

THE PROCESS OF CONSTRUCTING ONTOLOGICAL MEANING  
BASED ON CRIMINAL LAW VERBS

Ángel Felices Lago

*Universidad de Granada*

[afelices@ugr.es](mailto:afelices@ugr.es)

Abstract

This study intends to account for the process involved in the construction of the conceptual meaning of verbs (#EVENTS) directly related to legal aspects of terrorism and organized crime based on the evidence provided by the Globalcrimeterm Corpus and the consistent application of specific criteria for term extraction. The selected 49 concepts have eventually been integrated in the Core Ontology of FunGramKB (Functional Grammar Knowledge Base), a knowledge base which is founded on the principles of deep semantics and is also aimed at the computational development of the Lexical Constructional Model ([www.fungramkb.com](http://www.fungramkb.com)). To achieve this purpose, key phases of the COHERENT methodology (Periñán Pascual & Mairal Usón 2011) are followed, particularly those which involve the modelling, subsumption and hierarchisation of the aforementioned verbal concepts. The final outcome of this

Ángel Felices Lago. 2016.

The process of constructing ontological meaning based on criminal law verbs.

*Círculo de Lingüística Aplicada a la Comunicación* 65, 109-148.

<http://www.ucm.es/info/circulo/no65/felices.pdf>

<http://revistas.ucm.es/index.php/CLAC>

[http://dx.doi.org/10.5209/rev\\_CLAC.2016.v65.51983](http://dx.doi.org/10.5209/rev_CLAC.2016.v65.51983)

research shows that most of the apparently specialised conceptual units should eventually be included in the Core Ontology instead of the specific Globalcrimeterm Subontology, due to the fact that the semantic content of their corresponding lexical units can be found in widely used learner's dictionaries and, consequently, this conceptual information is not only shared by the experts in the field but also by the layperson and the average speaker of the language.

Key words: FunGramKB, ontology building, conceptual modelling, term extraction

## Contents

1. Introduction	111
2. Ontology building and legal ontologies	113
3. The architecture of FunGramKB	116
4. Corpus design and terminological extraction	120
4.1 Corpus compilation	121
4.2 Term extraction process	122
5. Conceptual modelling	126
6. Hierarchisation process	130
7. Conclusions	135
References	136
Appendix	143

## 1. Introduction<sup>1 2</sup>

It is a well-known fact that the main problem in the construction of natural language understanding systems is usually found in the lack of a robust semantic knowledge base and a powerful inference component (Vossen 2003). Moreover, a key aspect in knowledge engineering is the design and construction of an ontology model under a series of well-founded guidelines, particularly when you want to reuse it in different natural language processing (henceforth NLP) applications, e.g. document retrieval, information extraction, text categorization, etc. Consequently, ontology structuring must be supported by some theory about the elements in the domain, their inherent properties and the way in which these elements are related to each other. To reach that purpose, the comprehensive theory of constructional meaning known as the Lexical Constructional Model (Mairal Usón & Ruiz de Mendoza 2008, 2009; Ruiz de Mendoza & Mairal Usón 2008, among others) has incorporated as part of its architecture the knowledge base FunGramKB (see Functional Grammar Knowledge Base at [www.fungramkb.com](http://www.fungramkb.com)): an online lexical conceptual knowledge base that integrates semantic and syntactic information for the creation of NLP applications (Periñán Pascual & Arcas Túnez 2004, 2005, 2007). The main advantage of this knowledge base is its capacity to combine linguistic knowledge and human cognitive abilities within the same integrated system. The concept-oriented interlingua which is used (COREL) serves to describe the properties of the different modules that integrate FunGramKB in the cognitive level (Periñán Pascual & Arcas Túnez 2010). As a consequence, this knowledge base moves away from the traditional solutions based

---

<sup>1</sup> This article is based on research carried out within the framework of the projects FFI2014-53788-C3-1-P and FFI2010-15983, which are funded by the Spanish Ministry of Economy and Competitiveness.

<sup>2</sup> I wish to express my sincere gratitude to my colleagues Alba Luzondo Oyón and Pedro Ureña Gómez-Moreno that collaborated so generously in the conceptual modelling of verbal concepts included in the Globalcrimeterm corpus and collected in the appendix.

on surface semantics to offer a fully-fledged alternative in which linguistic information is grounded on conceptual structures representing human knowledge.<sup>3</sup>

However, the focus in this article is on the FunGramKB Ontology, which can be considered as the pivotal module for the whole architecture of the knowledge base. The Ontology, along with the Cognicon and the Onomasticon (see section 3), is presented as a hierarchical catalogue of the concepts that a person has in mind when talking about everyday situations and is also the repository where semantic knowledge is stored in the form of meaning postulates. (Periñán Pascual & Arcas Túnez 2007: 198). The Ontology consists of a general-purpose module (Core Ontology) and several domain-specific terminological modules (Satellite Ontologies or Subontologies). With reference to the latter, in the last few years a research project has been carried out in order to create a terminological subontology based on the international cooperation against terrorism and organized crime (*Globalcrimeterm*) under the postulates of FunGramKB.<sup>4</sup> This domain-specific ontology combines a narrow and, at the same time, fuzzy terminological scope with diverse interdisciplinary sub-fields. However, it is helpful to note that in the following sections I will explain how this Subontology shares the same integrated structure as the Core Ontology, and both contain a well-structured body of concepts related to each other in an “IS-A” conceptual hierarchy. Furthermore, both types of ontologies distinguish between metaconcepts, basic concepts and terminal concepts; both have COREL as a common metalanguage for meaning representation, and both share and split the metaconcepts into three subontologies which arrange lexical units of a different part of speech; i.e. #ENTITIES for nouns, #EVENTS for verbs, and #QUALITIES for adjectives and some adverbs.

Within this context, the purpose of this paper is to account for the process involved in the construction of the conceptual meaning of verbs (#EVENTS) directly related to the aforementioned domain-specific Ontology and the application of the COHERENT

---

<sup>3</sup> A conceptual approach to meaning construction is proposed, being based on the methodological principles which have been essential for both formal and functional linguistic models, e.g. Jackendoff (1990), Pustejovsky (1995), Levin & Rappaport (2005), Van Valin (2005) or Reinhart (2006).

<sup>4</sup> The project referred to above was funded by the Spanish Ministry of Economy and Competitiveness, code no. FFI2010-15983 and the results have been included in the FunGramKB editor.

methodology (Periñán Pascual & Mairal Usón 2011): a stepwise method for forming basic or terminal concepts and their subsumption under the Core Ontology or other satellite ontologies. In doing so, this study furnishes substantial evidence on the modelling, subsumption and hierarchization of a set of verbal concepts borrowed from the domain of legal ontologies (cf. Breuker, Valente & Winkels 2005; Valente 2005; Breuker et al. 2008, Sartor et al. 2011), particularly those included in the *Globalcrimeterm* Corpus (Ureña Gómez-Moreno, Alameda Hernández & Felices Lago (2011); Felices Lago and Ureña Gómez-Moreno (2012, 2014). Consequently, this article is organized as follows: sections 2 and 3 deal with an introduction to legal ontologies, followed by an overview of FunGramKb and the building of ontological meaning under the principles of deep semantics; section 4 explains the methodology used for (a) compiling the corpus, (b) designing the term extractor, and (c) analysing the verbal units finally selected; sections 5 and 6 describe and discuss the results of the conceptualisation and hierarchisation phases in the application of the COHERENT methodology and section 7 offers some concluding remarks; finally, an appendix of the selected 49 concepts (under #EVENT) is provided, with a full description of their meaning postulates.

## 2. Ontology building and legal ontologies

The origin of the term *ontology* comes from philosophy and bears no relation with the concept of ontology in NLP (e.g. Musen 1992, Gruber 1993), even if both share the human endeavour to comprehend the structure of knowledge and reality. From a more linguistic perspective, Sowa (2000: 492) defines *ontology* as “a catalogue of the type of things that are assumed to exist in a domain of interest *D*, from the perspective of a person who uses a language *L* for the purpose of talking about *D*”. However, it is our concern the interpretation of the concept *ontology* in the framework of knowledge engineering (Gruber 1993), which consists of a hierarchy of concepts, attributes and their associations in order to allow the establishment of a semantic network of relations. In this vein, a domain-specific ontology of concepts within a certain field, along with their relations and properties, is a new medium for the storage and propagation of specialised knowledge (Hsieh et al. 2010).

As regards the contemporary development of specialised ontologies in connection with their meaning in Artificial Intelligence (AI) or NLP it is worth mentioning, among many others, projects based on specialised knowledge representation through corpus such as *Oncoterm* (Faber 2002), *Ecolexicon* (Faber 2014), *Genoma-KB* (Cabr  et al. 2004), *PoCeHRMOM* ( Kerremans, Temmerman et al. 2007), *Prolex* (Maurel 2008), or the pioneering *Cogniterm* prototype (developed by Sku e between 1991 and 1997) and its management system of data knowledge bases called CODE (*Conceptually Oriented Design Environment*). However, the description of ontologies as conceptual schemas within the legal domain arises in 1995 (Valente & Breuker 1994, Valente 1995), as a result of the growing necessity to formalize the information exchange and linkage among all the components which make up a legal system. In 1997 this field will emerge as a new area of research in the *First International Workshop on Legal Ontologies (LEGONT`97)* held within the biennial *International Conference on Artificial Intelligence and Law (ICAIL-97)*. The objective has always been to provide the adequate instruments for accessing and managing a growing amount of legal information which is rapidly produced in electronic format every day (Breuker, Casanovas, Klein & Francesconi 2008).

Concerning the origins and development of ontologies used in the legal field, I should cite Liebwald (2007: 140), who concluded that “the formalization of implicit [legal] knowledge proved to be especially difficult”. “The cross-linking of different domains and the connection between legal concepts and world concepts is still problematic. Contrary to e.g. a biological taxonomy, a legal ontology is not language and country independent”<sup>5</sup>. In consequence, the most feasible options are application-oriented or specific domain ontologies. He also adds that “... ontology developers should always consider the specific needs of the intended application area(s) and user group(s)”. Due to the complexity of this new and heterogeneous field, Valente (2005: 72) proposed a classification of the set of types and roles of ontologies in order to account for the legal

---

<sup>5</sup> From this point of view, law is a dynamic, normative field and its conceptualization would necessarily include those aspects, together with the representation of world knowledge or common-sense knowledge (see, for example, Lame (2002) and Breuker and Hoekstra (2004)).

domain ontologies developed since the 1990s. He collected a catalogue of 24 ontologies<sup>6</sup> and concluded that “different authors mean different things by the term ‘ontology’” and that “ontologies are used in very different ways”<sup>7</sup>. In addition, Periñán Pascual & Arcas Túnez (2007) assert that the large majority of these “misnamed” ontologies are, in fact, lexical taxonomies which do not give formal representation of meaning to each of their terms, but which are rather infra-defined as regards their subsumptive relation with other terms (and sometimes with other semantic relations such as synonymy, meronymy, etc.). Some of the so called ontologies (Casanovas, Sartor, Biasiotti & Fernández-Barrera 2011: 5-7):

- (1) organize and structure information, as in the case of projects such as *Jur-Wordnet* (Gagemi, Sagre & Tiscornia 2005) or the Italian ontology of crimes (Asaro et al. 2003; Lenci 2008);
- (2) have a reasoning and a problem solving engine, such as the ontology *CLIME* for maritime law (Boer, Hoekstra & Winkels 2001) or *Argument Developer*, which works with different types of legal data bases (Zeleznikow & Stranieri, 2001);
- (3) have semantic indexing and search, such as the ontologies of French codes (Lame 2002), ontologies which represent cases of financial fraud (Leary, Vandenberghe & Zeleznikow 2004) or which develop an intelligent FAQ (Frequently Asked Questions) system for judges (Benjamins et al. 2004; Casanovas, Casellas & Vallbé 2009) or young legal professionals ([OPJK: Ontology of Professional Judicial Knowledge], Casanovas, Casellas & Vallbé 2009).
- (4) understand a domain, such as those which are more generally applied in law, e.g. the functional ontologies of law (based on *Ontolingua*) by Valente and Breuker (1994, 1999), and those of language of legal discourse by McCarty (1989) or

---

<sup>6</sup> Breuker et al. (2008) increased the previous list to 33 ontologies. In Casanovas, Sartor, Biasiotti & Fernández-Barrera (2011) the list reached more than 60 references.

