La estructura del componente léxico en una base de conocimiento léxico-conceptual
The structure of the lexical component in a lexico-conceptual knowledge base

Resumen: En los últimos años se ha desarrollado un modelo lingüístico, el Modelo Léxico
Construccional (Mairal y Ruiz de Mendoza, 2008, 2009; Ruiz de Mendoza y Mairal, 2008) que
proporciona una descripción detallada de de todos los aspectos del significado, incluso de los
aspectos pragmáticos y discursivos. En trabajos recientes en el marco del MLC (Periñán y Areas,
2004, 2005, 2007a, 2010a, 2010b) se ha diseñado una base de conocimiento léxico-conceptual,
FunGramKB, que permite desarrollar aplicaciones para el procesamiento del lenguaje natural.
FunGramKB se estructura en dos niveles (un nivel léxico y un nivel conceptual) que a su vez
comprenden varios módulos. El presente artículo se centra en el lexicon, uno de los dos módulos
del nivel léxico. Este trabajo tiene dos objetivos: explicar el diseño del componente léxico en
FunGramKB, y mostrar la representación semántica en el lexicon de los conceptos incluidos en la
ontología mediante un estudio contrastivo de las entradas léxicas de los predicados de habla en
francés e inglés.

Palabras clave: Modelo Léxico Construccional, FunGramKB lexicon, predicado, entrada léxica,
Aktionsart.

Abstract: In the last years a language model, the Lexical Constructional Model (Mairal & Ruiz
de Mendoza, 2008, 2009; Ruiz de Mendoza & Mairal, 2008) has been developed to account for
all facets of meaning, including pragmatic and discourse phenomena. Recent research in the
LCM (Periñán & Arcas, 2004, 2005, 2007a, 2010a, 2010b) has built an online lexico-conceptual
base, FunGramKB, for natural language processing (NLP) applications. FunGramKB consists of
two levels of information (i.e. a lexical level and a conceptual level) which in turn comprise
several modules. The present contribution focuses on the lexicon, one of the two modules in the
lexical level. The aim of the paper is twofold: first, to analyse the FunGramKB lexical
component; second, to illustrate the semantic representation in the lexicon of the concepts stored

1 Financial support for this research has been provided by Project no. FFI2010-17610, Ministry of Science and
Innovation, Spain. We would like to thank Prof. Mairal Usón for his guidance in shaping this paper.
in the ontology through a contrastive analysis of the French and English lexical entries for speech act predicates.

**Key words:** Lexical Constructional Model, FunGramKB lexicon, predicate, lexical entry, Aktionsart.

1. Introduction

The Lexical Constructional Model (hereafter LCM[^2]; Mairal & Ruiz de Mendoza, 2008, 2009; Ruiz de Mendoza & Mairal, 2008, among others) is a meaning construction model elaborated in the research group Lexicom (www.lexicom.es). As mentioned in Ruiz de Mendoza & Mairal (2008: 355), the LCM approach is based on moderate functional models of language (especially Role and Reference Grammar) and cognitive linguistics (particularly Goldberg’s Construction Grammar and Lakoff’s Cognitive Semantics). The model provides a thorough semantic description inasmuch as it accounts for all those aspects involved in meaning construction, including traditional implicature, illocutionary force and discourse coherence.

The LCM posits the following modules of semantic description (Ruiz de Mendoza & Mairal 2008; Mairal & Periñán, 2009b):

(i) The level 1 or argumental module is concerned with the semantic representation of predicates in terms of lexical templates.

(ii) The implicational or level 2 layer deals with low-level inferences.

(iii) The level 3 or illocutionary module accounts for illocutionary force.

(iv) The level 4 or discourse layer addresses discourse aspects, especially cohesion and coherence phenomena.

These four layers are interrelated by two cognitive processes: subsumption and cueing. This means that each level is either subsumed into a higher level module or acts as a cue for the activation of relevant conceptual structures that yields an implicit meaning derivation.

