lexical resource for English whose objective is to document the range of semantic and syntactic combinatory possibilities of each sense of a word. After providing the reader with an overview of the two resources under scrutiny, we address their similarities and differences by focusing on the following issues: (1) methodology; (2) information at the lexical and conceptual levels; (3) relations between frames and concepts; (4) information management; and (5) multilingualism. To illustrate this comparison, we analyze how the verb *dry* is represented in each project.

Key Words: NLP, FrameNet, FunGramKB, frames, concepts.

1. Introduction

Meaning representation has always been a major issue in the field of linguistics. More

---

[1] The research on which this chapter is based has received financial support from the Spanish Ministry of Economy and Competitiveness, grants no. FFI 2010-17610/FILO and no. FFI2011-29798-C02-01.

recently, with the development of natural language processing (NLP) systems, attention has also been paid to computational resources capable of handling this kind of information, since semantic knowledge plays a fundamental role in solving tasks that involve successful reasoning. As Periñán and Arcas (2007a: 197) point out, some NLP systems (e.g. information extraction or automatic indexing) do not really require representations that mirror the structure and organization of the cognitive system human beings are endowed with. By contrast, other tasks such as machine translation or text understanding call for the incorporation of reasoning. Although it is true that surface semantics (Velardi et al. 1991) may be sufficient for the former type of systems, the construction of a solid knowledge base guarantees its use in most NLP task (cf. Periñán and Arcas 2007b: 279).

In this context, this chapter compares FunGramKB (Periñán and Arcas 2010a, 2010b; Mairal and Periñán 2010a, 2010b; Periñán and Mairal 2010), a multipurpose lexico-conceptual knowledge base for NLP systems, with FrameNet (FN; Petruck 1996; Baker et al. 1998; Fillmore et al. 2003a, 2003b; Boas 2005; Ruppenhofer et al. 2010), a lexical database that lately has also been involved in NLP applications (cf. Ovchinnikova et al. 2010: 3157). Consequently, by contrasting these two projects we aim to provide a detailed description of the ways in which a *deep* approach and a *shallow* approach to knowledge representation vary (Periñán 2013). As will be shown, FunGramKB, which formally describes the conceptual content of lexical units, is an example of the former approach. By contrast, in a shallow account like the one adopted in FN, the cognitive content of a lexical item is characterized “by means of a simple feature value matrix of conceptual relations” (Periñán 2013: 89). Thus, owing to the fact that FN and FunGramKB have been designed with different aims in mind, we should expect to find many aspects in which these resources differ. Nevertheless, FN and FunGramKB also share a similar view in other respects such as their treatment of polysemy, their use of organizational units larger than words, among others.

The reminder of this chapter is organized as follows. Section 2 provides the reader with a general overview of the two computational resources under examination. Section 3, which comprises the core of the study, details the similarities and differences between FN and FunGramKB with respect to the methodology employed (subsection 3.1), the information stored at the lexical and conceptual levels (subsection 3.2), the relations between frames and concepts that each computational resource posits (subsection 3.3), how information is managed by these projects (subsection 3.4), and their stances on the creation of multilingual resources (3.5). The case of *dry* will be used to illustrate this comparison. Finally, some concluding remarks are offered in section 4.

2. Main tenets of FN and FunGramKB

2.1. FN

The FN project builds on *Frame Semantics* (cf. Fillmore 1976, 1982; Fillmore and Atkins 1992; see also Boas 2005, 2009a: 15 and references therein). As the following quote evidences, the main idea underlying this theory is that word meanings provide access to (or are described as belonging to) organized pieces of knowledge, i.e. *frames*:

> A word's meaning can be understood only with reference to a structured background of experience, beliefs, or practices, constituting a kind of conceptual prerequisite for understanding the meaning. […] Within such an approach, words and word senses are not related to each other directly […] but only by way of their links to common background frames and indications of the manner in which their meanings highlight particular elements of such frames (Fillmore and Atkins 1992: 76-77).

Frame Semantics has been applied to the construction of FN, an online lexical resource "that extracts information about the linked semantic and syntactic properties of English words" from electronic text corpora like the British National Corpus (BNC) (Fillmore et al. 2003a: 235). According to Fillmore et al. (2012), FN aims to account for the syntactic and semantic properties of lexical units, paying especial attention to their meanings, their syntactic and combinatorial affordances, as well as their relation to frames.

There are two main units of analysis in FN, namely, frames and lexical units (LU(s)). The former, as characterized by Fillmore et al. (2003b: 305), are schematic representations of a situation type (e.g. "going to the movies", "drying something", "buying", etc.), which can be defined in terms of participants and their functions. A frame is in turn *evoked* by a set of LUs (i.e. words taken in one of their senses), which involves that the separate senses of a polysemous word are connected to different semantic frames. For example, the verbal predicate *run* evokes, among others, the frames labeled "Leadership" and "Self motion" as exemplified in *It is she who runs that business* and *We ran into the house to get Mame*, respectively.

In FN, valence information is specified both semantically and syntactically via: (1) *frame elements* (FEs) or the entities taking part in the situation depicted by a given frame, and (2) phrase types (e.g. NP, PP, etc.) and their corresponding grammatical functions (Subject, Object, etc.) (cf. Fillmore et al. 2003a: 236-237). FEs, which are frame-specific instantiations of broader semantic roles (e.g. AGENT, PATIENT, EXPERIENCER, etc.; see Boas 2010: 61), may either be core (i.e. conceptually prominent components of the frame) or non-core/peripheral (i.e. elements that do not uniquely characterize a frame), as shown in (1a) and (1b):[2]

(1) Frame: "Becoming dry" ("An entity loses moisture with the outcome of being in a dry state").

a. <u>Core FE</u>:

Entity [Ent]: "The thing that loses moisture and becomes dry".

b. <u>Some peripheral FEs</u>:

Place [Pla]: "The location where the Entity becomes dry".

Time [Tim]: "When the entity becomes dry".

Manner [Man]: "The way in which the Entity becomes dry".

According to Boas (2005: 143), one of the primary goals of this project is to document all of the lexicographically relevant uses of a word as found in the corpus, in addition to the syntactic environments in which they occur. To illustrate this point, we now examine how the information that a given verb can display is formally handled in FN. Let us consider the case of *dry* one more time. The lexical entry for one of the senses of *dry* comprises the following elements: (1) the FN definition given to this specific verb sense (i.e. "to lose moisture"), (2) the larger frame it evokes, i.e.

---

[2] According to Atkins et al. (2003: 268), there are two types of non-core FEs. The first kind includes adjuncts such as place, time, manner, purpose, etc. The second group comprises *extra-thematic* FEs (e.g. Iteration, Subregion, Frequency, etc.), i.e. those which do not conceptually belong to the frames they appear in, but which situate an event against a backdrop of another state of affairs (Ruppenhofer et al. 2010: 27-28).

"Becoming dry", (3) a list of core (i.e. Entity) and non-core FEs (i.e. Manner, Place, etc.) and their syntactic realizations (see Table 1), as well as the valence patterns in which they may occur (see Table 2 below).

Frame Element	Number Annotated	Realization(s)
Entity	(11)	NP.Ext (11)
Manner	(4)	AVP.Dep (4)
Particular_iteration	(1)	AVP.Dep (1)
Place	(1)	PP[on].Dep (1)
Time	(1)	PP[in].Dep (1)

Table 1: FEs for *dry* and their syntactic realizations.

In the middle column in Table 1 we are provided with the number of corpus-based annotated instances for this word sense, in which, for example, the core FE labeled Entity is expressed eleven times as a NP, which grammatically works as a Subject (so-called External argument).[3] For reasons of space, only two examples are supplied here:

(2) a. [Her tears]$_{\text{ENTITY}}$ DRIED$^{\text{TARGET}}$ as suddenly as they had appeared.

b. [The slab]$_{\text{ENTITY}}$ DRIES$^{\text{TARGET}}$ [very quickly]$_{\text{MANNER}}$ and faces virtually due south (…).

In much the same vein, the Place FE is found once as a PP with *on*, grammatically functioning as a Dependent:

(3) (A)llow [the excess water]$_{\text{ENTITY}}$ to DRY$^{\text{TARGET}}$ [on a flat surface between clean towels]$_{\text{PLACE}}$.

[3] According to Ruppenhofer et al. (2010: 89-90), constituents occupying core syntactic slots fulfill the functions of Subject and Object. The rest of constituents accompanying a syntactic head are considered Dependents.

Finally, Table 2 itemizes the valence patterns of this verb, that is, the different combinations of FEs and their syntactic realizations which may appear in a particular sentence (cf. Fillmore et al. 2003b: 330):

Number Annotated	Patterns		
4 TOTAL	Entity		
(4)	NP Ext		
2 TOTAL	Entity	Entity	Manner
(2)	NP Ext	NP Ext	AVP Dep
2 TOTAL	Entity	Manner	
(2)	NP Ext	AVP Dep	
1 TOTAL	Entity	Particular_iteration	
(1)	NP Ext	AVP Dep	
1 TOTAL	Entity	Place	
(1)	NP Ext	PP[on] Dep	
1 TOTAL	Entity	Time	
(1)	NP Ext	PP[in] Dep	

Table 2: Valence patterns for *dry* in "Becoming dry".

But besides the information discussed thus far, FN includes frame-to-frame relations of different types (cf. Boas 2010: 61), which we will address in more detail in section 3.3. Moreover, the role of "semantic types" (Ruppenhofer et al. 2010: 111-120) is that of recording those features that are not specified in their frame and frame element hierarchies. For instance, in the "Cause to be dry" frame, a semantic type labeled "sentient" delimits the nature of the Agent FE. This further clarifies that the Agent is necessarily a human entity (e.g. *Michael dehydrated the apples*), except for those examples in which the presence of the Cause FE (i.e. "the animate or inanimate entity, force or event that produces the state of dryness") excludes the Agent, as in *The wind dried the porch*.

Clearly, one of the main advantages of the FN project has to do with its manifest desire to offer a highly detailed account of phrase types and grammatical functions for each FE. For that matter, the linguist can benefit greatly from this lexicographic tool to indentify verb classes and to learn about the idiosyncratic semantic and syntactic behavior of an LU (see Boas 2011). Seemingly, the fact that the various FEs are highlighted in different colors in the list of annotated examples makes FN a user-friendly environment intended to facilitate the comparison of the often disparate syntactic behavior of verbs belonging to the same cognitive domain. However, although it is contended that, unlike commercial dictionaries, the set of examples displayed in the FN annotation reports "illustrate *all* of the combinatorial possibilities of a LU" (Ruppenhofer et al. 2010: 7, emphasis in the original), this is not always the case. At least for some frames, the sentences selected seem to instantiate the most typical combinatory affordances of an LU as attested in the corpus (cf. Fillmore et al. 2012), thus disregarding the occurrence of other, more or less frequent, configurations. Take some of the examples gathered from the annotation report of the predicate *drink*, respectively grouped under the labels *beer*, *bottle*, *coffee*, *cup*, etc.:[4]

(4) a. Jed DRANK two more beers.

b. They DRANK the whole bottle in perfect accord.

c. We DRANK coffee and talked some more.

d. He DRANK from the little blue-and-white cup.

Even though these sentences probably represent some of the most commonly employed collocations in which *drink* is embedded, this annotation report does not cover the occurrence of this predicate within, for example, the resultative construction. Surprisingly, the BNC retrieves realizations of the type *She drank*

---

[4] The interested reader is referred to Newman and Rice (2005) for a corpus-based study of *drink* and *eat* in terms of preferred objects of consumption, degrees of affectedness of the Object, the most favored Subjects, etc.

herself to death (BNC ABS), *She drank herself into stupidity* (BNC H7E), *They drank themselves blind* (BNC EUU), *You can drink yourself silly* (BNC KCT), *You have both drunk the barrel dry* (BNC HHC), etc. In principle, these realizations should have been incorporated into their annotation report, given that other instances of the resultative such as "freeze solid" or "push oneself free" can actually be found in other frames (i.e. "Change of Phase" and "Cause motion", respectively).[5] Consequently, if a construction grammarian wishes to study the ability of this lexical item to fuse with the resultative construction or the *way* construction (e.g. *I drank my way through the whole pot* (BNC FAP)), FN appears to fall short of this task. For this reason, the FN Constructicon is currently under development (Fillmore et al. 2012). Its explanation, however, goes beyond the scope of the present chapter.