<sup>7</sup> This author also considers that the term ‘ontology’ should not be used when referring to domain-independent knowledge representations—representation languages—. Also, although the origins of ontologies were related to knowledge sharing and reuse, most ontologies are built “with some application in mind.”

those more general ontologies used for knowledge representation (*Frame Ontology*) by Van Kralingen (1995). They all use general language for expressing legal knowledge.

The number of specialists who are working at present on legal ontologies is very high, although, as far as I know, none of the applications which have been designed so far is formally based on deep semantics or, in other words, on a functional linguistic model similar to the Lexical Constructional Model or the architecture offered by FunGramKB (see section 3). Moreover, none of the so called legal ontologies contains any development which covers the area of terrorism and organized crime from a procedural or criminal law perspective. Consequently, the methodology used in the construction of the *Globalcrimeterm* Subontology will be explained in sections 4 to 6,<sup>8</sup> focusing both on a brief description of the corpus collection/term extraction process and on the conceptual modelling, subsumption and hierarchisation of verbs related to procedural law and criminal events.

### 3. The architecture of FunGramKB

Over 20 years ago Velardi et al. (1991: 156) distinguished two well-defined strategies when describing meaning in NLP: the cognitive content in a lexical unit can be described by means of semantic features or primitives (conceptual meaning), or through associations with other lexical units in the lexicon (relational meaning). Strictly speaking, the latter doesn't give a real definition of the lexical unit, but it describes its usage in the language via 'meaning relations' with other lexical units. Bender (2009) and Perrián Pascual (2012) maintain that it is certainly easier to state associations among lexical units in the way of meaning relations than describing the cognitive content of lexical units formally, but the inference power of conceptual meaning is much stronger. Surface semantics can be adequate in some NLP systems, but the construction of a robust knowledge base guarantees its use in most NLP tasks,

---

<sup>8</sup> I refer the reader to the aforementioned project (see footnote 2): the development of the satellite ontology on criminal law in the domain of transnational terrorism and organized crime to be implemented in intelligent systems which aim to understand legal discourse automatically.

reinforcing thus the concept of resource reuse. This crucial distinction set up the foundations for FunGramKB, which can be defined as a multipurpose lexico-conceptual knowledge base for natural language processing systems and natural language understanding. This knowledge base is made up of three major knowledge levels, consisting in turn of several independent but interrelated modules. As shown in Periñán Pascual & Arcas Túnez (2010b) and figure 1 below, these are:

- (a) The lexical level (linguistic knowledge) comprising the *Lexicon*, which stores morphosyntactic, pragmatics and collocational information about lexical units in a specific language, and the *Morphicon*, which handles cases of inflectional morphology.
- (b) The grammatical level (linguistic knowledge), formed by the *Grammaticon*, which stores and captures the properties that are specific to the most relevant constructional families in the languages selected.
- (c) The conceptual level (non-linguistic knowledge) which consists of three modules:
  1. The *Ontology*, a hierarchical catalogue of the concepts that a person has in mind, so here is where semantic knowledge is stored in the form of meaning postulates. The ontology consists of a general-purpose module (i.e. Core Ontology) and several domain-specific terminological modules or satellite ontologies).
  2. The *Cognicon*, a repository of procedural knowledge which is stored by means of scripts, that is, conceptual schemata in which a sequence of stereotypical actions is organized on the basis of temporal continuity, and more particularly on Allen's temporal model (Allen 1983; Allen and Ferguson 1994). The scripts help you, for example, to describe *how to pay online*. or *how to launder money*.
  3. The *Onomasticon*, a repository of information about instances of entities and events such as *Bin Laden* or *September 11*. This module stores two different types of schemata (i.e. snapshots and stories), since instances can be portrayed synchronically or diachronically.

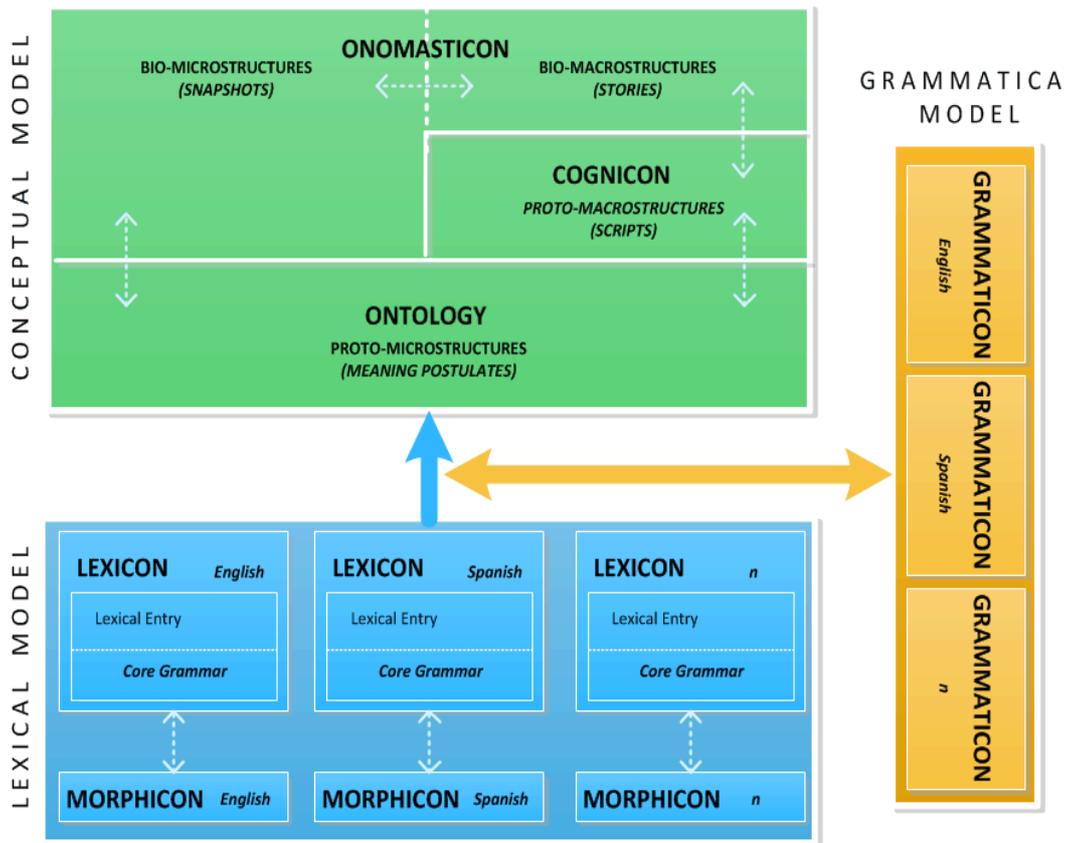


Figure 1. FunGramKB Architecture

Periñán Pascual & Arcas Túnez (2010a, 2010b) refer to the fact that only the lexical and grammatical models are language-dependent, while the modules that make up the conceptual level are all language-independent (shared by all the languages in the knowledge base). When working within the context of this knowledge base, this means that linguists will build one *Lexicon*, one *Morphicon* and one *Grammaticon* for each of the languages implemented in the knowledge base, whereas knowledge engineers will develop just one *Ontology*, one *Cognicon* and one *Onomasticon* to account for all the languages introduced in the system.

The *Ontology* is the module where conceptual meaning is modelled and is also the key component around which the whole knowledge base pivots. As can be seen in Figure 1, the FunGramKB Ontology distinguishes three different conceptual levels, each one of them with concepts of a different type (Periñán Pascual & Arcas Túnez, 2004: 38):

- (a) *Metaconcepts*, preceded by the symbol # (e.g. #ABSTRACT, #COMMUNICATION, #MATERIAL, #PHYSICAL, #PSYCHOLOGICAL, #QUANTITATIVE, #SOCIAL, etc.), constitute the upper level in the taxonomy and coincides with some of the most relevant linguistic ontologies, e.g. SIMPLE (Pedersen & Keson 1999, Lenci et al. 2000), SUMO (Niles & Pease 2001), DOLCE (Gangemi et al. 2005; Masolo et al. 2003), among others. The FunGramKB Ontology is divided into three subontologies, since subsumption (IS-A) is the only taxonomic relation permitted, and therefore each subontology arranges lexical units of a different part of speech: i.e. #ENTITIES for nouns (e.g. +JUDGE\_00, +SOUL\_00, +WINDOW\_00,...), #EVENTS for verbs (+BUILD\_00, +THINK\_00, etc.), and #QUALITIES for adjectives and some adverbs (i.e. +DIRTY\_00, +ALONE\_00...).
- (b) *Basic concepts*,<sup>9</sup> preceded by symbol + (e.g. +VIOLENT\_00, +CRUEL\_00, +CRIME\_00, +TRIAL\_00, +PUNISH\_00, +MURDER\_00, etc.), are used in FunGramKB as defining units which enable the construction of meaning postulates for basic concepts and terminals, as well as taking part as selectional preferences in thematic frames.
- (c) *Terminals* (e.g. \$ASSASSINATION\_00, \$FELONY\_00, \$GANGSTER\_00, \$CONSPIRE\_00, \$DISHONEST\_N\_00, etc.) are headed by the symbol \$. The borderline between basic concepts and terminals is based on their definitory potential to take part in meaning postulates. Hierarchical structuring of the terminal level is practically non-existent.