The basic organization of the LCM is given in figure 1.

[^2]: Abbreviations employed in this article: CLS ‘Conceptual Logical Structure’; FunGramKB ‘Functional Grammar Knowledge Base’; LCM ‘Lexical Constructional Model’; NLP ‘natural language processing’; RRG ‘Role and Reference Grammar’.
Figure 1. The overall architecture of the Lexical Constructional Model

LT = lexical template; CT = constructional template; CS = Conceptual Structure

In the present contribution we focus on the FunGramKB lexicon, which can be considered as one of the key components upon which the knowledge base is built. Along with the Morphicon, the
Lexicon, which stores morphosyntactic, pragmatic and collocational information about lexical units, is one of the linguistic modules that make up FunGramKB. It is necessary to highlight that one of the distinguishing features of FunGramKB, which makes it an incomparable model of linguistic representation, is that the lexicon provides a rich semantic description of lexical units including a fine-grained definition, grammatically relevant semantic information and pragmatic information.

Within this framework, our goal is to present the anatomy of the lexical component in FunGramKB, and to illustrate it through the analysis of the representation of speech act verbs in the French and English lexica.

The organization of the paper has the following format. Section 2 gives an overview of FunGramKB, with a special emphasis on the elements of the lexicon. Section 3 describes the organization of the FunGramKB lexicon and the features in FunGramKB lexical entries, as shown in www.fungramkb.com. This is followed by an account of the English and French lexical entries for the predicates belonging to the domain of SPEECH (or subsumed under the metaconcept #COMMUNICATION in FunGramKB). Finally, Section 4 presents a few concluding remarks.

2. The overall organization of FunGramKB

FunGramKB is a multipurpose (i.e. multifunctional and multilingual) lexico-conceptual knowledge base for NLP systems designed as part of the LCM (Periñán & Arcas, 2004, 2005, 2007a, 2007b, 2010a, 2010b; Mairal & Periñán 2009a, 2009b, 2010; Periñán & Mairal 2009). It is multifunctional and multilingual in the sense that it is meant to be reused in NLP tasks, e.g. document retrieval, information extraction and text categorization, and with many Western languages, including English, Spanish, French, German and Italian.

It is worth mentioning that the FunGramKB French version is an exceptional knowledge base in that it fills a gap in French computational lexicography, where the only language resources so far developed have been dictionaries and corpora. Further, it integrates rich morphological, syntactic, semantic and pragmatic information of lexical units.

FunGramKB comprises two interrelated information levels: a lexical level and a conceptual level, which in turn consist of several independent but interrelated modules.
The lexical level is made up of the lexicon and the morphicon, while the conceptual level consists of the ontology, the cognicon and the onomasticon.

**Lexical level (i.e. linguistic knowledge)**

- The lexicon stores morphosyntactic, pragmatic and collocational information about lexical units.
- The morphicon handles cases of inflectional morphology.

**Conceptual level (i.e. non-linguistic knowledge)**

- The ontology – the key module in FunGramKB – is presented as a hierarchical catalogue of all the concepts that a person has in mind when talking about everyday situations.
The cognicon stores procedural knowledge by means of conceptual macrostructures.

The onomasticon stores information about named entities and events.

A point worth mentioning is that while the lexical level is language-specific, the conceptual level is universal (Periñán & Arcas, 2010a, 2010b). Accordingly, the ontology feeds the different lexica, as pictured in figure 2.

3. The FunGramKB Lexicon

The FunGramKB lexicon presents two clear advantages:

a) In line with the LCM concern to account for all dimensions of meaning, FunGramKB lexical entries capture conceptual, semantic, syntactic and pragmatic information.

b) The amount of semantic information meets the needs of Artificial Intelligence systems.

Lexical entries in the FunGramKB lexicon are arranged alphabetically and include conceptual information. This is a major difference between FunGramKB and multilingual database such as SIMPLE and EuroWordnet, which describe the meaning of lexical units through associations with other lexical units, thus increasing redundancy through the knowledge base.