2.2 FunGramKB

FunGramKB is a "user-friendly online environment for the semiautomatic construction of a multipurpose lexico-conceptual knowledge base for NLP systems" (Mairal and Periñán 2009: 219). FunGramKB is multipurpose in two ways. First, it has been designed to be reused in other NLP tasks, in particular in those that focus on language understanding (e.g. machine translation). Second, it faces the challenge of developing a multilingual tool that includes several western languages.[6] Most importantly, this project is constructed on the basis of a deep semantics approach (cf. Periñán and Arcas 2007b), since it actually defines each concept individually through

[5] The BNC offers a total of 8 examples for sentences containing "freeze solid", "freezes solid" or "froze solid". In turn, resultatives like "drank *(= someone) into" (e.g. *He drank himself into a stupor*), and "drink *(= someone) into" (e.g. *They drink themselves into a stupor on black PVC bar stools*) add up to total of 12 instances. It thus comes as a surprise that in the annotation report of *drink* these examples are not considered.
[6] Currently, English, Spanish and, to some extent, Italian, are fully supported. Work on the rest of languages (e.g. German, French, Catalan, etc.) is on its early stages.

a machine-readable metalanguage called COREL (i.e. *Conceptual Representation Language*).

One of the central methodological assumptions sustained by FunGramKB is that there is a division of labor between the linguistic level (i.e. the various Lexica and the Grammaticon) and the conceptual component (namely, the Onomasticon, the Cognicon and the Ontology), which in Figure 1 is represented with a dashed line:[7]

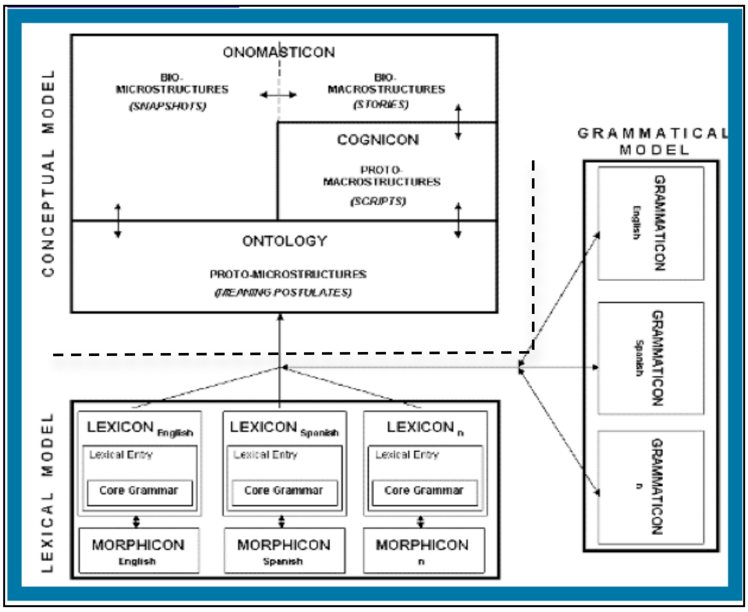


Figure 1: The architecture of FunGramKB and its different modules.

As can be seen in Figure 1, FunGramKB comprises three knowledge levels consisting of several independent, albeit interrelated, modules:

1. The lexical level covers what is commonly referred to as linguistic knowledge, which thus becomes language-dependent. In this module, the various lexica store information regarding morphosyntactic and pragmatic factors, as well as any possible collocational preferences of a given lexical unit. Likewise, some of the basic assumptions of Role and Reference Grammar (RRG; Van Valin and LaPolla 1997;

---

[7] In the Cognicon procedural knowledge in the form of conceptual macrostructures is stored (e.g. how to christen a baby). In turn, the Onomasticon stores information about instances of both events and entities (e.g. Johnny Cash, the Himalayas, or 9/11). In this chapter, only the Ontology will be discussed in detail.

Van Valin 2005) are preserved (i.e. logical structures, macroroles and the linking algorithm). In turn, the role of the Morphicon is to handle cases of inflectional morphology.

2. The grammatical level, which is also language-specific, stores the constructional schemata helping RRG to build the semantics-to-syntax algorithm. In the Grammaticon, the four constructional layers of the Lexical Constructional Model (LCM; Ruiz de Mendoza and Mairal, 2008, 2011; Mairal and Ruiz de Mendoza, 2009; Ruiz de Mendoza, fc.) are being implemented computationally.[8]

3. The conceptual level, which will be the focus of our attention, deals with non-linguistic (i.e. language independent) knowledge. Within this level, the Ontology, (i.e. a hierarchical structure of conceptual units), functions as the pivot around which the whole knowledge base revolves. For that matter, FunGramKB adopts a conceptualist approach which brings with it the acceptance of two major principles. Firstly, given that the conceptual level is the pillar that supports the different lexica, the semantic description of lexical items must rest upon the Ontology (Mairal and Periñán 2010b: 160). Consequently, FunGramKB is populated in a top-down fashion, since, in order for the description of a lexical item to be possible, there should previously exist a concept in the Ontology to which such lexical piece can be linked. Secondly, we cannot but stress that whereas the different lexica and the Grammaticon are idiosyncratic, the conceptual module, although being linguistically motivated, is nonetheless language independent. This entails that, while computational lexicographers will work on one Lexicon and one Morphicon for each of the

[8] The four constructional layers or levels of the LCM are the following: level 1 deals with argument structure constructions; implicational constructions are part of level 2; level 3 is devoted to illocutionary constructions; and finally, level 4 is concerned with discourse analysis. The interested reader is referred to Periñán and Mairal (2009) for more information on the relation between FunGramKB and the LCM.

languages currently supported, there is only one Ontology, one Onomasticon and one Cognicon to process any (culturally similar) language input.

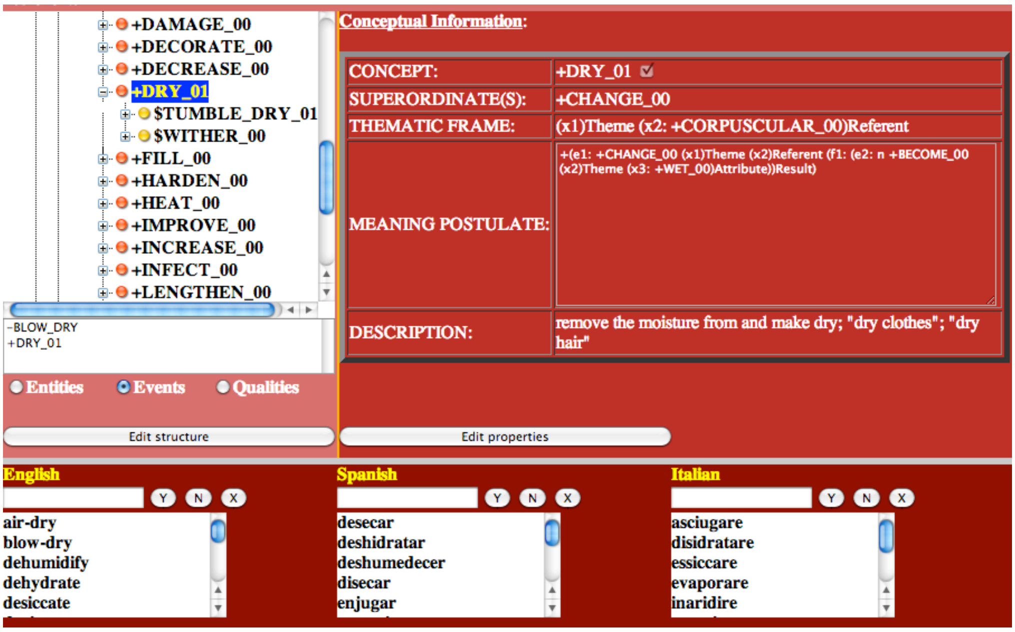
There are three conceptual levels in the Ontology of FunGramKB, i.e. *metaconcepts, basic* concepts and *terminal* concepts (cf. Periñán 2013 for more details). Metaconcepts (e.g. #ABSTRACT, #MOTION, #COMMUNICATION, etc.), which represent different cognitive dimensions, are the most abstract level in the hierarchy. In turn, basic concepts (e.g. +DIRTY_00, +DANCE_00, +BOOK_ 00, etc.) are the building blocks on the basis of which knowledge engineers define the meaning of a basic and/or terminal concept via Meaning Postulates (MPs). At last, the lowest level of the FunGramKB Ontology is made up of terminal concepts (e.g. $MOP_00, $GRASP_00, $SPORT_00, etc.), which contrast with basic concepts in that they cannot appear in MPs because they lack defining potential.[9] But essential to our discussion here is the fact that basic and terminal concepts are provided with semantic properties, namely, TFs and MPs. In the Ontology, events (e.g. +DRY_01, +SWIM_00, +ANNOY_00), are assigned one TF, i.e. "a prototypical cognitive construct which states the number and type of participants involved in an event" (Periñán and Arcas 2007a: 200).[10] Partially inspired in Halliday's (1985) and Dixon's (1991) work, participants in FunGramKB, also called *thematic roles*, are expressed through indexed variables in the TF (i.e. (x1) (x2) (x3), etc.) and their nature is usually (although not compulsorily) delimited by basic concepts (e.g. +HUMAN_00, +ANIMAL_00, etc.) referred to as *selectional preferences*. For example, the TF of the

---

[9] There is one exception to this general rule. As shown in (13), terminal concepts can appear in an MP when the event to be defined and the instrument used are the same (e.g. $TUMBLE_DRY_01).

[10] Although this chapter revolves around the analysis of events (#EVENT), one of the Subontologies of FunGramKB, there are two additional subcomponents, namely, #ENTITIES (e.g. +BIRD_00, +SOUL_00, +FREEDOM) and #QUALITIES (e.g. +AFRAID_00, +HAPPY_00, +ALONE_00). While #ENTITIES and #QUALITIES correspond to nouns and adjectives respectively, #EVENTS account for verbs.

basic concept +SWIM_00 (located within the metaconcept #MOTION), contains five cognitively necessary participants: (x1)Agent, (x2)Theme, (x3)Location, (x4)Origin and (x5)Goal. While Agent, Origin and Goal may be left unspecified, both the Theme and Location roles can be further spelled out by introducing the selectional preferences (x2: +HUMAN_00 ^ +ANIMAL_00)Referent and (x3: +WATER_00)Location.[11] As a result, the TF of +SWIM_00 describes a prototypical cognitive scenario in which a human or an animal moves in the water from an origin towards a destination. It is clear, however, that by merely stating the TF of a concept, we are not taking account of its meaning. Thus, the importance of MPs may now become apparent to the reader. According to Mairal and Periñán (2009: 224), MPs are sets of one or more logically connected predications (e_1, e_2, e_3, etc.), that is, conceptual constructs carrying the generic features of concepts. For instance, it is through the MP of +DRY_01 in Figure 2 that its conceptual meaning is represented in the Ontology. MPs are in turn lexicalized by various LUs in the languages currently supported, as shown in the boxes located at the bottom of Figure 2.



---

[11] For an exhaustive account of the identification of selectional preferences in FunGramKB and its possible implications for the RRG logical structures, see Jiménez and Pérez (2011).

Figure 2: +DRY_01 and its related LUs.