Basic and terminal concepts in FunGramKB are provided with semantic properties which are captured by *thematic frames* (TF) and *meaning postulates* (MP). Every event (or quality) in the ontology is assigned one single TF, e.g. a conceptual construct which states the number and type of participants involved in the prototypical cognitive

---

<sup>9</sup> The examples of basic and terminal concepts indicated here have been obtained from FunGramKB Core Ontology and the *Globalcrimeterm* Subontology. The original source for most of the basic concepts in the Core Ontology was a scrutinised reclassification of the defining vocabulary in the *Longman Dictionary of Contemporary English* (Procter 1978).

situation portrayed by the event (Periñán Pascual & Arcas Túnez 2007)<sup>10</sup>. Moreover, an MP is a set of one or more logically connected predications ( $e_1, e_2, \dots e_n$ ), e.g. conceptual constructs that represent the generic features of concepts. As stated above, the basic concepts are the main building blocks of these types of constructs in the Core Ontology.

Since metaconcepts and basic concepts are already defined in FunGramKB, it is worth noticing the importance of building adequate terminal concepts for a fine-grained knowledge base which is based on deep semantics. As a consequence, knowledge engineers have to cope with the modelling of ontological meaning which means not only deciding on the creation of terminal concepts, but also formalizing these concepts in COREL interface language or determining which lexical units should be linked to them. In the following sections, I will briefly explain the methodology used for the design of the *Globalcrimeterm Corpus* (henceforth GCTC)<sup>11</sup> and *FunGramKB Terminology Extractor* (henceforth FGKBTE) as a previous step towards a detailed description of the method employed for the conceptual modelling of the selected verbs (EVENTS) related to procedural and criminal law.

#### 4. Corpus design and terminological extraction

The initial stages in the process of corpus compilation included a number of decisions and selections that helped us to collect and organize the GCTC coherently and efficiently (Bowker & Pearson 2002, Koester 2010). To begin with, the legal subdomain of organized crime and terrorism was selected for its current international relevance and for the scarce NLP references on the topic, particularly with the purpose of populating ontologies. Therefore, the winning terms extracted from the GCTC helped us to populate both, the specific-domain subontology and the Core Ontology in the system of FunGramKB.

---

<sup>10</sup> We refer the reader to Periñán Pascual & Mairal Usón (2010) for examples of conceptual representation in the form of *thematic frames* and *meaning postulates*.

<sup>11</sup> This is a text collection which was compiled for the creation of the *Globalcrimeterm* Subontology: a satellite ontology on terrorism and organized crime integrated in FunGramKB.

#### 4.1 Corpus compilation

As previously stated in Felices Lago & Ureña Gómez-Moreno (2014), the *first step* in the compilation lies in the selection of sources, e.g. academic and professional repositories containing specialised documents on the topic(s) of interest. This step is of vital importance, since it will determine to a great extent whether the corpus is optimal, both qualitatively and quantitatively, for the purpose of term extraction. The selected sources must therefore meet high scientific standards or be highly regarded by the professional community. For example, the GCTC contains a selection of more than 10 sources, such as the *European Union* (EU), the *Council of Europe*, the *Organization for the Cooperation and Security in Europe* (OSCE), *Eurojust* or the *International Criminal Court* (ICC), which offer reliable information concerning cooperation against criminal and terrorist activities. In addition to the data sources, another important decision is the representativeness of the corpus (Biber 1993). In this respect, the GCTC consists of approximately 5,600,000 tokens from a wide variety of text types, including international treaties, fact sheets, rules, resolutions, conventions and acts, among others.<sup>12</sup> The corpus is also reasonably balanced with respect to the number of texts on the domains under study, i.e. 49% are focused on terrorism, while 35% deal with organized crime and 16% account for texts on both types of subject areas. Other sources, such as academic reference works and journal articles, were also considered due to the usual high concentration of specialised terms in their texts.

Once the relevant documents were selected and downloaded, a *second step* in the compilation of the corpus refers to text editing. A series of manual and semiautomatic editing tasks were required in order to filter out typographical mistakes resulting from the reformatting of original formats (usually *pdf*) to plain text. This preparatory pre-processing of the texts was necessary because of the characteristics of the term extractor tool (part of the *FunGramKB suite*)<sup>13</sup>, which only works with raw texts. Thus, whenever

---

<sup>12</sup> All the sources included in the GCTC were in English.

<sup>13</sup> *FunGramKB Suite* is the name used to refer to the knowledge engineering tool and FunGramKB is the resulting knowledge base. *FunGramKB Suite* was developed in C# using ASP.NET 2.0 platform and a MySQL database (Periñán Pascual & Arcas Túnez 2010b: 2667).

necessary, manual editing included tasks such as linking words which the automatic text converter had separated previously or subdividing longer documents when they exceeded file size.

Besides the compilation, formatting and editing of the corpus, the *third step* lies in the creation of a database which includes a record of all the relevant data of the documents in the collection such as the title, a brief description, the date in which the document was accessed and stored, as well as the name of the original source. Figure 2 shows the database created in the compilation of the GCTC<sup>14</sup>:

ID	Language	Brief description	Title	Topic	Type of document	Source
1	English	Fight against organised crime	EOAct (joint) law enforcement cooperation	Organised Crime	Joint Action	Eur-Lex
2	Spanish	Fight against organised crime	SOAct (joint) law enforcement cooperation	Organised Crime	Joint Action	Eur-Lex

Figure 2. Sample of corpus database

The data gathered in the database had three main objectives. First, they served as a guide to monitor criteria such as corpus balance and representativeness. Second, some of the data registered in the database could be used during the uploading of texts onto the extractor and had to be conveniently stored. Finally, the database also provided the documentary basis for the calculation of simple descriptive statistics about the corpus.

Once the GCTC was completed and closed, the following stage comprised the extraction of specialised terms, whose process is described in the following section.

#### 4.2 Term extraction process

Terminological extraction in the FGKBTE is based on corpus data, since this information can contribute to finding the relevant terminology used by professionals

---

<sup>14</sup> The first field, “ID”, assigns a unique numeric code to each text. The field “Language” contains information about the language in which the text is written. “Brief description” offers very succinct information about the contents of the text. “Title” provides a title that summarises the specific topic of the document. The “Topic” field, on the other hand, records the subdomain the text belongs to; in the case of GCTC, a distinction is drawn between “Organized crime”, “Terrorism” or “Both”. Finally, the field “Type of document” contains information about the text type (e.g. joint action, agreement, green paper, proceedings, etc.), while “Source” adds a reference on the source from which the original document was extracted.

and practitioners within a domain. Once the textual repository is set up, term extraction is the following step. As can be seen in figure 3, this process comprises two stages, an automatic phase and a manual phase:

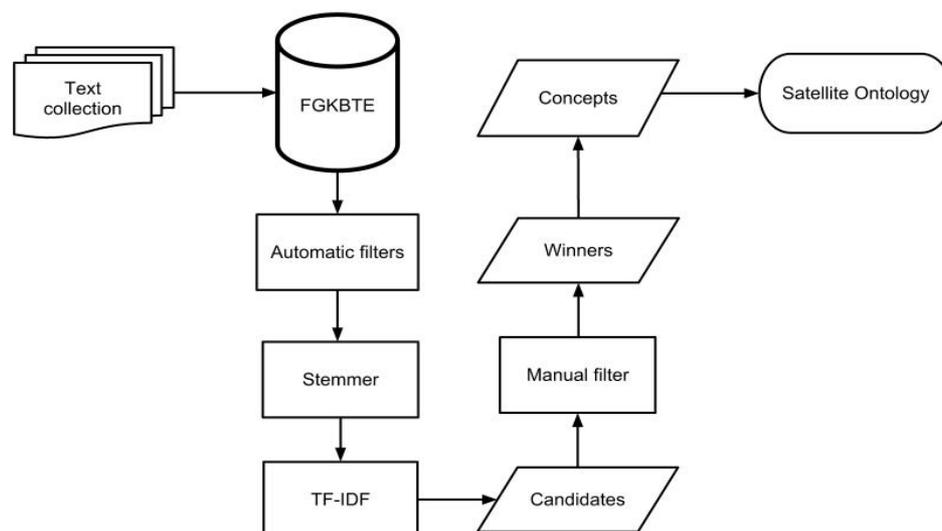


Figure 3. Flowchart of semi-automatic term extraction in FGKBTE

Automatic phase:

FGKBTE is the cornerstone in the process of creating satellite ontologies or subontologies (Felices Lago & Ureña Gómez-Moreno 2014, Periñán Pascual & Arcas Túnez, 2014). The main function of this tool is to retrieve terms, here roughly defined as specialised lexical units within a knowledge domain and despite the fact that they do not have clear boundaries. This definition should not ignore the distinction between technical terms, semi-technical terms and general vocabulary (Gómez González-Jover 2007) and the necessary exhaustive scrutiny of the selected units to decide their inclusion or not in a specific domain (for a more detailed discussion, see Cabré 1999, Temmerman 2000, Temmerman & Kerremans 2003 and L'homme 2004, 2006). The extractor applies a series of filters to an input corpus, mainly removal of non-textual characters, numbers and punctuation marks. It is upon this cleaned up text that the statistical extraction process operates. FunGramKB Extractor (FGKBTE) calculates a *tf-idf* score for each lexical unit in the corpus. As a result, the terminologist can work on a list of candidate terms ranked according to their semantic weight, so that candidates that appear higher in the list are statistically more relevant specialised terms, while elements

that show a *tf-idf* index below 3 are not statistically specialised. It is important to notice that the extraction process in FunGramKB is semi-automated and that the ultimate decision of what counts as a specialised term relies on the criterion of the terminologist. Figure 4 below shows the main menu of FGKBTE containing the principal functions of the tool:



Figure 4. Main menu of FunGramKB Term Extractor

From the top leftmost button: the “Pre-processing” tab contains an area for testing new features for the extractor. The “Processing (indexing)” tab is used for uploading texts of a corpus to the extractor. “Processing (statistics)” is a key function allowing the terminologist to automatically obtain the list of candidate terms from the corpus. “View” allows the terminologist to filter false terms by means of a series of removal options. The “Search” tab is a secondary tool for searching strings of text in a corpus. Finally, “Corpus” shows basic descriptive statistics concerning the number of indexed texts making up a given corpus as well as the number of tokens included. This tab also shows a terminological box containing a list of false candidates that were discarded during the filtering process tackled in the “View” function.

One of the most outstanding features of FGKBTE lies in its potential for filtering false candidates. The “View” mode contains for each term candidate an option for “simple removal”, so that if the terminologist chooses this option, a bigram such as “avoid transact” would be sent to the list of false candidates in “Corpus”. More interestingly, the extractor can also make complex removal of lexical bigrams and trigrams. For example, the nested removal of “avoid transact” will result in the removal of “avoid transact” as a bigram as well as in the removal of each component individually (“avoid” and “transact”).

Previous results emphasize the utility of this approach for term extraction. After uploading the components of the GTC to the extractor, which contains roughly 5,500,000 tokens, and applying the preparatory filters and the statistical processor

afterwards, the initial count was reduced to a set of approximately 5,700 candidate terms, a comparatively much smaller quantity of acceptable terms. It is important to emphasise that such a reduced set of candidates was reached in a short period of time, if compared to other approaches such as manual inspection of concordances or collocations.