The incorporation of conceptual information into the FunGramKB lexical entries is explained by the heavy influence of the ontology on the lexical level to the extent that the lexicon is grounded on the ontology. The conceptual content of a lexical unit is expressed in terms of the concept to which the lexical unit is linked, the thematic frame, the meaning postulate and the meaning description. FunGramKB distinguishes three conceptual levels (Periñán & Arcas, 2004):

- Metaconcepts, which are preceded by symbol #, constitute the upper level in the FunGramKB ontology. FunGramKB posits forty-two metaconcepts distributed in three subontologies: #ENTITY for nouns (e.g. #COLLECTION), #EVENT for verbs (e.g. #MOTION, #POSSESSION) and #QUALITY for adjectives and some adverbs (e.g. #ABSTRACT, #TEMPORAL)

- Basic concepts, which are headed by symbol + (e.g. +BOOK_00, +DIRTY_00, +FORGET_00 etc), are used as defining units which enable the construction of meaning

---

3 For a full description of the lexical module we refer to Mairal and Periñán (2009a, 2009b).

4 A complete description of the conceptual level is provided in Mairal & Periñán (2009a) and Periñán & Mairal (2009).
postulates (cf. below) for basic concepts and terminals, as well as taking part as selectional preferences in thematic frames (cf. below): e.g. +COLD_00, +MONEY_00, +MOVE_00 etc.

- Terminals, which are preceded by symbol $ (e.g. $JUNE_00, $SENILE_00, $GLEAM_00), are those concepts which have no definitory potential to take part in meaning postulates.

Entries in the lexicon and concepts in the ontology are linked by means of the feature <concept> in such a way that a) lexical entries sharing the same headword are mapped to different concepts and b) lexical entries sharing the same meaning are mapped to the same concept.

On the other hand, as advanced above, in FunGramKB basic and terminal concepts are employed in the definition of lexical units through the thematic frame and the meaning postulate. Every event and quality is assigned one thematic frame, i.e. a conceptual schema which states the number and type of participants involved in the prototypical cognitive scenario described by the event or quality (Periñán & Arcas, 2007a: 267).

The information stated in the TF is integrated into the meaning postulate. A MP comprises a group of one or more logically connected predications (e₁, e₂... eₙ), which are conceptual constructs carrying the generic features of concepts (Periñán & Arcas, 2004: 39). The MP of a subordinate concept consists of a genus or definiens (i.e. the superordinate concept, a metaconcept or a basic concept in the FunGramKB ontology) and one or more distinctive features (or differentiae).

The meaning description of a lexical unit provides the natural language equivalent of the meaning postulate.

The following instances exemplify the thematic frame, the meaning postulate and the meaning definition of two basic concepts belonging to the metacognitive dimension #COMMUNICATION> #EVENT, and which some of the lexical units under analysis are linked to.

1. Basic concept +SING_00, to which lexical units like English *sing*, *yodel* or French *chanter, chantonner, fredonner* are linked.
(1) TF = (x1: +HUMAN_00 ^ +BIRD_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal

As we see, TFs also account for those selectional preferences typically involved in the cognitive situation being described (Jiménez & Pérez, 2010): +HUMAN_00, +BIRD_00 and +ANIMAL_00 for the cognitive scenario of ‘singing’. Therefore, TF (1) describes this prototypical scenario as involving three participants: (i) entity x1 (Theme), being typically a human or a bird, which in the COREL notation is expressed with the basic concepts +HUMAN_00 and +BIRD_00 connected with the exclusion logical connector “^”; (ii) entity x2 (Referent); (iii) entity x3 (Goal), typically identified with a human or an animal.

(2) MP = +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: +MUSIC_00)Manner)

The MP for (1) presents a semantic representation in which a human being or a bird (x1 = Theme) typically makes sounds with the voice (x2 = Referent) using music (Manner).