Let us explain the different components of Figure 2. We first observe that +DRY_01 is a subordinate of the parent concept or superordinate +CHANGE_00. Through inheritance –one of the reasoning mechanisms connecting concepts in FunGramKB (cf. section 3.3)–, the subordinate concept +DRY_01 shares the MP of the superordinate one, i.e. +CHANGE_00. More concretely, +CHANGE_00 (i.e. the *genus*) must be specified in the first predication (e_1) of all of its subordinate concepts (cf. *e1: +CHANGE_00 (x1)Theme (x2)Referent*). Furthermore, the MPs of every subordinate concept need to contain a distinguishing feature (i.e. *differentia*) not present in the MP of the more generic or superordinate concept. In FunGramKB, such semantic distinctions or differentiae are codified through *satellites* (e.g. Manner, Purpose, Result, Instrument, etc.; see Periñán and Mairal 2010 for the complete list), such as (*f1: (e2: n +BECOME_00 (x2)Theme (x3: +WET_00)Attribute))Result*) in Figure 2 above. Hence, the final representation of this concept reads as follows: "an entity (x1) changes a three dimensional entity (x2: +CORPUSCULAR_00) and as a result (satellite f1), (x2) will not be wet". To conclude, words such as *dry, dehydrate*, Spanish *desecar*, or Italian *asciugare*, are some of the LUs lexicalizing this concept. The description of each of the LUs attached to the concept +DRY_01 is specified at the lexical level (cf. section 3.2).

Having provided the reader with an overview of FN and FunGramKB, we are in a position to compare them in terms of the way in which each resource handles semantic knowledge. Likewise, we tangentially discuss the implications for the creation of NLP tasks.

3. Comparing a lexically-oriented and a conceptually-driven approach

In this section, we address the similarities and differences between FN and FunGramKB by paying attention to the following issues: (1) methodology; (2) information at the lexical and conceptual levels; (3) relations between frames and concepts; (4) information management; and (5) multilingualism. To illustrate this comparison, we concentrate on the representation of *dry*, which in FN would correspond to the frames "Becoming dry" and "Cause to be dry". As noted in (1), the former frame is defined as "an entity loses moisture with the outcome of being in a dry state". In such a scenario, the Entity FE is a core constituent, whereas the rest of FEs (i.e. Circumstance, Concessive, Degree, Manner, Particular iteration, Place and Time) are peripheral. This frame inherits from a frame labeled "Becoming" and is inchoative of "Being dry". "Becoming dry" is evoked by the following verb senses: *dehydrate, dry up, dry* and *exsiccate*. In turn, the causative counterpart of this frame, i.e. "Cause to be dry", is described as "an Agent causes a Dryee (either a surface or an entire entity, inside and out) to become dry". Here, Agent, Cause, and Dryee are core FEs, whereas notions such as Degree, Duration, Instrument, Manner, Means, Place, Purpose, Subregion, Temperature and Time are non-core. Moreover, this frame specifies two frame-to-frame relations: (1) it inherits from the "Transitive action" frame, and (2) it is causative of "Being dry". The verbal predicates *anhydrate, dehumidify, dehydrate, desiccate, dry, dry off, dry out*, *dry up*, and the nouns *desiccation* and *dehumidification* are defined against such a scenario.[12] As mentioned in section 2.2, FunGramKB approaches the case of *dry* through the basic concept

[12] These frames are available at https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Becoming_dry and at https://framenet2.icsi.berkeley.edu/fnReports/data/frameIndex.xml?frame=Cause_to_be_dry.

+DRY_01.

3.1. Methodology

We begin by pointing out two unquestionable similarities holding between FunGramKB and FN. First, both make use of organizational conceptual units that are larger than words, i.e. concepts and frames. Second, in both approaches the different senses of a word are assigned to different frames/concepts. For example, the "stop talking" sense of *dry up* in *Everyone became embarrassed and conversation dried up* (Longman dictionary) would not be part of the FunGramKB concept +DRY_01 examined herein.

FN and FunGramKB, however, seem to differ in other respects. When populating the FunGramKB Ontology, knowledge engineers invariably depart from a list of basic concepts, which is later on enriched through the addition of more specific terminal concepts and/or LUs from various languages. A thorough examination of dictionaries, thesauri, and corpora guides this process. In turn, FN begins by defining frame descriptions (based on corpus evidence) for the family of words to be analyzed (Boas 2009a: 16). The limited inventory of approximately 1,300 basic concepts in FunGramKB was identified by means of the defining vocabulary extracted from the *Longman Dictionary of Contemporary English* (Procter 1978). According to Periñán and Mairal (2011) and Boas (2009a), such a catalogue of basic vocabulary is a useful source for artificial language. Nevertheless, this reduced inventory has not been straightforwardly transformed into FunGramKB basic concepts. Instead, a deep revision was performed following the four-stepped COHERENT methodology, a label that stands for the phases termed COnceptualization, HIErarchization, REmodeling and refinemeNT (see Periñán and Mairal 2011 for more details). As such, a solid methodological approach supports the existing inventory of basic concepts on the grounds of which the subsequent creation of terminal concepts and analysis of related LUs will be carried out. By the same token, given that knowledge modeling is a fairly creative process, the population of the Ontology is guided by a protocol of seven ontological commitments, which work by preventing inconsistencies and reducing a number of common errors (Periñán and Arcas 2010a).

Going through the literature devoted to FN, we wonder about the strength of the methodology followed in FN. Drawing on Fillmore et al.'s (2003b) analysis of the "Attaching" frame, the following is a non-exhaustive summary of how the FN workflow unfolds (see also Atkins et al. 2003). The process begins with an informal description of the frame and an elaboration of a list of possible verbs that, according to the lexicographers' intuition, may evoke the frame. Dictionaries and thesauri are also consulted in this phase, along with the BNC, which is checked to extract sentences in which the central verb of the frame appears (e.g. *tie* in the "Attaching" frame) so that its syntax and semantics are fully accounted for. Once the initial analysis is completed, findings are compared with those of the *Concise Oxford Dictionary of the English Language* (COD). On the basis of examples and their informal characterization of the frame, the analyst can proceed to select and characterize the core and peripheral FEs of the frame. At last, with these FEs in mind, the lexicographer finally defines the frame under study. It should be emphasized that the steps described thus far are not rigid and modifications can be made at any given time, as much as it is possible to reevaluate and readjust concepts in FunGramKB.

At this point, however, a basic methodological question arises: it is somewhat unclear how the FN lexicographer arrives at a new frame. That is, if it is the case –as may be deduced from the above explanation– that frame characterizations depart from one LU (e.g. *tie*), we may ask ourselves on what grounds that verb is deemed to be a central member. Likewise, we gather from Fillmore et al.'s (2003b) description that, although corpus sentences are examined, some of the work is, to a certain extent, based on intuitive assessments. Eventually, initial conclusions are later on checked against the COD. This may, nevertheless, be insufficient in that sometimes dictionaries vary in the number of senses acknowledged to exist for a given verb and/or the way they account for their definitions (cf. Boas 2009b: 61 for a similar perspective). A case in point is that of *wither*, one of the hyponyms of *dry*:

(5) Cobuild (http://dictionary.reverso.net/english-cobuild/):

a. If somebody or something withers, they become very weak: *Industries unable to modernize have been left to wither.*

b. If a flower or plant withers, it dries up and dies: *The flowers in Isabel's room had withered.*

(6) Longman (http://www.ldoceonline.com/dictionary/):

a. If plants wither, they become drier and smaller and start to die.

(7) Oxford (http://oxforddictionaries.com/):

a. Of a plant, become dry and shriveled: *The grass had withered to an unappealing brown.*

b. Fall into decay or decline: *It is not true that old myths either die or wither away.*

c. Humiliate someone with a scornful look or manner: *She withered him with a glance.*

For instance, the Longman dictionary does not consider the figurative sense of *wither* expressed in (5a) or (7b). Likewise, neither the Longman nor the Cobuild provide an entry for the “humiliate” sense of this verb, which is, nonetheless, distinguished by the Oxford dictionary. These factors evidence the need to resort to more than one single dictionary in order to arrive at a more precise conclusion of the concept/frame to be described. For this reason, FunGramKB relies heavily on the use of dictionaries, thesauri and corpora. As such, more than three different English and Spanish dictionaries, together with the above-specified resources, are consulted in the process of creating terminal concepts and of assigning LUs to both basic and terminal concepts (see Jiménez and Luzondo 2011). It should be noted that, although we are by no means implying that the current frame classifications are inaccurate, there is a paucity of information as to how exactly the analyst arrives at a specific class of semantically related verbs (from which more general elements are abstracted away).

Nevertheless, in more recent work by Ruppenhofer et al. (2010), intuitive assessments appear to have been replaced by more explicit criteria:

> The core of the process has always been looking at corpus attestations of a group of words that we believe to have some semantic overlap, and dividing these attestations into groups. *In the past, the criteria for such grouping have been informal and intuitive, but recently, the criteria have become more explicit* (Ruppenhofer et al. 2010: 11, our emphasis).

Thus, according these authors, causatives and inchoatives should be split. Surprisingly enough, FN lexicographers do not appear to consistently apply this splitting approach to every LU. For example, *dry out* and *dry off* only evoke the “Cause to be dry” scenario, but not the inchoative one. Yet, online dictionaries provide users with both causative and non-causative definitions, as shown in (8) and

(9):

(8) *Dry out*

a. Longman: to become completely dry or to make something completely dry, especially after it has been very wet.

b. Cobuild: if something dries out or is dried out, it loses all the moisture that was in it and becomes hard.

c. Cambridge (http://dictionary.cambridge.org/dictionary/british/): to make something dry, or to become dry.

(9) *Dry off*

a. Longman: to become dry or to make something dry, especially on the surface.

b. Cobuild: if something dries off or if you dry it off, the moisture on its surface disappears or is removed.

c. Cambridge: to make someone or something dry, or to become dry, especially on the surface.

In FunGramKB, instead of selecting more than one *Aktionsart* for *dry*, we treat it in its basic form, i.e. a causative accomplishment, while its inchoative counterpart is accounted for through the selection of the inchoative construction.[13]

3.2. Lexical and conceptual information

Both FN and FunGramKB share the aim of offering detailed semantic and syntactic characterizations for each LU. The basic difference, however, is the way in which

---

[13] As shown in 3.2, the lexical level contains a list of argument structure constructions. The lexicographer working on a given LU selects the grammatical configurations in which such an LU may be embedded, since these serve as pointers to the constructional information stored in the Grammaticon (see Periñán fc.).

each project deals with this information. As previously noted, for FN this mainly entails creating and describing frames and FEs, listing possible frame-evoking items, recording the range of semantic and syntactic combinatory possibilities of each word sense, providing corpus-attested examples, establishing connections between frames, etc. From our point of view, the lexical level of FunGramKB, which is divided into three subsections (i.e. "morphosyntax", "LCM core grammar" and "miscellaneous"), covers a broader spectrum of information (see Mairal and Periñán 2009 and Guerra and Sacramento 2011 for a full explanation):

1. In the morphosyntax subcomponent, we specify the following information for the headword:[14] (1) possible graphical variants (e.g. *favor/favour*), (2) abbreviations (e.g. *personal computer*/*PC*), (3) the head in a phrasal constituent or idiom (e.g. *dry* in the phrasal verbs *dry out* or *dry up*), (4) the particle(s) of a phrasal verb and whether they can be detached from the verb, in the sense of allowing the embedding of an object (e.g. *out* is the detachable particle in *dry out*), (5) category (e.g. verb), (6) verb paradigm, which records features regarding the regularity and irregularity of the verb, along with any possible constraints in voice or tense, and (7) pronominalization information (i.e. issues related to clitization and reciprocality, if applicable).