Manual filtering<sup>15</sup>:

For identifying terms it is not enough to apply the previous statistical processor and decide about units' *termhood* on the bases of their statistical significance, since there are other theoretical problems to be faced. To facilitate term identification, terminologists should consider three additional criteria during the manual filtering process (Felices Lago & Ureña Gómez-Moreno 2014: 264-266):

- (1) *Ontological criterion*: To decide whether a candidate is a specialised unit, the speaker's mere introspection is sometimes a valid criterion. Within the framework of FunGramKB, introspection is carried out at the level of the Core Ontology, so that the question of whether a candidate is actually a term can be answered by means of another more specific question: does the Core Ontology contain a concept that could include this candidate as a possible lexical realisation?
- (2) *Lexicological criterion*: It relates to the lexicological features of the candidate terms. Terms were traditionally characterised by a univocal, unambiguous monosemic meaning. This misconception has been successfully overcome in the last decades. However, aspects such as meaning banalisation or the acquisition of new terminological senses in general language lexical units through processes of metaphorisation and metonymic mapping require the terminologists to check

---

<sup>15</sup> The whole process involved in the manual extraction (preceded by the automatic extraction) would correspond to the selection and acquisition phases described in the methodology for the construction of satellite ontologies (Periñán Pascual & Arcas Túnez 2014) and the COHERENT methodology (Periñán Pascual & Mairal Usón 2011).

whether a candidate term is polysemous or homonymous and, if so, decide which sense is technical and discard common knowledge meanings.

- (3) *Lexicographical criterion*: The most important criterion, nevertheless, is the consultation of specialised dictionaries, since they reflect the necessary knowledge for the understanding of expert knowledge. It is necessary to note that trained terminologists and lexicographers (with the advice and support of domain-specific experts and practitioners) are the best placed to determine and define terms, since they know how to concisely formulate a definition in a systematic way.

Consequently, considering all the criteria involved above the terminologist must conclude the analysis by determining the specific nature of the lexical candidate and the place it should occupy inside the Ontology.

## 5. Conceptual modelling

This section refers to the main procedural aspects concerning the transduction of terminological units into conceptual constructs, and the use of the latter in the population of the FunGramKB Ontology.<sup>16</sup> It is essential to emphasize that I will only be dealing with #EVENTS, and thus both #QUALITIES and #ENTITIES will be disregarded. Results expected at this stage include a set of terminal concepts, as well as a group of lexical units representing the linguistic expression of each of these cognitive categories. The terminal concepts will occupy slots within the Ontology (be it Core or Satellite), while the linguistic representations of terminal concepts (lexical units) will fill the corresponding slots in the lexicon of the language selected.

The final output of “events” linked to the domain of the international cooperation against terrorism and organized crime amounts to 49 concepts and their corresponding lexical units in the English lexicon. This was the result of the exhaustive scrutiny of the

---

<sup>16</sup> This process corresponds to the CONCEPTUALIZATION phase of the COHERENT methodology referred to above.

verbal concept candidates and their definition in natural language as a previous step to their conceptual modelling and hierarchisation.<sup>17</sup> The transduction of these definitions into conceptual constructs in COREL interface language creates the semantic properties which are captured by *thematic frames* (TF) and *meaning postulates* (MP) and represented in basic or terminal concepts in FunGramKB ontologies.<sup>18</sup> A complete list of the fully defined 49 concepts can be seen in Appendix 1. In order to illustrate how to define terms, let us consider the following example of the concept \$BRIBE\_00 included in the criminal law domain:

(1) *Term in the English lexicon:* Bribe

*Concept:* \$BRIBE\_00

*Thematic frame:* (x1: +HUMAN\_00) Agent (x2: +MONEY\_00 ^ +GIFT\_00)

Theme (x1) Origin (x3: +HUMAN\_00 & +OFFICIAL\_00) Goal

*Meaning postulate:* +(e1: +GIVE\_00 (x1: +HUMAN\_00)Agent (x2: +MONEY\_00 ^ +GIFT\_00)Theme (x1)Origin (x3: +HUMAN\_00 & +OFFICIAL\_00)Goal

\*(e2: +DO\_00 (x4)Theme (x5)Referent (f1: \$LEGAL\_N\_00 | +DISHONEST\_00)Attribute

*Description in natural language:* A person gives someone (especially a public official), money or a gift in order to persuade them to do something illegal for the person.

The lexical conceptual information of terminological units is introduced in FGKBTE by means of the “Edit” tool included in the “View” tab (see figure 4 above). “Edit” appears

---

<sup>17</sup> The role of the knowledge engineer at this stage is to gather the semantic content of a term from a selected number of dictionaries and to produce a general description in natural language which encompasses all the different lexicographical definitions. The most common sources have been the updated editions of the following reference works: *Longman Dictionary of Contemporary English*, *Oxford Dictionary of Law* or *Black’s Law Dictionary*, among others.

<sup>18</sup> I refer the reader to Periñán Pascual & Mairal Usón (2010) for an exhaustive description of the notation system used in the grammar of COREL, particularly the diverse satellites and operator used for the thematic frames and the predications of the meaning postulates.

next to each candidate term and by clicking on it the terminologist accesses the screen in figure 5.

view	edit					
<a href="#">view</a>	<a href="#">edit</a>	<input type="checkbox"/>	counteract	0.00146	<input type="checkbox"/>	<a href="#">remove</a>
<a href="#">view</a>	<a href="#">edit</a>	<input type="checkbox"/>	lift	0.00146	<input type="checkbox"/>	<a href="#">remove</a>
<a href="#">view</a>	<a href="#">edit</a>	<input checked="" type="checkbox"/>	bribe	0.00144	<input type="checkbox"/>	<a href="#">remove</a>
<a href="#">view</a>	<a href="#">edit</a>	<input type="checkbox"/>		0.00142	<input type="checkbox"/>	<a href="#">remove</a>

Terms: 553 context: 50 fragments: 2

**Senses:**  
#BRIBE\_00

**CONCEPT:**  
+BRIBE\_00

**DESCRIPTION:**  
To give a price, reward, gift or favor bestowed or promised with a view to pervert the judgment of or influence the action of a person in a position of trust.

**METACONCEPT:**  
#TRANSFER

Entities  Events  Qualities

[bribe](#) Save

Rename Delete  
 Done  Mirror

**+BRIBE\_00**

**English**  
 Y N

**Spanish**  
 Y N

**Italian**  
 Y N

Figure 5: Lexical conceptual editing interface of \$BRIBE\_00 in FGKBTE.

In order to illustrate the editing function further, this figure captures a screen with the dialog boxes shown above filled with information related to the term *bribe*. To understand each subsection included here, it is helpful to describe them starting with “Senses” and showing the other subsections in a clockwise fashion. “Senses” is aimed at storing the several senses of homonymous and polysemous terms. Each sense shall carry a distinctive numerical index (e.g. +SEIZE\_00, +SEIZE\_01 and so on). It is an automatic dialog box, i.e., the information displayed here is generated automatically after the information in the other dialog boxes has been introduced. The “Delete” and “Rename” options allow the terminologist to make corrections before eventually validating the term at work together with its lexical conceptual information. “Concept” is the label or the COREL name that serves as a host cognitive category of the terminological unit. “Description”, as the name suggests, is a space set aside for entering a description in natural language that captures the meaning of the concept. It is worth recalling that FGKBTE uses English as a lingua franca for this purpose. “Metaconcept” is completed automatically with the ontological data selected among “entities”, “events” or “qualities”. This option is a first contribution to the hierarchical organization of concepts in the domain. Once the fields “Concept”, “Description” and

“Metaconcept” are completed, if you click on “Save” the online information is automatically stored. “Duplication” will serve the purpose of creating mirror concepts. If the concept that is about to be introduced in the Subontology is already included in the Core Ontology, a note will appear prompting the engineer to create a mirror concept or to warn them not to repeat information. The last element in the “Edit” tool is the Lexicon, which gathers the different lexical realisations, in this case, terminological units, instantiating a concept.

FGKBTE is currently designed to interpret and process information in seven languages: English, Spanish, Italian, French, German, Bulgarian and Catalan. Moreover, it also allows the assignment one or more terms to each concept. As it was mentioned above, it is necessary that there is at least one lexeme for each concept in any of these (or other) languages. Once all the data mentioned above have been introduced, the engineer must click on “Done” and all the information will be validated definitely, although this process can be reversed in case further changes or corrections are needed. The importance of “Done” is that only the terms so validated (in this final validation process) will be included in the Ontology (see figure 6), while the rest will be discarded.

The screenshot shows the FunGramKB Editor interface. At the top, the title is "FunGramKB Editor" and the ontology name is "Core". On the left, a tree view shows 485 events, with the concept \$BRIBE\_00 selected. The right panel displays the "Conceptual Information" for \$BRIBE\_00, which includes the following details:

CONCEPT:	\$BRIBE_00
SUPERORDINATE(S):	+GIVE_00
THEMATIC FRAME:	(x1)Agent (x2)Theme (x3)Origin (x4)Goal
MEANING POSTULATE:	+(e1: +GIVE_00 (x1: +HUMAN_00)Agent (x2: +MONEY_00 ^ +GIFT_00)Theme (x1)Origin (x3: +HUMAN_00)Goal *(e2: +DO_00 (x4)Theme (x5)Referent (f1: \$LEGAL_N_00   +DISHONEST_00)Attribute
DESCRIPTION:	A person gives someone, especially a public official, money or a gift in order to persuade them to do something illegal for the person

Figure 6: the concept \$BRIBE\_00 and its integration in the Ontology of FunGramKB editor.

Besides the guidelines just mentioned as to how to define terms in FGKBTE, it is necessary to enter three additional caveats: Firstly, terminologists and knowledge engineers must be careful not to include the *definiens* within the *definiendum*; in other

words, definitions should not take the shape of paraphrases in which the word being defined is a component part itself, such as “if someone commits X (...)”, where “X” is the *definiens*. Secondly, terminologists should also avoid including examples in the *definiens* showing how this term is used in natural language. Thirdly, definitions shall conveniently be expressed using simple syntactic structures such as “S+P+O”, and, whenever possible, support themselves on the reiteration of keywords. The example of \$BRIBE\_00 in figure 6 illustrates a definition with a simple syntactic outline and the recursive use of simple but key concepts.

## 6. Hierarchisation process

The new stage to be reached is the hierarchisation phase, which deals with the establishment of hierarchical meaning relations among concepts in the domain. Designing a networked hierarchy will endow FunGramKB with the capacity to derive relevant and meaningful inferences, as well as to *understand* and produce knowledge for a specific user-defined goal. The present section deals with the details of conceptual-hierarchy construction.