The genus of the MP is the superordinate +SAY_00, which is modified by one differentiae (f1), which expresses the manner (+MUSIC_00) of the action.

(3) Meaning description: to make musical sounds with the voice, usually a tune with words.

2. Basic concept +WHISPER_00, to which lexical units like English whisper or French chuchoter, murmurer, susurrer are linked.

(4) TF = (x1: +HUMAN_00)Theme (x2)Referent (x3: +HUMAN_00 ^ +ANIMAL_00)Goal

The MP of (4) provides a semantic representation in which a human being (x1 = Theme) typically transmits a message (x2 = Referent) to a human or an animal (x3 = Goal) (Location) without making noise (manner), which employing COREL is formalized as follows:

(5) MP = +(e1: +SAY_00 (x1)Theme (x2)Referent (x3)Goal (f1: (e2: +CREATE_00 (x1)Theme (x4: +NOISE_00)Referent))Manner)

(6) Meaning description: to speak softly, in a low voice.

We now explain the format of a lexical entry as represented in the FunGramKB lexical module. The meaning representation of lexical units in the FunGramKB lexicon does not only encode those aspects of the meaning of a word that are grammatically relevant but also semantic and
pragmatic properties. It is presented in terms of features or parameters. Table 1 contains the types of features being present in FunGramKB lexical entries for English and French\(^5\). 

\(^5\) The “en” and “fr” tags represent English and French languages respectively.
<table>
<thead>
<tr>
<th>Section</th>
<th>Noun</th>
<th>Adjective</th>
<th>Verb</th>
<th>Adverb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- Basic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2- Morphosyntax</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1- Graphical variant</td>
<td></td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2- Abbreviation</td>
<td></td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5- Category</td>
<td>en/fr</td>
<td>en/fr</td>
<td>en/fr</td>
<td>en/fr</td>
</tr>
<tr>
<td>2.6- Number</td>
<td>en/fr</td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.7- Gender</td>
<td>en/fr</td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.8- Countability</td>
<td>en/fr</td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.9- Degree</td>
<td>en/fr</td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.10- Adjectival position</td>
<td>en/fr</td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>constraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.12- Pronominalization</td>
<td>en/fr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3- LCM Core Grammar</td>
<td></td>
<td>en/fr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1- Aktionsart</td>
<td></td>
<td></td>
<td>en/fr</td>
<td></td>
</tr>
<tr>
<td>3.2- Lexical template</td>
<td></td>
<td></td>
<td>en/fr</td>
<td></td>
</tr>
<tr>
<td>3.3- Construction</td>
<td></td>
<td></td>
<td>en/fr</td>
<td></td>
</tr>
<tr>
<td>4- Miscellaneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Features in FunGramKB lexical entries.

Our focus will be on the core grammar module, which consists of elements of syntactically relevant semantic information. As shown in Figure 3, the LCM Core Grammar in the lexicon
contains those attributes whose values allow the system to build automatically the conceptual logical structures of lexical units after the application of the CLS Constructor Algorithm.\(^6\)

Lexical entries include the *Aktionsart* of the predicate and the argument constructions in which the predicate participates. Each predicate is assigned one or more *Aktionsarten* from Vendler’s (1967) catalogue of verb classes which is divided into states, activities, achievements, semelfactives, and accomplishments, together with their corresponding causatives.