2. The LCM core grammar subsection is only available for verbs. Here, the RRG *Aktionsart* distinctions are adopted, which, suffice it to say, have been proved to be valid crosslinguistically in languages as typologically different as Italian, Tagalog, Lakota or Japanese. Thus, for each verbal predicate we select one *Aktionsart*. In order to build the lexical templates of the verbs included in the Lexicon, we shall begin by identifying the number of prototypical arguments or variables that they take. *See*, for

---

[14] In this chapter, only information related to English verbs is supplied. It should be emphasized that other categories and languages will require the specification of different morphosyntactic data. To illustrate, English nouns must specify information about features such as their number and countability, and in the case of the Spanish lexicon, the gender of the noun needs to be recorded as well.

example, would be a two-place state predicate (i.e. **see'** (x, y)), *run* is a monovalent activity (**do'** (x, [**run'** (x)]), while *dry*, which takes two arguments (x, y), is a causative accomplishment ([**do'** (x, ∅)] CAUSE [BECOME [**dried'** (y)]). We then state the idiosyncratic features or macroroles assigned to the verb, namely, actor and undergoer. In particular, we need to provide the number of macroroles that the verbal predicate selects (0, 1, or 2) and determine which variable will function as undergoer.[15] The next step consists in linking the variables of a predicate to one and only one participant in the TF of the concept to which such lexical item is connected. This is known as "thematic frame mapping" (Periñán and Mairal 2009: 268).[16] As for the bivalent predicate *dry*, the *x* variable maps onto the Theme, while *y* maps onto the Referent. At this point, it is necessary to note that all this linguistic information (i.e. *Aktionsart*, predicate variables, macroroles and the thematic frame mapping) serves as the input for the automatic generation of a computationally-tractable structure labeled "Conceptual Logical Structure" (CLS; Mairal, Periñán and Pérez 2012; Periñán fc.), i.e. a lexically-oriented schema containing concepts through which the lexical, conceptual, and grammatical FunGramKB levels are connected.

Lexical templates also need to accommodate information about the idiosyncratic prepositions a verb may require and the collocations with which it frequently co-occurs. Hence, computational lexicographers will record those prepositions that obligatorily introduce the arguments of a verb (e.g. *on* for the *y*

---

[15] In RRG macroroles are groups of semantic roles which function as umbrella semantic notions covering other more specific thematic relations (Van Valin 2005: 60). For example, the actor corresponds to the most agent-like argument, whereas the undergoer could be identified with the most patient-like participant. Basically, they could roughly be equated with what other theories label *logical subject* and *logical object*, respectively. The incorporation of macroroles into the FunGramKB Lexicon turns out to be crucial in the linkage between the lexical and the grammatical modules, which consists in binding the variables of the lexical template of a verb with those of the specific constructions into which it is integrated. Such constructions are stored in the Grammaticon.

[16] It should be highlighted that not all the participants in the TF must be obligatorily mapped onto the variables of the LU.

argument of the verb *depend*), as well as their more regular collocations (e.g. the verb *dry* typically collocates, to list just a few, with nouns such as *tears*, *hair*, *eyes* or *hands*).

Adding to this, another subdivision within the LCM core grammar contains a list of grammatical constructions which the linguist selects on the basis of corpus evidence. For example, *dry*, regardless of frequency, may be embedded into the following constructions: caused-motion (e.g. *Her blood has already dried into a brown crust* (COCA, 2001)); inchoative (e.g. *He dried the coat >>The coat dried*); middle (e.g. *Silverrood dries well* (COCA, 1994)); or what Levin (1993: 80) labels the natural force subject alternation (e.g. *I dried the clothes in the sun >> The sun dried the clothes*). These very same syntactic configurations are also listed in the FunGramKB Grammaticon, in particular, in the Level 1 Constructicon (i.e. one of the repositories in the four constructional layers of the LCM), where their syntactic, semantic, and categorial information is formalized. Although a more detailed examination of the Grammaticon is beyond the scope of this chapter, we would still like to briefly comment on the rather sketchy and unequal treatment of constructions in FN. To the best of our knowledge, FN makes no explicit claims on how argument structure constructions are to be dealt with. For instance, we can infer that the incorporation of the extra-thematic FEs Beneficiary and Recipient (cf. Ruppenhofer et al. 2010: 140), which are constituents not introduced by the main evoked frame but by an external one, capture the traditional benefactive and dative constructions. However, the unexpressed object construction (*John eats at noon*) is indeed treated in Atkins et al. (2003: 269) as a construction, in particular as a case of Indefinite Null Instantiation (INI). In FN, INI points to an unexpressed core FE that needs no extra information to be fully understood because it is inferred from its surrounding context (as opposed to the Definite Null Instantiation or DNI) or through the grammatical structure in which it appears (unlike the Constructionally-licensed Null Instantiation or CNI). Finally, as mentioned in section 2.1 in relation to the resultative construction, the same construction is recorded as a valence pattern for some lexical units (e.g. *freeze*) but not for others (e.g. *drink*). Thus, in the light of the erratic treatment given to constructions in FN, we argue in favor of the unified account provided in both the FunGramKB Lexica and Grammatical modules. The former list the argument structure constructions into which each LU is embedded, whereas the latter describe them formally as machine-readable representations.[17]

3. At last, the “miscellaneous” subsection covers information regarding dialect, style (i.e. formal, informal, etc.), the knowledge domain in which a given LU may be located (e.g. architecture, medicine, anthropology, etc.), if any, and some representative examples extracted from dictionaries or corpora. Interestingly enough, FN is also recording these features by means of the use of semantic types, which can be equated to selectional preferences in FunGramKB but which, as Ruppenhofer et al. (2010: 111-112) stress, can function marking the types of fillers for FEs (e.g. Sentient, Degree, Manner, etc.), the types of frame (i.e. non-lexical), or the types of speech context differences of the LUs of a particular frame (e.g. Positive judgment, Negative judgment). Since an exhaustive list of these semantic types is not available and their annotation is not fully completed in all frames or LUs, we cannot venture to make an exact comparison between FN and FunGramKB in this respect. Suffice it to say that the semantic types employed in FN to classify the fillers of FEs could resemble

[17] That constructions are not given a central role in FN is also supported by the fact that they do not count as a factor to sort out words into different frames. As Ruppenhofer et al. (2010: 15-16) claim: “At all times, we lump together those groups whose semantic differences are due to general constructions of the language: passive, middle, tense/aspect constructions, composition with extra-thematic frame elements”. Ruppenhofer et al.’s (2010: 135-160) Appendix provides a detailed study of the independent status of extra-thematic FEs in FN along with their main types, functions, and the structures by which they are triggered.

FunGramKB's selectional preferences, with the proviso that the former had an ontological status and composed a closed class, as is the case of basic concepts. In fact, Ruppenhofer et al. (2010: 112) comment that "most of these [semantic] types correspond directly to synset nodes of WordNet, and can be mapped onto ontologies, e.g. Cyc." We wonder, however, how these semantic types (i.e. register, style, or evaluation of a particular LU) are going to be annotated in the lexical entry reports, since so far, apart from the dictionary definition, they only list the semantic and syntactic patterns of words (see Tables 1 and 2).

At this point, the reader may wonder how FunGramKB takes care of such vital issues as phrase types and grammatical functions, which as we have seen, is one of FN's strongest assets. Owing to the fact that FunGramKB is a knowledge base that has been developed with a view to creating NLP systems that demand language understanding, information related to phrases and syntactic functions is provided in the very first phases involved in text processing, namely, tokenization, morphological analysis, and syntax-semantics processing. As Periñán and Arcas (2011: 8-9) detail, the tokenizer divides the input text into sentences and orthographic words, whereas the morphological parser deprives words from their inflectional suffixes, which entails consulting the information stored in the Lexicon as well as in the Morphicon. Lastly, the syntax-semantics parser tags the parts of speech of the lemmas, assigning grammatical functions as well. At the end of this phase, the system outputs a structure where the lemmas have been replaced with concepts. Examples (10a)-(10d) illustrate the three phases responsible for text processing in FunGramKB (cf. Periñán and Arcas 2011: 9 for the rest of phases):

(10) a. Sally dried the shirts

b. Sally | dried | the | shirts

c. Sally | dry | the | shirt

d. S(NP(n(%SALLY_00)), VP(v(+DRY_00)), (NP(det(the), n(+SHIRT_00)))

In sum, FN and FunGramKB offer complementary types of information. FN exhibits a highly detailed account concerning the semantic and syntactic properties of lexical items. The thorough list of phrase types and grammatical functions, the variety of illustrative examples given, and the number of LUs recorded (12,715) are certainly FN's strongest points. In this respect, FN clearly surpasses FunGramKB. In FunGramKB the phases described in (10) are not visible for the user, nor is this conceptualist account concerned with offering a notable number of representative examples (see Table 3 below). Rather, its main focus is on processing any input text. Likewise, the number of LUs in FunGramKB, which amounts to a total of 3,707 in the English lexicon, is still scarce if compared to those already available in FN. We believe, however, that the FunGramKB Lexica provide the user with a broader spectrum of linguistic data than that furnished in FN, e.g. the morphological paradigm of a verb, its *Aktionsart*, macroroles, prepositions, collocations, and the associated grammatical constructions, apart from features related to dialect, style, etc. By way of example, Table 3 captures all the information that the FunGramKB user has access to for the LU *dry*:

Headword	dry
Concept	+DRY_01
Thematic Frame	(x1)Theme (x2: +CORPUSCULAR_00)Referent
Meaning Postulate	+(e1: +CHANGE_00 (x1)Theme (x2)Referent (f1: (e2: n +BECOME_00 (x2)Theme (x3: +WET_00)Attribute))Result)
POS	v
Paradigm	Regular
Tense Constraints	no
Reflexivity	Optional
Reciprocity	Never
Aktionsart	CACC
Lexical template: variables	X-Theme, Y-Referent
Lexical template: restrictions	MR2, U = y
Collocations	Y = eyes, hair, hands, tears
Constructions	Caused-motion construction (transitive) Inchoative construction Middle construction Natural force subject construction
Dialect	Standard
Usage	Common
Topic	FACTOTUM
Examples	"Her mouth had dried." (GUE W_fict_prose) "She dried her hands on the kitchen towel, adding, "I worry about Ernest." (FPM W_fict_prose)

Table 3: FunGramKB's lexical entry for *dry*.

Adding to this, FunGramKB goes one step further in that all this linguistic information depends on the conceptual module, the pivot around which this knowledge base revolves, which explains why that is the very first piece of information that appears in an entry (see row 2 in Table 3). To illustrate this point, consider once again the two frames discussed throughout this chapter (i.e. "Becoming dry" and "Cause to be dry"). At the beginning of section 3, we introduced the list of LUs evoking the aforementioned scenarios, here repeated for the reader's convenience. Verbs such as *dehydrate*, *dry up*, *dry* and *exsiccate* evoke the "Becoming dry" frame, whereas one of the senses of *anhydrate, dehumidify*, *dehumidification*, *dehydrate*, *desiccate*, *desiccation*, *dry*, *dry off*, *dry out* and *dry up* are connected to “Cause to be dry”. At least in the case of the two frames under scrutiny, FunGramKB currently posits a more exhaustive list of the hyponyms of *dry*, with the added advantages brought about by the adoption of a purely conceptualist approach, namely: (1) although the Ontology is linguistically motivated as a result of its involvement with the semantics of the predicates, the knowledge stored is not language-specific, contrarily to what is the case in FN; (2) because meaning definitions are provided at the conceptual level, management of multilingual knowledge is facilitated; (3) since concepts are described through the metalanguage COREL, a reasoning engine will actually be capable of processing and understanding the information of all the modules included in the knowledge base.

With this in mind, consider now the complete organization of this conceptual subdomain. +DRY_01 in (11) functions as the basic concept within which two other terminal concepts are contained, i.e. $WITHER_00 and $TUMBLE_DRY_00. Their definitions in COREL, together with their translation into natural language, are given in (12a) and (13a). Note too that several LUs from English, Spanish and Italian depend on each of these conceptual units:

(11) a. MP of +DRY_01: +(e1: +CHANGE_00 (x1)Theme (x2: +CORPUSCULAR_00)Referent (f1: (e2: n +BECOME_00 (x2)Theme (x3: +WET_00)Attribute))Result) = ‘An entity (x1) changes a three dimensional object (x2) so that as a result, (x2) will not be wet’.

b. Related LUs:

- English (En.): *air-dry, blow-dry, dehumidify, dehydrate, desiccate, drain, drip-dry, dry, dry off, dry out, dry up, exsiccate, torrefy*.

- Spanish (Sp.): *desecar, deshidratar, deshumedecer, disecar, enjugar, escurrir, orear, resecar, secar*.
- Italian (It.): *asciugare, disidratare, essiccare, inaridire, prosciugare, seccare*.