Hierarchisation consists of determining for each terminological concept defined in FGKBTE its corresponding hyperordinate, subordinate(s) and sister concept(s). Hyperordinates are the most general type of units in the hierarchy and work as host concepts for the classification of one or more subordinate concepts. Each subordinate concept can in turn have one or more sister concepts, which are characterised by sharing common semantic features inherited from the hyperordinate. This arrangement of concepts is called the “IS-A” subsumption. An illustrative example of how inheritance and subsumption operate within the hierarchy of concepts is the terminal concept \$BRIBE\_00:

```
(2)  #EVENT> #MATERIAL> #MOTION> #TRANSFER>+TRANSFER_00>
      +GIVE_00> $BRIBE_00
```

In (2) the concept \$BRIBE\_00 depends on the metaconcept #TRANSFER, and, as a consequence, it inherits the prototypical scheme from the metaconcept and the subsequent superordinate concepts +TRANSFER\_00 and +GIVE\_00. The inherited

information is shown schematically in the thematic frame represented in (3), which can be interpreted as “someone transfers something to someone”:

(3) TF: (x1) Agent (x2) Theme (x3) Origin (x4) Goal

The thematic frame is then complemented with the meaning postulate, which expresses non-inherited specific information about the event expressed by *bribe*, which can be paraphrased as “a person (x1) gives something [money or gift] (x2) to somebody [generally a civil servant] (x3) so that the goal is that the person (x3) does something (x4) [illegal or dishonest] for the person (x1). Example (4) shows the MP contained in the event \$BRIBE\_00:

(4) MP: +(e1: +GIVE\_00 (x1: +HUMAN\_00)Agent (x2: +MONEY\_00 ^  
+GIFT\_00)Theme (x1)Origin (x3: +HUMAN\_00 & +OFFICIAL\_00)Goal  
\*(e2: +DO\_00 (x4)Theme (x5)Referent (f1: \$LEGAL\_N\_00 | DISHONEST\_00)  
Attribute

As seen in examples (3) and (4), which represent the relevant conceptual information of \$BRIBE\_00<sup>19</sup>, inheritance is crucial for knowledge organization in FunGramKB. It is moreover of paramount importance in case the knowledge base is intended for reasoning tasks of the utmost precision, as in legal practice, since semantic features must be inherited without causing incongruence or deriving erroneous conclusions.

Hierarchies of specialised concepts show the same classification tenets and share the same upper conceptual level as the Core Ontology. Therefore, in order to build the hierarchy consistently, the first step is to select the basic hyperordinate concepts under which the remaining concepts will be classified. In the case of the #EVENT subontology for the domain of international cooperation against terrorism and organized crime the diverse conceptual paths for the selected 49 criminal actions or procedural steps are classified as follows:

- #COMMUNICATION>+SAY\_00: \$ACQUIT\_00, \$CONFESS\_00,  
\$DECLARE\_00, \$INTERROGATE\_00, \$SENTENCE\_01, \$TESTIFY\_00.

---

<sup>19</sup> Example (1) above offers a full-fledged representation of the conceptual information of this terminal concept.

- **#COMMUNICATION>+SAY\_00>+BLAME\_00:** \$INCRIMINATE\_00,  
\$PROSECUTE\_00.
- **#COMMUNICATION>+SAY\_00>+REQUEST\_01:** \$APPEAL\_01.
- **#COMMUNICATION>+SAY\_00>+THREATEN\_00:** \$COERCE\_00,  
\$EXTORT\_00.
- **#MATERIAL>+DO\_00:** \$ABUSE\_00, \$CORRUPT\_00, \$DEPORT\_00,  
\$FREEZE\_01, \$INCITE\_00, \$OFFEND\_00, \$TORTURE\_00.
- **#MATERIAL>+DO\_00>+CREATE\_00:** \$BOOTLEG\_00, \$FORGE\_00.
- **#MATERIAL>+DO\_00>+CHANGE\_00>+DAMAGE\_00:** \$MUTILATE\_00.
- **#MATERIAL>+DO\_00>+DECEIVE\_00:** \$DEFRAUD\_00.
- **#MATERIAL>+DO\_00>+HIDE\_00:** \$PREVARICATE\_00.
- **#MATERIAL>+DO\_00>+KILL\_00:** \$EXECUTE\_00, \$MASSACRE\_00,  
+MURDER\_00 (-ASSASSINATE\_00)<sup>20</sup>.
- **#MATERIAL>+DO\_00>+MOVE\_00:** \$DIVERT\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+LEAVE\_00:** \$RELEASE\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+LEAVE\_00>+ESCAPE\_00:**  
\$ABSCOND\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+PUT\_00:** \$LAUNDER\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+TAKE\_01:** \$ARREST\_00,  
\$KIDNAP\_00, \$SMUGGLE\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+TRANSFER\_00>+GIVE\_00>**  
**+SELL\_00:** \$TRAFFIC\_00.
- **#MATERIAL>+DO\_00>+MOVE\_00>+TRANSFER\_00>+SEND\_00:**  
\$EXTRADITE\_00.
- **#MATERIAL>+DO\_00>+OBTAIN\_00:** +STEAL\_00\*<sup>21</sup>.

---

<sup>20</sup> As indicated by Jiménez Briones & Luzondo Oyón (2011:18), there are cases in which the conceptual specification takes place exclusively inside the thematic frame of a basic or terminal concept, without varying the MPs. When this occurs, “subconcepts” come into existence in FunGramKB, being codified in COREL by a preceding minus symbol and in capital letters. Consequently, -ASSASSINATE\_00 is a subconcept linked to +MURDER\_00, because (x2) or the referent entity will be +HUMAN\_00 & +FAMOUS\_00.

<sup>21</sup> The conceptual units having the asterisk are mirror concepts.

- **#MATERIAL>+DO\_00>+OBTAIN\_00>+TAKE\_00:** \$CAPTURE\_00, +SEIZE\_00\*.
- **#MATERIAL>+DO\_00>+STOP\_00:** \$CONNIVE\_00.
- **#MATERIAL>+DO\_00>+TRY\_00>+ATTACK\_00:** \$MOLEST\_00.
- **#MATERIAL>+DO\_00>+WORK\_00:** \$MANIPULATE\_00.
- **#MATERIAL>#MOTION>#TRANSFER>+TRANSFER\_00>+GIVE\_00:** \$BRIBE\_00, EMBEZZLE\_00.
- **#PSYCHOLOGICAL>#COGNITION>+THINK\_00:** \$CONSPIRE\_00.
- **#PSYCHOLOGICAL>#COGNITION>+THINK\_00>+EXAMINE\_00>+CHOOSE\_00>:** \$JUDGE\_00.
- **#PSYCHOLOGICAL>#PERCEPTION>+PERCEIVE\_00>+SEE\_00>+EXAMINE\_00:** \$INVESTIGATE\_00.
- **#STATIVE>#RELATIONAL>#IDENTIFICATION>+BE\_00:** \$DISBAND\_00.
- **#STATIVE>#RELATIONAL>#POSSESSION>+HAVE\_00:** \$IMPRISON\_00.

The possible disconnection between the diverse hierarchies of conceptual paths shown above and the way the “domain-specific” concepts are classified is only apparent and can be explained with a brief account of NLP in relation with the different approaches to ontology building.

In surface semantics, legal ontology engineers have been producing taxonomies and have established connections among units (or concepts) basing their assumptions on expert extra-linguistic information, for example, legal theories or deontic logic, but the reasoning capacity has been generally limited to very specific tasks. However, the way the concepts relate to each other in this proposal is based on deep semantics, which combines an extensive commonsense knowledge base (FunGramKB) and a reasoning engine. Consequently, the Ontology of FunGramKB (and the other two modules: Cognicon and Onomasticon) can work with two reasoning processes: MicroKnowing (Microconceptual Knowledge Spreading) and MacroKnowing (Macroconceptual-

Knowing Spreading).<sup>22</sup> Microknowing is performed by two types of reasoning mechanisms: inheritance and inference. Inheritance, for instance, strictly involves the transfer of one or several predications from a superordinate concept to a subordinate one in the ontology. On the other hand, inference is based on the structures shared between predications linked to conceptual units which do not take part in the same subsumption relation within the ontology. The application of these two mechanisms on the MPs allows FunGramKB to minimize redundancy and maximize the informative capacity of the knowledge base. Outside the scope of this article is the role played by macroknowing, which may allow the integration of the MPs of the “events”, “entities” or “qualities” in the Ontology with the cognitive macrostructures in the Cognicon or the episodic knowledge stored in the Onomasticon.

Another apparent contradiction in the building of this domain-specific ontology is the full inclusion of the “specific” verbal concepts in the Core Ontology. This methodological decision is based on a series of unpredictable results during the compilation phase.

As previously explained in section 4, the basic concepts of the Core Ontology come from a scrutinized reclassification of the defining vocabulary in the *Longman Dictionary of Contemporary English* (Procter 1978). However, in the case of domain-specific ontologies, the engineer must decide to what extent a percentage of basic commonsense concepts from the Core Ontology are useful for a given subontology and, on the other hand, to what extent it will be necessary to create new basic concepts for the proper classification of domain-based terminals. Precisely, the combination of lexicographical evidence and the hierarchical paths shown above furnish the evidence that all the relevant “events” in this supposedly specific domain need not be included in the *Globalcrimeterm* Subontology. Two reasons are given:

- (1) The way in which concepts relate to each other within a domain and the way in which this relationship should be represented in a hierarchical taxonomy is not always clear. Precisely for this reason, the role of the ontology engineer is to

---

<sup>22</sup> For an account of these two reasoning processes, see Perinián Pascual & Arcas Túnez (2005, 2007).

find out common ontological properties and to discern differences among the selected units. This premise may lead me to conclude that most of the apparently specialised conceptual units referred to criminal law in the area of international cooperation against terrorism and organized crime have eventually been included in the Core Ontology as terminal concepts, due to the fact that the semantic content of their corresponding lexical units can be found in widely used learner's dictionaries and, consequently, this conceptual information is generally known and used by the layperson.<sup>23</sup>

- (2) In the same vein, the expert knowledge in the area of legal and social sciences is more accessible to the non-specialised knowledge (common sense) of a layperson than in the case of natural or "hard" sciences. Moreover, the results of the GCTC term extraction have shown that most of the selected terms (specific verbs) included in this field are also well known by the general public.

## 7. Conclusions

A key factor for the development of this research has been the possibility to use FunGramKB, which was designed to cover many of the most noticeable problems currently faced by NLP and practitioners in the area of artificial intelligence. The main advantage of this knowledge base is its capacity to combine linguistic knowledge and human cognitive abilities within the same integrated system. The concept-orientated interlingua (COREL) serves to describe the properties of the different modules that integrate FunGramKB in the cognitive level. As a consequence, this knowledge base moves away from the traditional solutions based on surface semantics to offer a fully-fledged alternative in which linguistic information is grounded on conceptual structures representing human knowledge.

---

<sup>23</sup> The final outcome of the *Globalcrimeterm* project referred to in footnote 2 is that all of the specialised concepts included there have been "entities" (nouns), even if a large number of the relevant "entities" are also part of the Core Ontology.

Another relevant factor is the fact that most specialists who are working at present on legal ontologies are not developing applications formally inspired in deep semantics or, more specifically, in a functional linguistic model similar to the Lexical Constructional Model (LCM) or associated computational developments, such as FunGramKB. Moreover, none of the so called legal ontologies, as far as I know, contains any development which covers the area of terrorism and organized crime from a procedural or criminal law perspective. Consequently, the methodology used in the development of the *Globalcrimeterm* project focuses not only on a brief description of the corpus collection/term extraction process, but also on the conceptual modelling, subsumption and hierarchisation of verbs related to procedural law and criminal events. In this respect, this ontological construction based on the COHERENT methodology may contribute to a new perspective of analysis in the field of legal ontology building.