Examples of each verb class and their formal representation (cf. Van Valin, 2005: 45) are provided in table 2:

<table>
<thead>
<tr>
<th>VERB CLASS</th>
<th>LOGICAL STRUCTURE</th>
<th>EXAMPLE</th>
<th>INSTANTIATION OF LS</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>\textit{do}'(x, [\textit{predicate'}(x) or (x,y)])</td>
<td>\textit{run}</td>
<td>\textit{do}'(x, [\textit{run'}(x)])</td>
</tr>
<tr>
<td>Activity</td>
<td>\textit{INGR predicate'}(x) or (x,y) or \textit{INGR do}'(x, [\textit{predicate'}(x) or (x,y)])</td>
<td>\textit{pop} (burst into tears)</td>
<td>\textit{INGR popped'}(x)</td>
</tr>
<tr>
<td>Achievement</td>
<td>\textit{SEML predicate'}(x) or (x,y) or \textit{SEML do}'(x, [\textit{predicate'}(x) or (x,y)])</td>
<td>\textit{glimpse, cough}</td>
<td>\textit{SEML see'}(x, y)</td>
</tr>
<tr>
<td>Semelfactive</td>
<td>\textit{BECOME do}'(x, [\textit{predicate'}(x) or (x,y)])</td>
<td>\textit{receive}</td>
<td>\textit{BECOME have'}(x, y)</td>
</tr>
<tr>
<td>Accomplishment</td>
<td>\textit{do}'(x, [\textit{predicate'}(x, (y))] &amp; \textit{BECOME predicate'}(x, y or y)</td>
<td>\textit{drink}</td>
<td>\textit{do}'(x, [\textit{drink'}(x, y)]) &amp; \textit{BECOME consumed'}(y)</td>
</tr>
<tr>
<td>Causative accompl.</td>
<td>$\alpha$ \textit{CAUSES} $\beta$ where $\alpha$, $\beta$ are LS of any type</td>
<td>$\textit{kill}$</td>
<td>$\textit{[do}'(x, \emptyset) \text{ CAUSE } \text{[BECOME <a href="y">dead'</a>]}]$</td>
</tr>
</tbody>
</table>

Table 2. Inventory of RRG logical structures.

State: *L’enfant est blessé.*
Causative state: *Ses remarques m’ennuient.*
Activity: *La voiture a glissé sur l’autoroute.*
Causative activity: *Elle a glissé la note sous la porte.*
Accomplishment: *Les draps ont séché.*
Causative accomplishment: *Le soleil a séché le linge.*
Achievement: *Le pneu a éclaté.*
Causative achievement: *On m’a crevé deux pneus sur ma voiture.*

---

\(^6\)CLS are lexical representations resulting from the interaction of TFs and MPs with information stored in lexical entries. The CLS Constructor builds the CLS automatically from the information stored in the LCM Core Grammar together with conceptual knowledge stored in the FunGramKB ontology.
Semelfactive: Il toussait sans arrêt.
Causative semelfactive: Ne sonne pas.
Active accomplishment: Nous avons marché jusqu’à la gare.
Causative active accomplishment: J’ai promené le chien jusqu’au parc.

FunGramKB also incorporates into its core-grammar level of description an inventory of argument structure constructions of the kind proposed by Levin (1993) and Goldberg (1995, 2006). Such incorporation is based on the assumption that the argument structure of a predicate proves insufficient to explain the occurrence of some constituents and that constructions play an active role in determining the type of syntactic configuration in which a predicate is embedded (Mairal & Ruiz de Mendoza, 2009: 157). In this light, constructions contribute arguments to yield the final semantic interpretation of a predicate.
In the next section we discuss the lexicalisation of the metaconcept #COMMUNICATION in English and French.
4. Lexical representation of the metaconcept #COMMUNICATION in the English and French FG lexica

Speech act verbs belong to the SAY-type. SAY is a type of indefinable suggested by Faber and Mairal (1999) within the lexical domain of speech that serves to define the rest of the verbs in this class. In FunGramKB speech verbs belong to the metaconcept #COMMUNICATION and share the genus of “say”.

Figure 4 presents the hierarchical structure of the metaconceptual dimension #COMMUNICATION.

![Figure 4. The metaconcept #COMMUNICATION in the FunGramKB ontology.](image-url)
The metaconcept #COMMUNICATION agglutinates twenty-two basic concepts in the FunGramKB ontology. Basic concepts, which are headed by symbol +, are used 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.