(12) a. MP of $WITHER_00: +(e1: +DRY_01 (x1)Theme (x2: +PLANT_00)Referent (f1: (e2: +BECOME_00 (x2)Theme (x3: +SMALL_00 | +WEAK_00)Attribute))Result (f2: (e3: ing +DIE_00 (x2)Theme))Result) = 'An entity (x1) dries a plant (x2) so that (x2) becomes small and weak and starts ('ing') to die'.

b. Related LUs:

- En.: *wither, shrivel, scorch*.
- Sp.: *agostar, ajar, marchitar, mustiar*.
- It.: *appassire, avvizzire*.

(13) a. MP of $TUMBLE_DRY_01: +(e1: +DRY_01 (x1)Theme (x2: +CLOTHING_00)Referent (f1: $TUMBLE_DRY_00)Instrument) = 'An entity dries clothes using a tumble drier'.

b. Related LUs:

- En.: *tumble dry*.[18]

As opposed to the way in which FN organizes the frames "Becoming dry" and "Cause to be dry", the basic concept +DRY_01 in FunGramKB is semantically decomposed into two terminal concepts (together with their related LUs). Thus, although FunGramKB has at its disposal a limited inventory of basic concepts to work

[18] As is clear from the case of $WITHER_00, the tendency sought in FunGramKB is to create terminal concepts to which several words in English, Spanish, Italian, etc. can be linked, so that the process of modeling conceptual meaning in the Ontology is not biased towards a particular language (Jiménez and Luzondo 2011: 21). However, lexical gaps among the different languages should not impede the creation of terminal concepts (e.g. $TUMBLE_DRY_01) "whenever there is at least one lexical unit whose meaning does not match any of the MPs stored in the knowledge base" (Mairal and Periñán 2009: 223).

with, the semantic knowledge stored in this resource is quite fine-grained. Besides, since such knowledge is formalized employing COREL, the descriptions are machine-readable, which means that the computer will have the capacity to discriminate between conceptually related events such as +DRY_01, $WITHER_00 or $TUMBLE_DRY_01. As for FN, the proliferation of frames and FEs is directly favored by the fact that the database focuses on the semantics and syntax of LUs instead of envisaging a limited repertoire of defining units.[19] On the contrary, in FunGramKB the division of labor between linguistic knowledge into the Grammaticon, Morphicon, and Lexicon, and of non-linguistic knowledge in the Ontology, Cognicon, and Onomasticon modules, supplies a well-balanced lexico-conceptual resource. This is so because in order to have lexical items in the lexicon of a particular language, there must previously exist a concept in the Ontology to which those lexical pieces can be associated.

3.3. Relations between frames and concepts

As advanced in 2.1, FN establishes frame-to-frame relations of various types, among which we will only focus on those that are useful for machine processing: Inheritance, Using, Perspective on, Subframe, Precedes, Inchoative of, and Causative of.

Inheritance covers those cases in which a child frame is a more specific elaboration of a superordinate or parent frame. For instance, the "Attaching" frame inherits information from its parent, i.e. the "Intentionally affect" frame (Fillmore et

[19] Among others, an illustrative example of FN being semantically and syntactically driven is the case of *give*. This LU evokes the "Giving" frame but not to the "Causation" frame, purely because of linguistic reasons, that is, because *give* is not always synonymous with *cause*. Consequently, they cannot appear in the same syntactic patterns (cf. Ruppenhofer et al. 2010:115-116). However, based on conceptual grounds, FunGramKB links *give* to the basic concept +GIVE_00, whose superordinate parents are +MOVE_00 >>> +TRANSFER_00; to put it simply, the event of giving "in the real world" implies causing something to move from one location to another.

al. 2003b: 311). Inheritance in FN, which corresponds to the IS-A relationship employed in FunGramKB and other ontologies, is the most comprehensive, formally defined relation, since it not only organizes frames, but also FEs and semantic types (cf. Ruppenhofer et al. 2010: 104-105). The Using relation, however, makes reference to those situations in which the child frame employs the parent frame only as a background, that is, "when a part of the scene evoked by the subordinate frame refers to the parent frame" (Ruppenhofer et al. 2010: 110). For instance, the "Offering" frame utilizes the "Giving" frame, but the latter does not presuppose the former. As a further elaboration of the Using relation, the Perspective on relation has just been introduced in FN. Perspective on applies to those frames that express a different point of view of a given event (e.g. the "Get a job" and "Hiring" frames are perspectives of the "Begin employment" frame; see Ruppenhofer et al. 2010: 107). In turn, the Subframe relation allows FN lexicographers to characterize complex frames. Thus, Subframe is utilized to designate the different sequential parts of a complex event, such as the "Inchoative attaching" frame within the complex "Attaching" one (Fillmore et al. 2003b: 313). Similarly, the Precedes relationship organizes the occurrence of the component frames of a single complex one in terms of their precedence; that is to say, the "Criminal process" frame is preceded by the subframe "Criminal investigation" which, in turn, is preceded by the "Committing a crime" subframe (Ruppenhofer et al. 2010: 109).[20] Finally, the Causative of and Inchoative of relations capture the non-inheritance connection between, for example, the "Position on a scale", "Change position on a scale", and "Cause change of scalar position"

[20] Rosca (2012: 257) rightly points out that the FN Precedes relation parallels the before relation proposed in Allen's (1983) interval sequences, which FunGramKB employs in the Cognicon. Likewise, Rosca also establishes a connection between the Subframe relationship and the number of subscripts that make up a complex script in the Cognicon. The interested reader is referred to Periñán and Arcas (2010b) and Garrido and Ruiz de Mendoza (2011) for further details on the nature of the FunGramKB Cognicon.

frames (see Ruppenhofer et al. 2010: 110). It is interesting to note that all these frame-to-frame relations are shown in the frame reports and can be viewed through the FrameGrapher tool available on the main FN website. By way of illustration, Figure 3 portrays a partial snapshot of frame-to-frame relations:

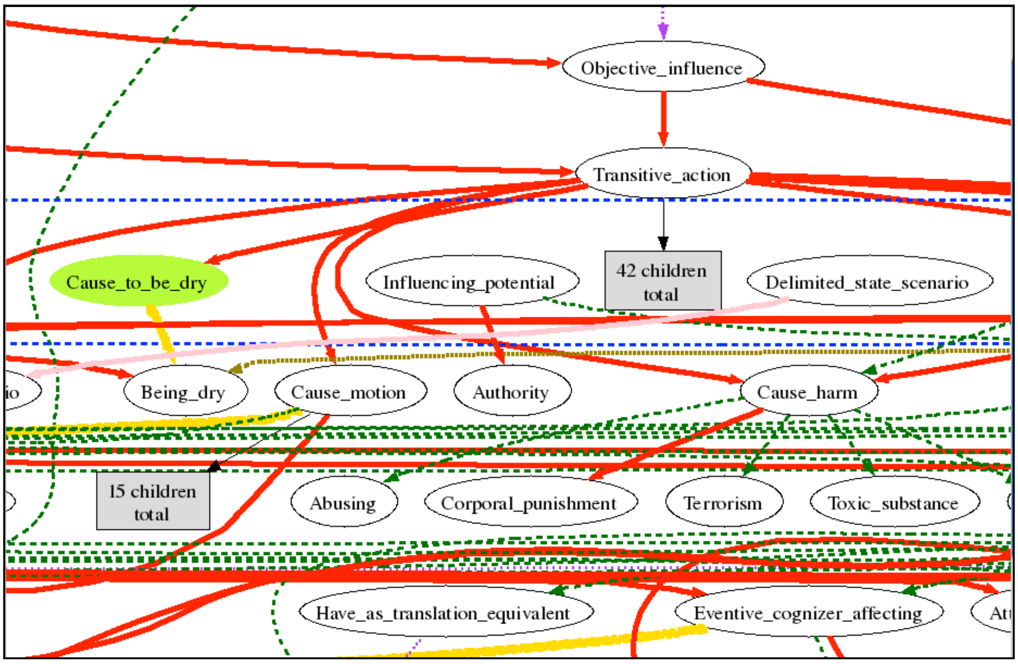


Figure 3: Frame-to-frame relations using FrameGrapher.

Although FN was not originally designed for NLP tasks, this issue has lately gained ground in this project. Hence, Ruppenhofer et al. (2010: 16-18, 121-127) argue that the above-mentioned relations are extremely valuable for paraphrasing, inferencing, and the propagation of information.[21] Since FN does not directly tackle the near-paraphrasability between sentences or utterances but only between words, frame relations like Causative of or Using explicitly connect LUs that may, however, belong to different frames. To illustrate, examples like *The paste hardened due to hydration of the cement* and *The hydration of the cement hardened the paste* can be

[21] It is commonly agreed in the literature that a powerful NLP system should be able to (1) paraphrase an input text, (2) translate it into another language, (3) answer questions about its contents, and (4) draw inferences from it (cf. Liddy 2001).

regarded as paraphrases of each other thanks to the Causative of relationship that links the two frames in which *harden* is stored (Ruppenhofer et al. 2010: 17). By the same token, frame relations, especially Inheritance, can explain some cases of (near) paraphrasability where inference is called upon. For example, *You needed my help so I got on the bus* and *You needed my help so I came* are considered near-paraphrases due to the fact that the LUs *bus* and *come* start off the inference process. As Ruppenhofer et al. (2010: 18) state, *bus* belongs to the "Vehicle" frame, which is associated to the "Motion" frame, this frame being employed by the "Arriving" frame, which is the one that lists the verb *come* among its LUs. Although we have tried to follow the conceptual path between *bus* and *come*, in the current "Vehicle" frame there is no link to the "Motion" frame but rather to the "Artifact" frame through Inheritance and, via the Is used by relation, to the "Convoy" and "Vehicle" subpart frames. We can venture here that inference may be triggered because the noun *bus* has in its lexical entry report the verb *come* as one of its governors, along with *catch*, *miss*, *leave*, etc., and thus *come* evokes the "Arriving" frame, among others. However, how the machine could arrive at such a connection is not made explicit in FN at this stage. Finally, Inheritance is also held responsible for the propagation of information both in terms of the semantic types and of the expected syntactical realization of the parent frame and its FEs down to the child frame and its FEs (see Ruppenhofer et al. 2010: 127-128).

Before moving on to FunGramKB, a number of issues need to be addressed here. As a glance at Figure 3 reveals, the baroque organization originated from the adoption of various types of frame-to-frame relations would most likely fail in a NLP environment as far as a reasoning engine is concerned. Furthermore, as Ruppenhofer et al. (2010: 121) contend, "more specific relationships between the words usually need to be ascertained to recognize or generate actual instances of paraphrases". Even though they claim that the wide range of relations proposed by WordNet could easily be mapped onto the FN LUs and vice versa, we are rather skeptical as to the efficiency of a reasoning engine that is required to handle not only the relations between frames but also between LUs and FEs, as "the relations of FEs and LUs to frames are intertwined with our Frame relations in various ways" (*ibid.* 126). We endorse here Ovchinnikova et al.'s (2010) proposal of improving these FN relations by employing an ontological analysis which, among other things, would entail connecting frames to categories in DOLCE (Masolo et al. 2003). As they rightly detail, such a methodological movement would solve issues such as FN's coverage poverty (i.e. the fact that semantically related frames that one would expect to be connected are, however, not captured that way in the FN taxonomy) which, as pointed out before in the *bus-come* example, hinder adequate inferences. Likewise, some of the conceptual discrepancies spotted in FN by Ovchinnikova et al. (2010) are claimed to stem from the fact that the FN Inheritance relation equally applies to frames, FEs, and semantic types. Consequently, serious reasoning problems can arise. Take, for instance, the example *Many [[parts]*^{PART}*]*$^{PART\ \ WHOLE}$ *[of tsunami-battered Aceh province]*^{WHOLE} *were not safe* (*ibid.*: 3158). Since in FN, the role WHOLE is connected to the "Substance" frame, which implies that its fillers will also be substances, the machine will wrongly process that the "Aceh province" is a substance. Therefore, Ovchinnikova et al. (2010: 3163) propose the axiomatization and restructuring of these FN relations, which will certainly improve the accuracy of natural language reasoning. In order to achieve this goal, an ontological analysis needs first to be applied to the current frames. In other words, frames are classified in terms of three of the categories employed by DOLCE, namely, perdurant, endurant, and quality, so that the relationships established between them can then be easily enriched, constrained, axiomatized, and cleaned up. As must be apparent to the reader, these DOLCE categories correspond to the three Subontologies already contemplated in FunGramKB (e.g. #EVENTS, #ENTITIES, and #QUALITIES), which highlights the benefits of an ontologically-driven approach to NLP systems vs. a lexically-oriented one.