At a more specific level, the concept \$BRIBE\_00 has been used as a canonical instantiation of conceptual modelling and a similar process has been followed to represent the meaning of the remaining 48 events collected in Appendix 1. Moreover, the hierarchisation phase has demonstrated how the apparently 49 specialised concepts replicate the same classification tenets and share the same upper conceptual level as the basic concepts of the Core Ontology. In fact, the “specific” verbal concepts are eventually included in the Core Ontology and not in the domain-specific ontology, as previously calculated. In this respect the selected events collected here clearly differ from the selected “entities” or terminological nouns, which are generally integrated in the *Globalcrimeterm* Subontology. Among the reasons that could explain this unexpected result, it is worth noting the evidence provided by lexicographical sources, which show how the semantic content of the units linked to the selected concepts is not only known by legal practitioners but also shared by the average speaker of the language.

## References

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

- Allen, J. F. & Ferguson, G. (1994). Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4 (5), 531-579.
- Asaro, C. et al. (2003). A Domain Ontology: Italian Crime Ontology. In *Proceedings of the ICAIL 2003 Workshop on Legal Ontologies & Web Based Legal Information Management*, (pp. 1-7).  
([https://www.academia.edu/8305368/A\\_Domain\\_Ontology\\_Italian\\_Crime\\_Ontology](https://www.academia.edu/8305368/A_Domain_Ontology_Italian_Crime_Ontology))
- Bender, E. (2009). Linguistically naïve != Language independent: why NLP needs linguistic typology. In *Proceedings of the European Chapter of the ACL 2009 Workshop on the Interaction between Linguistics and Computational Linguistics* (pp. 26-32). Association for computational linguistics.  
(<http://www.aclweb.org/anthology/W09-0106>)
- Benjamins, V. R. et al. (2004). Ontologies of Professional Legal Knowledge as the Basis for Intelligent IT Support for Judges. *Artificial Intelligence and Law*, 12 (4), 359-378.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257.
- Boer, A., Hoekstra, R. & Winkels, R. (2001). The CLIME Ontology. In *Proceedings of the Second International Workshop on Legal Ontologies* (pp. 37-47). Amsterdam: Jurix.
- Bowker, L. & Pearson, J. (2002). *Working with Specialized Language. A Practical Guide to Using Corpora*. London and New York: Routledge.
- Breuker, J. & Hoekstra (2004). Core concepts of law: taking common-sense seriously. In *Proceedings of Formal Ontologies in Information Systems FOIS-2004* (pp. 210–221). Amsterdam: IOS-Press.
- Breuker, J., Valente, A. & Winkels, R. (2005). Use and Reuse of Legal Ontologies in Knowledge Engineering and Information Management. In V.R. Benjamins et al. (Eds.), *Law and the Semantic Web* (pp. 36-64). Berlin: Springer.
- Breuker J., Casanovas, P., Klein, M.A.C. & Francesconi, E. (Eds.). (2008). *Law, Ontologies and the Semantic Web*. Amsterdam: IOS Press.

- Breuker, J. & Hoekstra, R. (2011). A cognitive science perspective on legal ontologies. In G. Sartor, P. Casanovas, M. A. Biasiotti & M.Fernández-Barrera (Eds.), *Approaches to legal ontologies, theories, domains, methodologies* (pp. 69-81). Berlin: Springer.
- Cabré Castellví, M. T. (1999). *Terminology. Theory, Methods and Applications*. Amsterdam: John Benjamins.
- Cabré Castellví, M. T. et al. (2004). The GENOMA-KB project: towards the integration of concepts, terms, textual corpora and entities. In *Proceedings of the LREC 2004* (pp. 87-90). Lisbon: European Languages Resources Association.
- Casanovas, P., Casellas, N. & Vallbé J.J. (2009). An ontology-based Decision Support System for Judges. In H. Breuker et al. (Eds.) *Legal Ontologies and the Semantic Web. Channeling the Legal Information Flood* (pp. 165-175). Amsterdam: IOS Press.
- Casanovas, P., Sartor, G., Biasiotti, M.A. & Fernández Barrera, M. (2011). Theory and methodology in legal ontology engineering: Experiences and future directions. In G. Sartor, P. Casanovas, M. A. Biasiotti & M.Fernández-Barrera (Eds.), *Approaches to legal ontologies, theories, domains, methodologies* (pp. 3-14). Berlin: Springer.
- Faber, P. (2002). ONCOTERM: sistema bilingüe de información y recursos oncológicos. In A. Alcina & S. Gamero (Eds.), *La traducción científico-técnica y la terminología en la sociedad de la información* (pp. 177-188). Castelló: Universitat Jaume I.
- Faber, P. & Buendía Castro, M. (2014). EcoLexicon. In A. Abel, Ch. Vettori & N. Ralli (Eds.), *Proceedings of the XVI EURALEX International Congress*, (pp. 601–607). Bolzano: EURALEX.
- Felices Lago, Á. & Ureña Gómez-Moreno, P. (2012). Fundamentos metodológicos de la creación subontológica en FunGramKB. *Onomázein*, 26/2, 49-67.

- Felices Lago, Á. & Ureña Gómez-Moreno, P. (2014). FunGramKB Term Extractor: a key instrument for building a satellite ontology based on a specialized corpus. In B. Nolan & C. Perrián Pascual (Eds.), *Language processing and grammars: The role of functionally oriented computational models* (pp. 251-269). Amsterdam: John Benjamins.
- Gangemi, A., Sagri, M. and Tiscornia, D. (2005). A Constructive Framework for Legal Ontologies. In V.R. Benjamins et al. (Eds.), *Law and the Semantic Web* (pp. 97-124). Berlin: Springer.
- Gómez González-Jover, A. (2007). Léxico especializado y traducción. In E. Alcaraz Varó, J. Mateo Martínez & F. Yus Ramos (Eds.), *Las lenguas profesionales y académicas* (pp. 27–40). Barcelona: Ariel.
- Gruber, Th. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220.
- Hsieh, S. et al. (2010). Enabling the development of base domain ontology through extraction of knowledge from engineering domain handbooks. In *Advanced Engineering Informatics*. D.O.I.:10.1016/j.aei.2010.08.004.
- Jackendoff, R. (1990). *Semantic Structures*. Cambridge (Massachusetts): MIT Press.
- Jiménez Briones, R. & Luzondo Oyón, A. (2011). Building ontological meaning in a lexico-conceptual knowledge base. *Onomázein*, 23, 11-40.
- Kerremans, K., De Baer, P., Temmerman, R. (2007). Dealing with terminological variation in Termonography: examples from the “PoCeHRMOM”project. In *Proceedings of the XVIth European Symposium on Language for Special Purposes* (pp. 27–31).Hamburg: University of Hamburg.
- Koester, A. (2010). Building small specialised corpora. In A. O’ Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 66-79). London: Routledge.
- Kralingen, R.W. van (1995). *Frame-based Conceptual Models of Statute Law*. The Hague: Kluwer Law International.

- Lame, G. (2002). Construction d'ontologie à partir de textes. Une ontologie de droit dédiée à la recherche d'information sur le Web. PhD dissertation, Ecole des mines de Paris, Paris. (<http://www.cri.ensmp.fr/>).
- Leary, R.M., Vandenberghe, W. & Zeleznikow, J. (2004). Towards a financial fraud ontology: a legal modelling approach. In ICAIL 2003 Workshop on Legal Ontologies & Web based legal information management (pp. 1-33). (To unload it, go to [Leibnizcenter.org](http://Leibnizcenter.org)).
- Lenci, A. et al. (2000). SIMPLE: A general framework for the development of multilingual lexicon. *International Journal of Lexicography*, 13 (4), 249-263.
- Lenci, A. (Ed.). (2008). From context to meaning: distributional models of the lexicon in linguistics and cognitive science. *Italian Journal of Linguistics*, 20/1.
- Levin, B. & Rappaport, M.H. (2005). *Argument realization*. Cambridge: Cambridge University Press.
- L'Homme, M-C. (2004). *La terminologie: principes et techniques*. Montreal: Les Presses de l'Université de Montréal.
- L'Homme, M-C. (2006). Processing of Terms in Specialized Dictionaries: New Models and Technique. Special issue of *Terminology* 12:2 (2006). Amsterdam: John Benjamins.
- Liebwald, D. (2007). Semantic Spaces and Multilingualism in the Law: The Challenge of Legal Knowledge Management. In *LOAIT* (pp. 131-148).
- Mairal Usón, R. & Ruiz de Mendoza Ibáñez, F. J. (2008). New challenges for lexical representation within the Lexical-Constructional Model. *Revista Canaria de Estudios Ingleses*, 57, 137-158.
- Mairal Usón, R. & Ruiz de Mendoza Ibáñez, F. J. (2009). Levels of description and explanation in meaning construction. In Ch. Butler & J. Martín Arista (Eds.) *Deconstructing Constructions* (pp. 153–198). Amsterdam: John Benjamins.
- Masolo, C. et al. (2003). *WonderWeb Deliverable D18: Ontology Library*. Laboratory for Applied Ontology, ISTC-CNR.  
(<http://wonderweb.man.ac.uk/deliverables/documents/D18.pdf>)

- Maurel, D. (2008). Prolexbase: A multilingual relational lexical database of proper names. In Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008 (pp. 335-338). ELRA: Marrakech.
- McCarty, L.T. (1989). A Language for Legal Discourse, I. Basic Features. In Proceedings of the 2nd International Conference on Artificial Intelligence and Law (pp. 180-189). New York: ACM.
- Musen, M.A. (1992). Dimensions of knowledge sharing and reuse. Computers and Biomedical Research, 25, 435-467.
- Niles, I. & Pease, A. (2001). Towards a standard Upper Ontology. In Proceedings of the Second International Conference on Formal Ontology in Information Systems. Ogunquit. (<http://www.adampease.org/professional/FOIS.pdf>)
- Pedersen, B.S. & Keson, B. (1999). SIMPLE—Semantic information for multifunctional plurilingual lexica: some examples of Danish concrete nouns. Proceedings of the SIGLEX-99 Workshop. Maryland. ([http://clair.eecs.umich.edu/aan/paper.php?paper\\_id=W99-0507#pdf](http://clair.eecs.umich.edu/aan/paper.php?paper_id=W99-0507#pdf))
- Periñán Pascual, C. (2012). The situated common-sense knowledge in FunGramKB. Review of Cognitive Linguistics, 10(1), 184-214.
- Periñán Pascual, C. & Arcas Túnez, F. (2004). Meaning postulates in a lexico-conceptual knowledge base. Proceedings of the 15th International Workshop on Databases and Expert Systems Applications (pp. 38-42). Los Alamitos (California): IEEE.
- Periñán Pascual, C. & Arcas Túnez, F. (2005). Microconceptual-Knowledge Spreading in FunGramKB. Proceedings of the 9th IASTED International Conference on Artificial Intelligence and Soft Computing (pp. 239-244). Anaheim-Calgary-Zurich: ACTA Press.
- Periñán Pascual, C. & Arcas Túnez, F. (2007). Cognitive modules of an NLP knowledge base for language understanding. Procesamiento del Lenguaje Natural, 39, 197-204.
- Periñán Pascual, C. & Arcas Túnez, F. (2010a). Ontological commitments in FunGramKB. Procesamiento del Lenguaje Natural, 44, 27-34.