There are five terminal concepts (i.e. $CONGRATULATE, $FLATTER, $GOSSIP, $PROPHESY and $SWEAR) located under the basic concept +SAY_00. Terminals, which are preceded by symbol $, are those concepts which have no definitory potential to take part in meaning postulates.

The following meaning definitions provided in the conceptual module show that the semantic particulars of the object permeate the lexical encoding of SAY-type predicates:

- **congratulate / féliciter**: to say someone that you are happy because they have achieved something or because something nice has happened to them.
- **request / demander**: to express the need or desire for something.
- **accuse / accuser**: to say that someone did something wrong or is responsible for something bad happening.
- **greet / saluer**: to say hello to someone.
- **gossip / potiner**: to say about people’s behavior and private lives remarks that are unkind or untrue.

The comparative analysis of the lexicalization of the metaconcept #COMMUNICATION reveals significant differences between French and English. This is in accord with the idea that while the lexical level is language-specific, the conceptual level is universal (cf. above). Accordingly, the lexical module codes particular idiosyncrasies such as those discussed below.

A quantitative analysis (cf. table 3) shows two facts: (i) A similar proportion of predicates are linked to the same concept in English and French; (ii) Terminal concepts are less richly lexicalized in both languages. This fact points to the lexicon-ontology interface.

---

7 In FunGramKB basic and terminal concepts are provided with language-independent conceptual properties such as a thematic frame and a meaning postulate. A thematic frame is a conceptual schema which states the number and type of participants involved in the prototypical cognitive scenario described by the event or quality (Periñán & Arcas, 2007a: 267). A meaning postulate comprises a group of one or more logically connected predications \( (e_1, e_2, \ldots, e_n) \), which are conceptual constructs carrying the generic features of concepts (Periñán & Arcas, 2004: 39).
<table>
<thead>
<tr>
<th>Concept</th>
<th>English lexical units</th>
<th>French lexical units</th>
</tr>
</thead>
<tbody>
<tr>
<td>+SAY_00</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>+CONGRATULATE_00</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>+FLATTER_00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>+GOSSIP_00</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>+PROPHESEY_00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+SWEAR_00</td>
<td>1</td>
<td>_</td>
</tr>
<tr>
<td>+ADVISE_00</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>+AGREE_00</td>
<td>11</td>
<td>6</td>
</tr>
<tr>
<td>+ANSWER_00</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>+APPROVE_00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+BLAME_00</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>+BLESS_00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+COMPLAIN_00</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>+DESCRIBE_00</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>+EXPLAIN_00</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>+FORBID_00</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>+GREET_00</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>+LIE_00</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>+PERMIT_00</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>+PRAY_00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+REFUSE_0</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>+REPEAT_00</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>+REQUEST_00</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td>+SHOUT_00</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>+SING_00</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>+THANK_00</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>+WHISPER_00</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3. Lexicalization of the metaconcept #COMMUNICATION in the English and French FG lexica
The lexicalization of the metaconcept #COMMUNICATION reveals quantitative differences concerning a few concepts. Such differences can be explained by orthographic variations, pragmatic features or semantic parameters relevant to meaning description. Orthographic variations account for the higher proportion of English verbs codified by the basic concept +SAY_00 (enquire / inquire, verbalize / verbalise).

Other differences in the lexicalization of the metaconcept +SAY_00 are accountable for on the basis of pragmatic features. Thus, the value <formal> is encoded in a few English verbs subsumed under the basic concepts +ANSWER_00 (rejoin, retort, riposte), +BLAME_00 (reproach, charge, fault, impute, tax), +EXPLAIN_00 (elucidate, explicate, expound), +FORBID_00 (debar, interdict, prohibit), +GREET_00 (hail, salute), +REPEAT_00 (iterate, recapitulate, rehearse, reiterate, restate) and +AGREE_00 (accede, accord, acquiesce, assent, concur, consent). In contrast, none of the French lexical items linked to these concepts are diastratically marked.