As for FunGramKB, IS-A is the only relation permitted. Figure 4 shows the hierarchical organization of the conceptual route of +DRY_01. Clearly, FunGramKB would benefit from a visual tool like FrameGrapher that could make explicit the inheritance flow of semantic knowledge in the taxonomy.

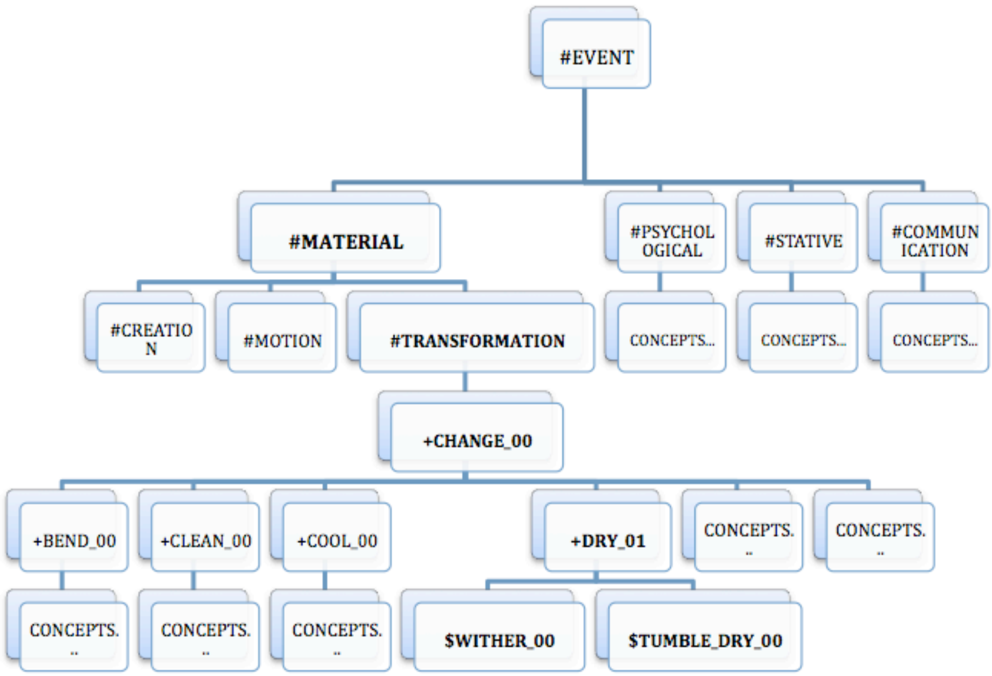


Figure 4: Conceptual path organization of +DRY_01 and its subordinates in FunGramKB.

But the exhaustive use of IS-A in FunGramKB is not free from problems. For example, despite the application of the OntoClean methodology to the Subontology of entities (see Periñán and Mairal 2011: 25-27), instances of taxonomic misclassifications and redundancy are present. A case in point is that of +FUEL_00 and its subordinate concepts +COAL_00, +PETROLEUM_00, +PETROL_00 and +WOOD_00. Since this basic concept includes three conceptual parents in its MP (i.e. +LIQUID_00 ^ +GAS_00 ^ +SOLID_00), +FUEL_00 and all of its subordinates appear three times in the taxonomy, once under +LIQUID_00, once connected to +GAS_00, and once linked to +SOLID_00. Therefore, although the IS-A relation is claimed to minimize redundancy and maximize information, redundancy is still present at least in the case of entities. In our view, these hierarchical problems cannot be successfully tackled until the FunGramKB Ontology is completely populated and thoroughly revised for possible inconsistencies.

Regarding FN, the use of inheritance is merely descriptive. As already discussed, the “Cause to be dry” frame inherits from the “Transitive action” frame and it is also the causative counterpart of “Being dry” (see Figure 3 above). For practical purposes, however, there seems to be no real inheritance of information, since each actual frame is re-defined in its own terms. That is, although the “Cause to be dry” frame defined as “An Agent causes a Dryee to become dry” is a subordinate of the “Transitive action” frame (i.e. “this frame characterizes, at a very abstract level, an Agent or Cause affecting a Patient”), such a specification is not fully incorporated into the definition of the causative frame itself, but rather, the connection between the two schemas is simply made explicit. A clearer example is that of the “Setting fire” frame, which also inherits from the “Transitive action” frame. Its definition is, however, the following: “this frame describes either the creation of Flame by a Kindler or the igniting of Flammables by a Kindler”.

But besides inheritance, FunGramKB also envisages a second reasoning mechanism, i.e. inference, which is based on the semantic knowledge shared between concepts that do not form part of the same IS-A relation. Figure 5, taken from Periñán and Arcas (2005: 241), portrays both mechanisms as if they were applied to any MP within the conceptual route between the nuclear concept and its metaconcept:

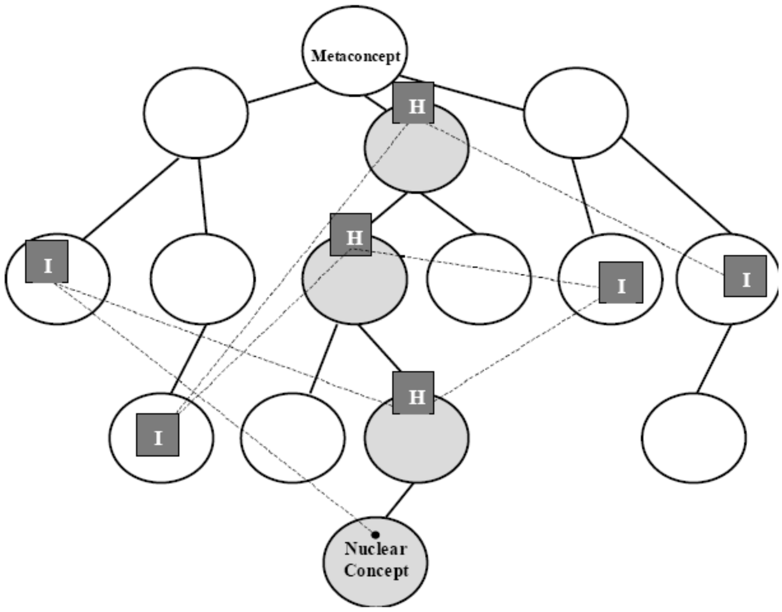


Figure 5: Inheritance (H) and Inference (I) mechanisms in FunGramKB.

We are fully aware that FN does not have NLP systems as one of its main goals. Therefore, we do not intend to criticize it for failing to provide what it was not designed to do. However, it is worth noting that FunGramKB already contemplates the existence of a reasoning engine that is able to work with the information and knowledge stored in the Ontology and in the Cognicon. Even though this reasoning engine is still being developed, there are two components that make it up: the Microconceptual-Knowledge Spreading or MicroKnowing and the Macroconceptual-Knowing Spreading or MacroKnowing. While the former runs on the Ontology, spreading the semantic knowledge contained in the MPs through the iterative application of inheritance and inference, the latter incorporates these MPs in the scripts stored in the Cognicon with a view to spreading the procedural knowledge recorded in FunGramKB.[22] Let us illustrate how the MicroKnowing process would expand the MP of + LIQUID_00 in (14) to include other concepts such as $BOTTLE_00, $GLASS_00, etc.

(14) MP of the entity +LIQUID_00: +(e1: +BE_00 (x1: +LIQUID_00)Theme (x2: +SUBSTANCE_00)Referent)

At the first spreading level, the inference mechanism is triggered, searching through the Ontology for any MP that contains the nuclear concept, that is, +LIQUID_00, as the selectional preference of the participants of a predication. For instance, the following predication from the MP of $BOTTLE_00 is inferred:

(15) *(e3: +CONTAIN_00 (x4: +LIQUID_00)Theme (x1)Location)

At the second spreading level, inheritance is applied so that those predications in the MP of the immediate superordinate concept of the nuclear concept are incorporated, namely, +SUBSTANCE_00:

(16) MP of the entity +SUBSTANCE_00: +(e1: +BE_00 (x1: +SUBSTANCE_00)Theme (x2: +ARTIFICIAL_OBJECT_00 ^ +NATURAL_OBJECT_00)Referent)

From this stage on, inference and inheritance are triggered iteratively, which results in the expansion of the conceptual meaning of +LIQUID_00 to incorporate the semantic knowledge from the predications of concepts such as +MELT_00, $GLASS_00, +WET_00, +TOUCH_01, +TASTE_01, +SPOON_00, etc.

This leads us to another crucial feature that distinguishes FN and FunGramKB, that is to say, how encyclopedic meaning is approached and dealt with by each resource. As far as frames describe types of situations, encyclopedic

---

[22] Periñán and Arcas (2005) provide the formal representation, the various spreading levels, and the resolution of incompatibilities of the MicroKnowing. The MacroKnowing does not as yet exist, but the interested reader can find a proper description in Periñán and Arcas (2007a).

knowledge is accounted for in FN. However, the level of detail achieved is different from that attained in FunGramKB. As illustrated above with +LIQUID_00, the FunGramKB conceptual definitions are more comprehensive than frames, since, thanks to inference and inheritance, they can contain an ever-growing web of encyclopedic knowledge. Consequently, for the conceptual unit +LIQUID_00, apart from learning that it is a substance (cf. its MP in (14)), we also infer that it is typically contained in a bottle or a glass, that as a result of a melting process a solid entity turns into liquid, that it has the quality of being wet, that it can be touched, tasted or that a human entity can use a spoon to ingest it. By contrast, the "Substance" frame evoked by the noun *liquid* is redundantly characterized as: "This frame concerns internally undifferentiated Substances". The Substance is the core FE, described as "The undifferentiated entity which is presented as having a permanent existence". As for the "Substance by phase" frame to which the adjective *liquid* is linked, the definition is: "in this frame an Undergoer is in a phase brought about after a change of phase. This can be its inherent phase or indicate that it is returned to its inherent state". However, this definition does not seem to provide the fine-grained analysis and the amount of encyclopedic information supplied by FunGramKB, since FN does not envision the reasoning mechanisms described thus far which are, nevertheless, crucial in an NLP environment.

3.4. Information management

FunGramKB's thematic roles and satellites can be equated to FEs, although the way in which each project manages this information is different.

As previously stated, the "Cause to be dry" frame evoked by *dry* contains three core FEs: an Agent, the Cause, which excludes the Agent, and a Dryee.[23] Being non-causative, the only core FE in the semantic frame "Becoming dry" is the Agent. Each of these FEs needs to be defined with respect to the semantic frame to which they belong. As such, whereas in the former the Agent is "the person performing the intentional act causing the Dryee to become dry", in the latter the Agent is characterized as "the thing that loses moisture and becomes dry". In the case of *bend*, which evokes a frame labeled "Change posture", we find that the Core FE is a Protagonist, that is, "the individual whose posture changes". The "Cause change of phase" frame (evoked by one sense of *melt*) is composed by three core FEs are: Agent (i.e. the entity that "causes the Undergoer to change phase"), Cause ("the non-Agent which brings about the change in the Undergoer") and Undergoer (the entity that "undergoes a change of phase brought about by the Agent"). Here, Cause also excludes the Agent.