- Periñán Pascual, C. & Arcas Túnez, F. (2010b). The architecture of FunGramKB. In Proceedings of the 7th International Conference on Language Resources and Evaluation, European Language Resources Association (pp. 2667-2674). European Language Resources Association (ELRA).
- Periñán Pascual, C. & Arcas Túnez, F. (2014). La ingeniería del conocimiento en el dominio legal: La construcción de una Ontología Satélite en FunGramKB. *Revista Signos. Estudios de Lingüística*, 47(84), 113-139.
- Periñán Pascual, C. & Mairal Usón, R. (2010). La gramática de COREL. Un lenguaje de representación conceptual. *Onomázein*, 21, 11-45.
- Periñán Pascual, C. & Mairal Usón, R. (2011). The COHERENT methodology in FunGramKB” *Onomázein*, 24, 13-33.
- Procter, P. (Ed.) (1978). *Longman Dictionary of Contemporary English*. Harlow, Essex: Longman.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge (Massachusetts): MIT Press.
- Reinhart, T. (2006). *Interface strategies: Optimal and costly computations*. Cambridge (Massachusetts): MIT Press.
- Ruiz de Mendoza, F. J. & Mairal Usón, R. (2008). Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model. *Folia Linguistica*, 42/2, 355-400.
- Sartor, G., Casanovas, P., Biasiotti, M.A. & Fernández-Barrera, M. (Eds.), *Approaches to legal ontologies, theories, domains, methodologies*. Berlin: Springer.
- Sowa, J. F. (2000). Ontology, metadata, and semiotics. In ). B. Gantery & G. Mineau (Eds.), *Conceptual structures: logical, linguistics, and computational issues* (pp. 55-81). Berlin: Springer.
- Temmerman, R. (2000). *Towards New Ways of Terminology Description. The Sociocognitive-Approach*. Amsterdam: John Benjamins.
- Temmerman, R. & Kerremans, K. (2003). Termontography. Ontology building and the Sociocognitive Approach to terminology description. In Prague CIL17-conference. ([http://www.hf.uib.no/forskingskole/temmerman\\_art\\_prague03.pdf](http://www.hf.uib.no/forskingskole/temmerman_art_prague03.pdf))

- Ureña Gómez-Moreno, P., Alameda Hernández, Á. & Felices Lago, Á. (2011). Towards a specialised corpus of organized crime and terrorism. In M. L. Carrió et al. (Ed.), *La investigación y la enseñanza aplicadas a las lenguas de especialidad y a la tecnología* (pp. 301-306). Valencia: Universitat Politècnica.
- Valente, A. (1995). *Legal knowledge engineering: a modelling approach*. Amsterdam: IOS Press.
- Valente, A. (2005). Types and Roles of Legal Ontologies. In V.R. Benjamins et al. (Eds.), *Law and the Semantic Web* (pp. 65-76). Berlin: Springer.
- Valente, A. & Breuker, J. (1994). A functional view of law. In G. Bargellini and S. Binazzi, (eds.) *Towards a global expert system in law*. Padua: CEDAM Publishers.
- Valente, A. & Breuker, J.(1999). Legal Modelling and Automated Reasoning with ONLINE. *International Journal of Human-Computer Studies*, 51(6), 1079-1125.
- Van Valin, R.D. Jr. (2005). *The Syntax-Semantics-Pragmatics Interface: An Introduction to Role and Reference Grammar*, Cambridge: Cambridge University Press.
- Velardi, P., Pazienza, M.T. & Fasolo, M. (1991). How to encode semantic knowledge: a method for meaning representation and computer-aided acquisition. *Computational Linguistics*, 17/2, 153-170.
- Vossen, P. (2003). Ontologies. In R. Mitkov (Ed.). *The Oxford Handbook of Computational Linguistics* (pp. 464-482). Oxford: Oxford University Press.
- Zelevnikow, J. & Stranieri, A. (2001). An ontology for the construction of legal decision support systems. In *Proceedings of the Second International Workshop on Legal Ontologies* (pp. 67-76)

## Appendix

Verbal concepts (events) included in the Core Ontology after their previous extraction from the *Globalcrimeterm* Corpus and their subsequent conceptual modelling. In alphabetical order:

felices: verbs 144

\$ABUSE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +DO\_00 (x1)Theme (x3)Referent (f1: (e2: +BE\_01 (x1)Theme (x4: +CRUEL\_00 | +VIOLENT\_00)Attribute (f2: x2)Goal))Manner)

\$ABSCOND\_00

TF: (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

MP: +(e1: +ESCAPE\_00 (x1)Agent (x2)Theme (x3)Location (x4: +PLACE\_00)Origin (x5: +PLACE\_00)Goal (f1: +FAST\_00 ^ +SECRET\_00)Manner)

+(e2: n +BE\_01 (x6)Theme (x7: +LEGAL\_00)Attribute)

\*(e3: +ESCAPE\_00 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f2: (e4: +STEAL\_00 (x1)Theme (x3: +ARTEFACT\_00)Referent))Manner)

\$ACQUIT\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +SAY\_00 (x1: +JUDGE\_00)Theme (x4: (e2: past n +DO\_00 (x5: +HUMAN\_00) Theme (x2: +CRIME\_00)Referent))Referent (x3)Goal (f1: +TRIAL\_00)Scene)

\$APPEAL\_01

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +REQUEST\_01 (x1: +HUMAN\_00)Theme (x4: (e2: +CHANGE\_00 (x3: +JUDGE\_00) Theme (x2: +OPINION\_00)Referent))Referent (x3)Goal)

\$ARREST\_00

TF: (x1: +POLICE\_00)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

MP: +(e1: +TAKE\_01 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

(f1: (e2: +THINK\_00 (x1)Theme (x6: (e3: past +DO\_00 (x2)Theme

(x7: +CRIME\_00)Referent))Referent))Reason)

(-ASSASSINATE\_00) / +MURDER\_00

TF: (x1)Theme (x2: +HUMAN\_00 & +FAMOUS\_00)Referent

MP: +(e1: +KILL\_00 (x1)Theme (x2)Referent (f1)Instrument (f2: (e2: +EXIST\_00 (x3: +CRIME\_00)Theme))Result)

\$BOOTLEG\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +CREATE\_00 (x1: +HUMAN\_00)Theme (x2: +ALCOHOL\_00 & +PRODUCT\_00)Referent) +(e2: +SELL\_00 (x1)Agent (x2)Theme (x3)Origin (x4)Goal)

+(e3: +TAKE\_01 (x1)Agent (x2)Theme (x5)Location (x3)Origin (x4)Goal)

+(e4: n +BE\_01 (x2)Theme (x6: +LEGAL\_00)Attribute)

\$BRIBE\_00

TF: (x1) Agent (x2) Theme (x3) Origin (x4) Goal

MP: +(e1: +GIVE\_00 (x1: +HUMAN\_00)Agent (x2: +MONEY\_00 ^ +GIFT\_00)Theme (x1)Origin (x3: +HUMAN\_00 & +OFFICIAL\_00)Goal

\*(e2: +DO\_00 (x4)Theme (x5)Referent (f1: \$LEGAL\_N\_00 | +DISHONEST\_00)Attribute

\$CAPTURE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +TAKE\_00 (x1: +HUMAN\_00)Theme (x2: +HUMAN\_00)Referent (f1: (e2: +STORE\_00 (x1)Theme (x2)Referent))Purpose (f2: (e3: +BECOME\_00 (x2)Theme

(x3: +PRISONER\_00)Attribute))Result)

\$COERCE\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +THREATEN\_00 (x1: +HUMAN\_00)Theme (x2)Referent (x3: +HUMAN\_00)Goal (f1: +VIOLENT\_00)Manner (f2: (e2: +DO\_00 (x3)Theme (x4)Referent))Purpose)

## felices: verbs 145

### \$CONFESS\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +SAY\_00 (x1: +HUMAN\_00)Theme (x4: (e2: past +DO\_00 (x1)Theme (x2: +CRIME\_00)Referent))Referent (x3: +HUMAN\_00)Goal)

### \$CONNIVE\_00:

TF: (x1)Theme (x2)Referent

MP: +((e1: n +STOP\_00 (x1)Theme (x4: (e2: +DO\_00 (x3)Theme (x2)Referent))Referent)(e3: +BE\_01 (x2)Theme (x5: +WRONG\_00)Attribute))

### \$CONSPIRE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +THINK\_00 (x1)Theme (x2)Referent (f1: +SECRET\_00)Manner (f2: +HUMAN\_00)Company (f3: ((e2: +DO\_00 (x1)Theme (x3)Referent (f2)Company) (e3: n +BE\_01 (x3)Theme (x4: +LEGAL\_00)Attribute)))Purpose)

### \$CORRUPT\_00

TF: (x1)Theme (x2)Referent

MP: +((e1: +DO\_00 (x1)Theme (x3: (e2: ing +DO\_00 (x4: +HUMAN\_00)Theme (x2)Referent))Referent)(e3: +BE\_01 (x2)Theme (x5: +DISHONEST\_00)Attribute))

### \$DECLARE\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +SAY\_00 (x1)Theme (x2)Referent (x3)Goal) +(e2: +BE\_01 (x2)Theme (x4: +TRUE\_00)Attribute)

### \$DEFRAUD\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +DECEIVE\_00 (x1: +HUMAN\_00)Theme (x2: +HUMAN\_00)Referent (f1: (e2: +EXPERIENCE\_00 (x2)Theme (x3: +DAMAGE\_01)Referent))Result | (f2: (e3: egr +HAVE\_00 (x2)Theme (x4)Referent))Result)

### \$DEPORT\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +DO\_00 (x1)Theme (x2: (e2: +LEAVE\_00 (x1)Agent (x3: +HUMAN\_00)Theme (x4)Location (x5: +COUNTRY\_00)Origin (x6: +COUNTRY\_00)Goal))Referent) +(e3: +BE\_02 (x3)Theme (x5)Location (f1:(e4: n +BE\_01 (x5)Theme (x7: +LEGAL\_00)Attribute))Reason)

### \$DISBAND\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: egr +BE\_00 (x1)Theme (x2: +ORGANIZATION\_00)Referent)

### \$DIVERT\_00

TF: (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

MP: +(e1: +MOVE\_00 (x1)Agent (x2: +MONEY\_00)Theme (x3)Location (x4)Origin (x5)Goal (f1: (e2: n +PERMIT\_00 (x6)Theme (x7)Referent (x1)Goal))Scene)