The encoding of semantic parameters and the semantic specification of the object argument contribute to a richer lexicalization of some concepts in a language. It is the case of the concept +REQUEST_00 in French, as the following dictionary definitions show:

- *dicter*: demander à quelqu’un en secret et à l’avance de faire quelque chose (‘ask for something secretly beforehand’)
- *somer*: demander avec autorité de faire quelque chose (‘ask authoritatively for something’)
- *prier*: demander à quelqu’un faire quelque chose avec humilité ou politesse (‘ask for something humbly or politely’)
- *charger*: demander à quelqu’un de remplir une mission/fonction (‘ask someone to do a task’)

The manner parameter is a distinguishing feature in a set of French speech predicates for which there is no translational equivalent in English. Consider the following meaning definitions:

- *chantonner* (+SING_00): chanter à mi-voix.
- *vanter* (+SFLATTER_00): parler très favorablement de quelqu’un/quelque chose en exagérant ses mérites/qualités publiquement.
- *pester* (+COMPLAIN_00): manifester son mécontentement/sa colère par des paroles hargneuses et violentes.
- *hèler* (+SHOUT_00): appeler quelqu’un en se servant des mains comme un porte-voix.
We notice a lexical gap between English and French with regard to the terminal $SWEAR_00$. This concept is not lexicalized in French since there is no French lexical item whose meaning is ‘use rude and offensive words’. This meaning is expressed by means of the hypernym “say” and an object which is prototypically rude words (des jurons).

We will now focus on the most relevant features in the FunGramKB lexicon for the purposes of this study, i.e. pronominalization, Aktionsart, syntactic constructions and pragmatic features.

4.1. Pronominalization

Pronominalization serves as a differentiating feature in the domain under analysis. It is true that the bulk of the verbs are never cliticised in either language (e.g. gossip, agree, complain, pray-prier, request, shout-crier, beg-supplier) and several verbs are optionally cliticised in both languages, such as ask/demander, express/exprimer, repeat/répéter, congratulate/féliciter, let/laisser. Consider the following examples:

(1) I can express myself better in writing than verbally.
(2) Le Premier Ministre s’exprimait hier à Strasbourg.
(3) He congratulated himself for doing a good job.
(4) Le patron de la chaîne s’est félicité de ses résultats historiques.

 Nonetheless, the analysis of the feature pronominalization in SAY-type verbs yields two differences:

(i) More English verbs are optionally cliticised, while the corresponding French verbs are never cliticized. It is the case of the verbs linked to the concept +DESCRIBE and others (i.e. reproach, allow). Let us see a few examples:

(5) The accused gunman described himself as a terrorist.
(6) Le soldat a décrit les mauvais traitements qu’il avait subis.
(7) He characterizes himself as highly moral and tolerant.
(8) Qu’est-ce qui caractérise la France?
(9) Palin depicts herself as one of the Tucson victims.
(10) Cette affaire a dépeint la relation entre les grandes fortunes et les politiques.
(11) How could I allow myself to get manipulated?
(12) Sa foi lui permet de tenir.

(ii) The value <always cliticized>, which is only codified in French, differentiates the lexical items *se plaindre* and *s'enquérir, s'exclamer* from their English counterparts *complain* and *enquire/inquire* and *exclaim*.

4.2. Aktionsart

As regards the verb’s *Aktionsart*, the predicates pertaining to the domain of SPEECH denote either an activity or a causative accomplishment. Activities are actions defined by the semantic properties [+dynamic] and [-telic]. The dynamic means that the participant in the action does or causes someone to do something. The telic specifies the culmination of the action. Active accomplishments are the telic uses of activities.