From our point of view, the number of FEs in a given frame is rather unregulated, while it also seems that FEs are sometimes arbitrarily labeled. The nature of FN requires that each frame be provided with its own set of FE-types; some FEs usually occurring across frames (e.g. Manner, Place, Time, etc.), some largely varying from one frame to another, e.g. Importing area FE (frame "Importing"), Hot cold source FE ("Inchoative change of temperature" frame), Characteristic FE ("Path traveled" frame), Body part FE ("Grooming" frame), Food and Grower FEs ("Growing food" frame), Configuration FE ("Hair configuration" frame), etc. Except for the non-exhaustive list of extra-thematic FEs given in Appendix A of Ruppenhofer et al.'s (2010) book, there appears to be no clear policy that regulates

---

[23] In FunGramKB, instead of having a FE that excludes another FE, we would not specify a selectional preference in cases like this. Recall, by way of illustration, that in the MP of +DRY_00 in (11a), the (x1)Theme is left unspecified since different entities (e.g. a human being, the sun, etc.) can cause something to dry.

the choice of one FE over another.[24] A case in point is that of the FEs “Dryee”, “Patient” and “Item”:

(17) “Cause to be dry” frame:

Core FEs: Agent, Cause, Dryee (i.e. “the entity which has the water removed from it”, e.g. *Michael dehydrated* ***the apples***).

(18) “Cause to change strength” frame:

Core FEs: Agent, Patient (i.e. “entity that is being made stronger”, e.g. *Weight training strengthens* ***the muscles***).

Core unexpressed FE: Cause

(19) “Cause temperature change” frame:

Core FE: Agent, Cause, Hot cold source, Item (i.e. “the Item undergoes the temperature change”, e.g. *Ryan reheated* ***the pasta*** *in the microwave*).

While clearly Agent and Cause are applied consistently, we may ask ourselves on what grounds the more specific participants Dryee or Item are selected over that of the more inclusive notion of Patient in (18). Seemingly, the question arises as to where lexicographers draw the line when it comes to addressing the number of FEs that may be involved in each semantic frame. For instance, while we find a Result FE in the “Absorb heat” frame (e.g. *The lobster boiled to death*), the “Change phase” frame (e.g. *The lake froze solid*) or the “Emotion active” frame (e.g. *She was worrying herself sick about Biggles*), we have seen that this FE is not present in the “Ingestion” semantic frame to account for realizations of the type *They drank themselves blind* (BNC EUU), *The snake that ate itself to death* (BNC FBG), etc.

[24] Periñán (2013: 89) argues that FN lexicographers probably “opted for this excessive granularity in semantic roles in order to compensate for the deficiencies” resulting from the adoption of a shallow approach to lexical meaning.

Likewise, the "Ingestion" frame does not include the Iteration FE even though the BNC retrieves cases like *I fed her twice a day* (BNC C8U) or *She eats meat once a week* (BNC EG0).

In FunGramKB, thematic frames and satellites are managed in a significantly different fashion. First, the cognitive dimensions under which all concepts are organized amounts to a total of 15 metaconcepts in the Subontology of events (e.g. #MOTION, #COGNITION, #EXISTENCE, #EMOTION, etc., see Periñán and Mairal (2010) for the complete list). In turn, the catalogue of thematic roles is restricted to the following participants: Agent, Theme, Referent, Attribute, Location, Origin and Goal. Obviously, the number of thematic roles varies from one cognitive dimension to the other. #MOTION, for example, takes Agent, Theme, Location, Origin and Goal; #EXISTENCE only contains the Theme role, whereas Theme, Referent and Goal are the participants in the metaconcept #COMMUNICATION. Moreover, the characterization of this limited set of thematic roles is different depending on each dimension. For example, #CREATION, #TRANSFORMATION and #PERCEPTION include two thematic roles, i.e. a Theme (x1) and a Referent (x2). Their definitions, however, are the following:

(20) #CREATION

a. Theme: entity that creates another entity.

b. Referent: entity that is created by another entity.

(21) #TRANSFORMATION

a. Theme: entity that transforms another entity.

b. Referent: entity that is transformed by another entity.

(22) #PERCEPTION

a. Theme: entity that perceives another entity through any of the senses.

b. Referent: entity that is perceived through any of the senses.

Take now the three LUs discussed at the beginning of this section (i.e. *dry*, *bend* and *melt*) and their corresponding basic concepts in FunGramKB (i.e. +DRY_01, +BEND_00 and +MELT_00). These conceptual units, together with others like +TRANSLATE_00, +SPLIT_00, +REPAIR_00, +CLEAN_01, etc., are located within the metaconcept #TRANSFORMATION. This entails that, despite the obvious meaning differences among these concepts, they and their subordinates must take an (x1)Theme and an (x2)Referent as participants in their TFs. However, the definition of these thematic roles (cf. example (21) above) is not altered so as to comply with the specifics of any of these concepts. The same rationale applies to #PERCEPTION in (22), since every single TF of any of the concepts situated under this metaconcept will only display two participant roles whose characterization is exactly the same throughout this entire portion of the Ontology.

Both FEs and the participants in the TF need to be defined against a given frame or metaconcept. Nonetheless, in FunGramKB the number and type of categories (i.e. 15 cognitive dimensions), and by extension the number of thematic roles and their definitions, is much more limited than in the case of FN. Additionally, it should be noted that in some cases, a FE can also take the form of a selectional preference in FunGramKB. For example, the TF of the terminal concept $FRY_00, a subordinate of +COOK_00, specifies two selectional preferences via two basic concepts: a human (i.e. (x1: +HUMAN_00)Theme) fries food (i.e. (x2: FOOD_00)Referent). In the "Apply heat" frame, the selectional preference (x2: +FOOD_00)Referent corresponds to the core FE Food of the verb *fry*. As was the case with metaconcepts and thematic roles, basic concepts also make up a limited inventory (cf. section 3.1).

We have already explained that all subordinate concepts in the Ontology must specify a distinctive feature or *differentia*, which is not present in the MP of the superordinate. In the syntax of COREL, *differentiae* are expressed by employing satellites. FunGramKB thus posits a total of 16 satellites, namely, beneficiary, company, condition, duration, frequency, instrument, manner, means, position, purpose, quantity, reason, result, scene, speed and time. Some of these satellites are similar to the most widely occurring extra-thematic and peripheral FEs. However, as in the case of thematic roles, each of these satellites has one and only one definition independently of the peculiarities of any given concept. For comparative purposes, let us deal with one example. In FunGramKB, the manner satellite is described as the "entity or quality that describes the way in which the event occurs". By contrast, the Manner FE in FN may be defined as:

(23) "Abusing" frame:

Manner FE: Manner is used for any description of the abuse event which is not covered by more specific FEs, including secondary effects (*quietly*, *loudly*), and general descriptions comparing events (the same way). In addition, it may indicate salient characteristics of an Abuser that also affect the action (*presumptuously*, *coldly*, *deliberately*, *eagerly*, *carefully*), e.g. *It is fair to say that it was* ***malicious*** *abuse*.

(24) "Cause to make progress" frame:

Manner FE: A description of the progress, generally a description comparing events (the same way); in addition, it may describe characteristics and states of the Agent that also affect the development (no example is given).

(25) "Make noise" frame:

Manner FE: Manner expressions may be of lexicographic interest if they describe properties of a sound as such: *loudly*, *shrilly*, etc., e.g. *Jocelyn sobbed* ***loudly***.

From this explanation it follows that FunGramKB opts for an economic, parsimonious organization of both thematic roles and satellites, which allows for the efficient management of information in an NLP environment.

3.5. Multilingualism

As rightly pointed out by Boas (2009b: 59), creating multilingual lexical databases (or knowledge bases, in the case of FunGramKB) that may be employed in a variety of (NLP) applications is a particularly challenging endeavor. Despite its indisputable difficulty, Boas (2009b: 59-60; see also Boas 2009a) contends that the English FN database facilitates the creation of multilingual lexical databases capable of overcoming linguistic problems concerning issues like diverging polysemy structures or differences in lexicalization patterns, among others. Thus, it is argued that semantic frames have been successfully utilized by a number of FN-type projects for both western languages (e.g. Spanish and German) and non-western languages like Japanese.

In building parallel non-English FNs various phases need to be followed, although only a simplified summary is included here (see Boas 2009b: 72-87 for more details). First, after downloading the English FN MySQL database, all language-specific information needs to be removed. Second, once only information not specific to English remains (i.e. conceptual information in the form of a frame relation table, frame elements table, etc.), the lexicographer can repopulate the database with lexical descriptions from another language. This step involves selecting a semantic frame from the stripped-down original database to which, with the help of dictionaries and parallel corpora, a set of LUs evoking that frame are added. Having completed these steps, corpus-attested sentences illustrating the use of each of the frame-evoking LUs are inserted. But the real labor-intensive process begins when parallel lexicon fragments are linked in order to map lexical information from one language to another. This entails that within a given semantic frame the lexicographer should find LUs in, for example, Spanish, which have corresponding semantic and syntactic combinatorial possibilities in English. Having identified comparable structures, the linking is performed via semantic frames, which work as *interlingual representations* (i.e. representations that combine aspects of true interlinguas and transfer-based systems). So far, only Spanish, German and Japanese FNs are associating their entries to entries from the English FN. Hence, the next step would require "linking lexical entries across languages in order to test the applicability of semantic frames as a cross-linguistic metalanguage" (Boas 2009b: 92).

The following conclusions may be gathered from the account given above. Even though individual FNs for various languages exist independently of one another, the approach adopted in the creation of frame-based multilingual databases is in line with the conceptualist account advocated by FunGramKB: both depart from linguistically motivated concepts (or semantic frames, in the case of FN), on the grounds of which LUs from various languages are interconnected. Two differences may be highlighted. First, unlike FN, FunGramKB was originally designed as an inherently multilingual knowledge base. Since the Ontology is the central component of the whole architecture and, therefore, semantic descriptions are provided at the conceptual level, the management of multilingual knowledge is facilitated. This avoids the removal of language-specific information and subsequent repopulation described by Boas (2009b). Instead, the process of defining language-independent concepts and that of inserting the set of related LUs lexicalizing them (in various languages) is carried out simultaneously. Second, because FunGramKB is a resource for NLP, the definition of concepts, in contrast to semantic frames, is formalized via the universal metalanguage, i.e. COREL, thanks to which the computer is able to understand and process knowledge. In spite of these advantages, it should be noted that so far FunGramKB has less monolingual and multilingual information than FN.

In closing, it is worth emphasizing that both projects appear to point in the same direction:

> At this early point, semantic frames do not serve as a true interlingua in which a concept is realized independently of a source language (…) *Once more languages are described, we may arrive at true universal semantic frames (e.g. communication, motion, etc.), which may then serve as a true interlingua*" (Boas 2009b: 87, our emphasis).

As previously mentioned, not only does FunGramKB already make use of the COREL metalanguage, but also its cognitive dimensions or metaconcepts, e.g. #COMMUNICATION, #MOTION, #POSSESSION, #PERCEPTION, #COGNITION, etc., appear to correspond to the future universal semantic frames pursued in FN.

4. Conclusion

This study has made an elaborate comparison between FN and FunGramKB by focusing on how each of these projects handles the representation of semantic knowledge. The implications for the creation of NLP tasks have also been peripherally touched upon, allowing us to make explicit the differences between a shallow approach and a deep approach to knowledge representation.

As detailed throughout the chapter, FunGramKB and FN have originally been created to fulfil different goals, which certainly speaks of their differences. On the one hand, FN is a lexical resource based on Frame Semantics whose aim is to document the semantic and syntactic combinatory possibilities of each word sense. FunGramKB, on the other hand, is a knowledge-base designed for its use in NLP. This does not mean, however, that these projects cannot be compared, since as we have seen, FN and FunGramKB are different in many respects but are also alike in others. To illustrate, we could list their common usage of organizational conceptual units larger than words, their equal treatment of polysemy, their account of semantic types, selectional preferences or encyclopedic knowledge, and their interest in developing multilingual resources.