### \$EMBEZZLE\_00

TF: (x1)Agent (x2)Theme (x1)Origin (x3)Goal

MP: +(e1: +GIVE\_00 (x1: +HUMAN\_00 ^ +COMPANY\_00)Agent (x2: +MONEY\_00)Theme (x1)Origin (x3: +HUMAN\_00)Goal (f1: (e2: +USE\_00 (x3)Theme (x2)Referent (f2: (e3: n +BE\_01 (x2)Theme (x4: +LEGAL\_00)Attribute))Manner))Scene)

### \$EXECUTE\_00

TF: (x1)Theme(x2)Referent

MP: +(e1:+KILL\_00(x1)Theme(x2)Referent(f1:+PUNISHMENT\_00)Purpose)

## felices: verbs 146

### \$EXTORT\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +THREATEN\_00 (x1: +HUMAN\_00)Theme (x2)Referent (x3: +HUMAN\_00)Goal (f1: (e2: +GIVE\_00 (x3)Agent (x4: +MONEY\_00)Theme (x3)Origin (x1)Goal)))Purpose)

### \$EXTRADITE\_00

TF: (x1)Agent (x2: +CRIMINAL\_00)Theme (x3)Origin (x4)Goal

MP: +(e1: +TRANSFER\_00 (x1)Agent (x2)Theme (x3: +COUNTRY\_00)Origin (x4: +COUNTRY\_00)Goal)

### \$FORGE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +CREATE\_00 (x1: +CRIMINAL\_00)Theme (x2: +DOCUMENT\_00)Referent (f1)Instrument (f2:(e2: n +BE\_01 (x2)Theme (x3: +TRUE\_00)Attribute)))Result  
+(e3: +CHANGE\_00 (x1)Theme (x2)Referent (f3: (e4: +BECOME\_00 (x2)Theme (x4: +REAL\_00)Attribute)))Result)

### \$FREEZE\_01

TF: (x1)Theme (x2)Referent

MP: +(e1: +STOP\_00 (x1)Theme (x4: (e2: +USE\_00 (x3)Theme (x2: +MONEY\_00)Referent))) Referent)

### \$INCITE\_00

TF:(x1)Theme (x2)Referent

MP: +((e1: +DO\_00 (x1: +HUMAN\_00)Theme (x4: (e2: +DO\_00 (x3: +HUMAN\_00)Theme (x2)Referent)))Referent(e2: +BE\_01 (x2)Theme (x3: +VIOLENT\_00)Attribute))  
+(e3: n +BE\_01 (x2)Theme (x4: +LEGAL\_00)Attribute)

### \$IMPRISON\_00

TF: (x1)Theme (x2)Referent

MP:+(e1: +HAVE\_00 (x1)Theme (x2: +HUMAN\_00)Referent (f1: +PRISON\_00)Location (f2: (e2: past +DO\_00 (x2)Theme (x3: +CRIME\_00)Referent)))Reason)

### \$INCRIMINATE\_00

TF:(x1)Theme (x2)Referent (x3)Goal

MP:+(e1: +SAY\_00 (x1)Theme (x4: (e2: past +DO\_00 (x3)Theme (x2)Referent)))Referent (x3)Goal  
(e3: +BE\_01 (x2)Theme (x5: +BAD\_00)Attribute))

### \$INTERROGATE\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +SAY\_00 (x1: +HUMAN\_00)Theme (x2: m +QUESTION\_00)Referent (x3: +HUMAN\_00)Goal (f1: +LONG\_01)Duration (f1: (e2: +OBTAIN\_00 (x1)Theme (x4: +INFORMATION\_00)Referent)))Purpose)

### \$INVESTIGATE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +EXAMINE\_00 (x1)Theme (x2: +HUMAN\_00 ^ +CRIME\_00 ^ +STATE\_00)Referent (f1: (e2: +KNOW\_00 (x1)Theme (x3)Referent)))Purpose)

### \$JUDGE\_00

TF:(x1)Theme (x2)Referent

MP: +(e1: +CHOOSE\_00 (x1: +JUDGE\_00)Theme (x2: (e2: +BE\_00 (x3)Theme (x4: +CRIMINAL\_00)Referent)))Referent (f1: +COURT\_00)Location)

### \$KIDNAP\_00

TF: (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

MP: +(e1: +TAKE\_01 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1: (e2: perm n +ESCAPE\_00 (x1)Agent (x2)Theme (x6)Location (x5)Origin (x7)Goal)))Result (f2: (e3: +OBTAIN\_00 (x1)Theme (x8: +MONEY\_00)Referent)))Purpose)

felices: verbs 147

+(e1: +TAKE\_01 (x1)Agent (x2: +HUMAN\_00)Theme (x3)Location (x4)Origin (x5)Goal (f1: (e2: n perm +ESCAPE\_00 (x1)Agent (x2)Theme (x6)Location (x5)Origin (x7)Goal))Result (f2: (e3: +OBTAIN\_00 (x1)Theme (x8: +MONEY\_00)Referent))Purpose)

\$LAUNDER\_00

TF: (x1)Agent (x2)Theme (x3)Origin (x4)Goal

MP: \*((e1: +PUT\_00 (x1)Agent (x2: +MONEY\_00)Theme (x3)Origin (x4: s +BANK\_00 ^ s +COMPANY\_00)Goal) (e2: +SEEM\_00 (x2)Theme (x5: +LEGAL\_00)Attribute))

\*(e3: n +DISCOVER\_00 (x6: +GOVERNMENT\_00)Theme (x2)Referent)

\$MANIPULATE\_00

(x1)Theme (x2)Referent

MP: +(e1: +WORK\_00 (x1)Theme (f1: (e2: +OBTAIN\_00 (x1)Theme (x2)Referent))Purpose

(f2: (e3: +DESIRE\_01 (x1)Theme (x2)Referent))Condition)

\$MASSACRE\_00

TF: (x1:+HUMAN\_00)Theme(x2:+HUMAN\_00^+ANIMAL\_00)Referent

MP: +(e1:+KILL\_00(x1)Theme(x2)Referent(f1:+VIOLENT\_00&+CRUEL\_00)Manner)

\$MOLEST\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +ATTACK\_00 (x1: +ADULT\_00)Theme (x2: +CHILD\_00 ^ +ADULT\_00)Referent

(f1: +SEX\_00)Means)

+MURDER\_00

TF:(x1)Theme (x2)Referent

MP: +(e1: +KILL\_00 (x1)Theme (x2)Referent (f1)Instrument (f2: (e2: +EXIST\_00

(x3: +CRIME\_00)Theme))Result)

\$MUTILATE\_00

TF:(x1)Theme (x2)Referent

MP: +(e1: +DAMAGE\_00 (x1)Theme (x2: +HUMAN\_00)Referent (f1: (e2: egr +HAVE\_00 (x2)Theme

(x3: i +LIMB\_00)Referent))Result)

\$OFFEND\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +DO\_00 (x1: +HUMAN\_00)Theme (x2: +CRIME\_00)Referent)

\$PREVARICATE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +HIDE\_00 (x1)Theme (x2: +CRIME\_00)Referent (f1: (e2: +LIE\_01 (x3)Theme (x4)Referent (x5)Goal))Manner)

\$PROSECUTE\_00

TF:(x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +BLAME\_00 (x1)Theme (x2: +CRIME\_00)Referent (x3)Goal)

+(e2: +DEMONSTRATE\_00 (x4: +LAWYER\_00)Theme (x5:(e3: +BE\_00 (x3)Theme (x6: +CRIMINAL\_00)Referent))Referent (f1)Goal (f2: +COURT\_00)Location)

\$RELEASE\_00

TF: (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal

MP: +(e1: perm +LEAVE\_00 (x1)Agent (x2: +HUMAN\_00)Theme (x3)Location (x4)Origin (x5)Goal

(f1: (e2: past +STORE\_00 (x1)Theme (x2)Referent))Condition)

\$SENTENCE\_01

TF: (x1)Theme (x2)Referent (x3)Goal

MP: +(e1: +SAY\_00 (x1: +JUDGE\_00)Theme (x2: +CRIMINAL\_00)Referent

(x3: +PUNISHMENT\_00)Goal (f1:(e2: past +DO\_00 (x2)Theme (x3: +CRIME\_00)Referent)) Reason)

+SEIZE\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +TAKE\_00 (x1)Theme (x2)Referent (f1: +HAND\_00)Instrument (f2: +FAST\_00 & +STRONG\_00)Manner)

*Mirror concept:* +((e1: +TAKE\_ (x1: +HUMAN\_00)Theme (x2)Referent) (e2: n +HAVE\_00 (x1)Theme (x2)Referent) (f1: (e3: past +OBTAIN\_00 (x3)Theme (x2)Referent)) (f2: \$LEGAL\_N\_00)Manner) Reason)))

\$SMUGGLE\_00

TF: (x1: +CRIMINAL\_00 ^ +GROUP\_00)Agent (x2: s +MONEY\_00 ^ s +TOBACCO\_00 ^ s +DRUG\_00 ^ s +WEAPON\_00)Theme (x3)Location (x4: +COUNTRY\_00)Origin (x5: +COUNTRY\_00)Goal

MP: +(e1: +TAKE\_01 (x1)Agent (x2)Theme (x5)Location (x4)Origin (x3)Goal

(f1: (e2: n +PAY\_00 (x1)Agent (x6: s +TAX\_00)Theme (x7)Origin (x8: +GOVERNMENT\_00)Goal)) Scene)

\*(e3: +SEIZE\_00 (x9: +OFFICER\_00)Theme (x2)Referent)

+STEAL\_00

TF: (x1)Theme (x2)Referent

MP: +(e1: +OBTAIN\_00 (x1)Theme (x2)Referent (f1: +PLACE\_00)Origin (f2: +CRIME\_00) Result)

*Mirror concept:* +(e1: +USE\_00 (x1)Theme (x2: +VIOLENCE\_00)Referent (f1: (e2: +DO\_00 (x1)Theme (x3: +THEFT\_00)Referent))Purpose)

\$TESTIFY\_00

TF: (x1)Theme (x2)Referent (x3)Goal

MP: \*((e1: +SAY\_00 (x1: +HUMAN\_00)Theme (x2)Referent (x3)Goal (f1: +COURT\_00)Scene)

(e2: +BE\_01 (x2)Theme (x4: +TRUE\_00)Attribute))

\$TORTURE\_00

TF:(x1)Theme (x2)Referent

MP: \*(e1: +DO\_00 (x1: +HUMAN\_00)Theme (x2: (e2: +SUFFER\_00 (x1)Agent (x3: +HUMAN\_00)Theme (f1: +MUCH\_00)Quantity)))Referent (f2: (e3: +SAY\_00 (x3)Theme (x4)Referent (x1)Goal))Purpose)

\$TRAFFIC\_00

TF: (x1)Agent (x2)Theme (x3)Origin (x4)Goal

MP: +(e1: +SELL\_00 (x1: +CRIMINAL\_00)Agent (x2: +HUMAN\_00 | +PRODUCT\_00)Theme (x3)Origin (x4)Goal (f1: (e2: n +BE\_01 (x2)Theme (x5: +LEGAL\_00) Attribute))Manner)

Received: December 19, 2014

Accepted: September 29, 2015

Published: February 29, 2016

Updated: March 2, 2016