In sum, FN and FunGramKB offer complementary types of information in many aspects, especially in the way they approach linguistic knowledge. However, at present, since FunGramKB assumes a division of labor between linguistic knowledge and conceptual knowledge in which the Ontology is the component that connects them, this knowledge base is more suitable for language processing. Additionally, given that text understanding implies the incorporation of a situation model so that the gestalt meaning of the text can be recovered by the machine from the words that make it up (Zwaan and Radvansky 1998, *apud* Periñán and Arcas 2007a), lexical knowledge *per se* is not sufficient. Thus, the fact that FunGramKB links each LU to a concept in the Ontology and the knowledge therein, a step that FN has not taken yet, assures the creation of more accurate NLP systems for language understanding.

5. References

Allen, James F. 1983. Maintaining knowledge about temporal intervals. *Communications of the ACM* 26(11): 832-843.

Atkins, Sue, Fillmore, Charles J. & Johnson, Christopher R. 2003. Lexicographic relevance: Selection information from corpus evidence. *International Journal of Lexicography* 16(3): 251-280.

Baker, Collin F., Fillmore, Charles J. & Lowe, John B. 1998. The Berkeley FrameNet project. In *COLING-ACL '98 Proceedings of the Conference*, 86-90. Montreal: Canada.

Boas, Hans C. 2005. From theory to practice: Frame Semantics and the design of FrameNet. In *Semantik im Lexikon*, Stefan Langer & Daniel Schnorbusch (eds), 129-160. Tübingen: Narr.

Boas, Hans C. 2009a. Recent trends in multilingual computational lexicography. In *Multilingual FrameNets in Computational Lexicography: Methods and Applications*, Hans C. Boas (ed), 1-26. Amsterdam: John Benjamins.

Boas, Hans C. 2009b. Semantic frames as interlingual representations. In *Multilingual FrameNets in Computational Lexicography: Methods and Applications*, Hans C. Boas (ed), 59-100. Amsterdam: John Benjamins.

Boas, Hans C. 2010. Linguistically relevant meaning elements of English communication verbs. *Belgian Journal of Linguistics* 24: 54-82.

Boas, Hans C. 2011. A frame semantic approach to syntactic alternations with build-verbs. In *Morphosyntactic Alternations in English*, Pilar Guerrero Medina (ed), 207-234. London: Equinox.

Cambridge Advanced Learners Dictionary <http://dictionary.cambridge.org/dictionary/british/> (September 2012).

Davies, Mark. *The British National Corpus* (BNC) <http://corpus.byu.edu/bnc/> (September 2012).

Davies, Mark. *The Corpus of Contemporary American English* (COCA) <http://www.americancorpus.org/> (September 2012).

Dixon, Robert M. W. 1991. *A New Approach to English Grammar on Semantic Principles*. Oxford: Clarendon Press.

English Cobuild Dictionary <http://dictionary.reverso.net/english-cobuild/> (September 2012).

Fillmore, Charles J. 1976. Frame Semantics and the nature of language. In *Origins and Evolution of Language and Speech*, Stevan R. Harnad, Horst D. Steklis & Jane Lancaster (eds), 20-32. New York: New York Academy of Sciences.

Fillmore, Charles J. 1982. Frame Semantics. In *Linguistics in the Morning Calm*, The Linguistic Society of Korea (ed), 11-137. Seoul: Hanshin.

Fillmore, Charles J. & Atkins, Beryl T. 1992. Towards a frame-based organization of the lexicon: The semantics of RISK and its neighbors. In *Frames, Fields, and Contrasts: New Essays in Semantics and Lexical Organization*, Adrienne Lehrer & Eva Kittay (eds), 75-102. Hillsdale: Lawrence Erlbaum.

Fillmore, Charles J., Johnson, Christopher R. & Petruck, Miriam R. L. 2003a. Background to FrameNet. *International Journal of Lexicography* 16(3): 235-250.

Fillmore, Charles J., Lee-Goldman, Russell R. & Rhodes, Russell. 2012. The FrameNet constructicon. In *Sign-Based Construction Grammar*, Hans C. Boas & Ivan Sag (eds), 283-299. Stanford: CSLI Publications.

Fillmore, Charles J., Petruck, Miriam R. L., Ruppenhofer, Josef & Wright, Abby. 2003b. FrameNet in action. The case of Attaching. *International Journal of Lexicography* 16(3): 297-332.

Garrido, Nazaret & Ruiz de Mendoza, Francisco J. 2011. La modelación de conocimiento procedimental en FunGramKB. *Anglogermánica Online* 8: 106-120.

Guerra, Fátima & Sacramento, Elena. 2011. El modulo léxico de FunGramKB. *Anglogermánica Online* 8: 52-65.

Halliday, Michael A. K. 1985. *An Introduction to Functional Grammar*. London: Edward Arnold.

Jiménez, Rocío & Luzondo, Alba. 2011. Building ontological meaning in a lexico-conceptual knowledge-base. *Onomázein* 23(1): 11-40.

Jiménez, Rocío & Pérez, María Beatriz. 2011. An account of selection restrictions in Role and Reference Grammar. *Revista Canaria de Estudios Ingleses* 62: 99-122.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago/London: University of Chicago Press.

Liddy, Elizabeth D. 2001. Natural language processing. In *Encyclopedia of Library and Information Science* (2nd edition), Miriam A. Drake (ed). New York: Marcel Decker, Inc.

Longman Dictionary of Contemporary English <http://www.ldoceonline.com/> (September 2012).

Mairal, Ricardo & Periñán, Carlos. 2009. The anatomy of the lexicon component within the framework of a conceptual knowledge base. *Revista Española de Lingüística Aplicada* 22: 217-244.

Mairal, Ricardo & Periñán, Carlos. 2010a. Role and Reference Grammar and ontological engineering. In *Los Caminos de la Lengua. Estudios en Homenaje a Enrique Alcaraz Varó*, José Luis Cifuentes, Adelina Gómez, Antonio Lillo, José Mateo & Francisco Yus (eds), 649-665. Alicante: Universidad de Alicante.

Mairal, Ricardo & Periñán, Carlos. 2010b. Teoría lingüística y representación del conocimiento: Una discusión preliminar. In *Tendencias en Lingüística General y Aplicada*, Dolores García Padrón & María del Carmen Fumero Pérez (eds), 155-168. Berlin: Peter Lang.

Mairal, Ricardo, Periñán, Carlos & Pérez, María Beatriz. 2012. La representación léxica. Hacia un enfoque ontológico. In *El Funcionalismo en la Teoría Lingüística. La Gramática del Papel y la Referencia. Introducción, Avances y Aplicaciones*, Ricardo Mairal, Lilián Guerrero & Carlos González (eds), 85-102. Akal: Madrid.

Mairal, Ricardo & Ruiz de Mendoza, Francisco J. 2009. Levels of description and explanation in meaning construction. In *Deconstructing Constructions*, Christopher Butler & Javier Martín Arista (eds), 153-198. Amsterdam: John Benjamins.

Masolo, Claudio, Borgo, Stefano, Gangemi, Aldo, Guarino, Nicola & Oltramari, Alessandro. 2003. WonderWeb Deliverable D18: Ontology library. Technical report. Laboratory for Applied Ontology, ISTC-CNR.

Newman, John & Rice, Sally. 2005. Transitivity schemas of English EAT and DRINK in the BNC. In *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*, Stefan Th. Gries & Anatol Stefanowitsch (eds), 225-260. Berlin/New York: Mouton de Gruyter.

Ovchinnikova, Ekaterina, Vieu, Laure, Oltramari, Alessandro, Borgo, Stefano & Alexandrov, Theodore. 2010. Data-driven and ontological analysis of FrameNet for Natural Language Reasoning. In *Proceedings of the 7th Conference on International Language Resources and Evaluation (LREC'10)*, 3157-3164. Valletta: Malta.

Oxford Dictionaries <http://oxforddictionaries.com/> (September 2012).

Periñán, Carlos. 2013. A knowledge-engineering approach to the cognitive categorization of lexical meaning. *Vigo International Journal of Applied Linguistics* 10: 85-104.

Periñán, Carlos. Forthcoming. Towards a model of constructional meaning for natural language understanding. In *Liking Constructions into Functional Linguistics: The Role of Constructions in RRG Grammars* [Studies in Language], Brian Nolan & Elke Diedrichsen (eds). Amsterdam: John Benjamins.

Periñán, Carlos & Arcas, Francisco. 2005. Microconceptual-Knowledge Spreading in FunGramKB. In *Proceedings of the 9th IASTED International Conference on*

Artificial Intelligence and Soft Computing, 239-244. Anaheim-Calgary-Zurich: ACTA Press.

Periñán, Carlos & Arcas, Francisco. 2007a. Cognitive modules of an NLP knowledge base for language understanding. *Procesamiento del Lenguaje Natural* 39: 197-204.

Periñán, Carlos & Arcas, Francisco. 2007b. Deep semantics in an NLP knowledge base. *Proceedings of the 12th Conference of the Spanish Association for Artificial Intelligence*, 279-288. Universidad de Salamanca: Salamanca.

Periñán, Carlos & Arcas, Francisco. 2010a. Ontological commitments in FunGramKB. *Procesamiento del Lenguaje Natural* 44: 27-34.

Periñán, Carlos & Arcas, Francisco. 2010b. The architecture of FunGramKB. In *Proceedings of the 7th International Conference on Language Resources and Evaluation*, 2667-2674. Malta: European Language Resources Association.

Periñán, Carlos & Arcas, Francisco. 2011. Introducción a FunGramKB. *Anglogermánica Online* 8: 1-15.

Periñán, Carlos & Mairal, Ricardo. 2009. Bringing Role and Reference Grammar to natural language understanding. *Procesamiento del Lenguaje Natural* 43: 265-273.

Periñán, Carlos & Mairal, Ricardo. 2010. La gramática de COREL: Un lenguaje de representación conceptual. *Onomázein* 21: 11-45.

Periñán, Carlos & Mairal, Ricardo. 2011. The COHERENT methodology in FunGramKB. *Onomázein* 24(2): 13-33.

Petruck, Miriam R. L. 1996. Frame Semantics. In *Handbook of Pragmatics*, Jef Verschueren, Jan-Ola Östman, Jan Blommaert & Chris Bulcaen (eds), 1-13. Amsterdam: John Benjamins.

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

Rosca, Andreea. 2012. Bases for the Development of Ontological Semantics within the Conceptual Domains of Change and Possession. Implementations and Implications for the Lexico-Syntactic-Cognition Interface and the Development of Intelligent Agents. PhD dissertation, Universidad de La Rioja.

Ruiz de Mendoza, Francisco J. Forthcoming. Meaning construction, meaning interpretation and formal expression in the Lexical Constructional Model. In *Linking Constructions into Functional Linguistics: The Role of Constructions in RRG Grammars* [Studies in Language], Brian Nolan & Elke Diedrichsen (eds). Amsterdam: John Benjamins.

Ruiz de Mendoza, Francisco J. & Mairal, Ricardo. 2008. Levels of description and constraining factors in meaning construction: An introduction to the Lexical Constructional Model. *Folia Linguistica* 42(2): 355-400.

Ruiz de Mendoza, Francisco J. & Mairal, Ricardo. 2011. Constraints on syntactic alternations: Lexical-constructional subsumption in the Lexical Constructional Model. In *Morphosyntacic Alternations in English*, Pilar Guerrero (ed), 62-82. London: Equinox.

Ruppenhofer, Josef, Ellsworth, Michael, Petruck, Miriam R. L., Johnson, Christopher R & Scheffczyk, Jan. 2010. *FrameNet II: Extended Theory and Practice*. <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>.

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

Van Valin, Robert D. Jr. & LaPolla, Randy J. 1997. *Syntax, Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Velardi, Paola, Fasolo, Michela & Pazienza, María T. 1991. How to encode semantic knowledge: A method for meaning representation and computer aided acquisition. *Computational Linguistics* 17(2): 153-170.

Zwaan, Rof A. & Radvansky, Gabriel A. 1998. Situation models in language comprehension and memory. *Psychological Bulletin,* 123(2): 162-185.