Table 4 presents the distribution of the *Aktionsarten* encoded by speech predicates according to the concept the verbs are linked to:
Table 4. *Aktionsarten* encoded by speech predicates.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Causative accomplishment</th>
</tr>
</thead>
<tbody>
<tr>
<td>+SAY_00</td>
<td>+ADVISE_00</td>
</tr>
<tr>
<td>SCONGRATULATE_00</td>
<td>+FORBID_00</td>
</tr>
<tr>
<td>SFLATTER_00</td>
<td>+FORBID_00</td>
</tr>
<tr>
<td>SGOSSPIP_00</td>
<td>+REQUEST_00</td>
</tr>
<tr>
<td>SPROPHESY_00</td>
<td></td>
</tr>
<tr>
<td>SSWEAR_00</td>
<td></td>
</tr>
<tr>
<td>+AGREE_00</td>
<td></td>
</tr>
<tr>
<td>+ANSWER_00</td>
<td></td>
</tr>
<tr>
<td>+APPROVE_00</td>
<td></td>
</tr>
<tr>
<td>+BLAME_00</td>
<td></td>
</tr>
<tr>
<td>+BLESS_00</td>
<td></td>
</tr>
<tr>
<td>+COMPLAIN_00</td>
<td></td>
</tr>
<tr>
<td>+DESCRIBE_00</td>
<td></td>
</tr>
<tr>
<td>+EXPLAIN_00</td>
<td></td>
</tr>
<tr>
<td>+GREET_00</td>
<td></td>
</tr>
<tr>
<td>+LIE_00</td>
<td></td>
</tr>
<tr>
<td>+PRAY_00</td>
<td></td>
</tr>
<tr>
<td>+REFUSE_00</td>
<td></td>
</tr>
<tr>
<td>+REPEAT_00</td>
<td></td>
</tr>
<tr>
<td>+SHOUT_00</td>
<td></td>
</tr>
<tr>
<td>+SING_00</td>
<td></td>
</tr>
<tr>
<td>+THANK_00</td>
<td></td>
</tr>
<tr>
<td>+WHISPER_00</td>
<td></td>
</tr>
</tbody>
</table>

Both *Aktionsarten* are associated with three participants (or arguments): the Theme, the Referent and the Goal of the action. Syntactically, the goal argument of an activity is realized by a noun phrase, whereas the goal of a causative accomplishment is expressed as an infinitive. Let us consider the following examples:

- Activities:
  
  (13) He didn’t answer all the questions.
(14) The teacher explained the vocabulary in the text.
(15) What are they complaining about?
(16) Je lui ai reproché de ne pas travailler assez.
(17) La petite criait à pleins poumons.
(18) Je ne sais pas pourquoi il a menti.

The goal is an optional argument of a set of predicates (e.g. say/dire, shout/crier, approve/approuver, answer/répondre, repeat/répéter). We illustrate this point through examples:

(19) The boy shouted in pain.
(20) I don’t approve the project.
(21) Il répète toujours les mêmes histoires.

The predicates belonging to the conceptual dimensions +ADVISE_00, +FORBID_00, +PERMIT_00 and +REQUEST_00 designate causative accomplishments. The Goal argument is an obligatory argument, as illustrated in the following examples:

(22) He forbade his son to smoke in the house.
(23) She didn’t allow me to go.
(24) Je l’ai prié de bien y réfêchir.
(25) Je vous conseille d’aller voir un spécialiste.

The incorporation of the predicate’s Aktionsart into the lexical entries is particularly relevant for French since it accounts for semantic distinctions such as those exemplified below that are not explained in traditional grammatical descriptions (e.g. Grevisse, 2007, 2009):

(26) Je lui ai dit comment j’avais réussi. (Activity)
(27) Je lui ai dit de venir tout de suite. (Causative accomplishment)
(28) Nous n’avons pas demandé le prix. (Activity)
(29) Ils nous ont demandé de les aider. (Causative accomplishment)

In traditional grammatical accounts of French dire and demander would simply be ditransitive verbs whose direct object may be realized as a noun phrase (28), a nominal clause (26) or an infinitive (27) (29), the description of a verb’s syntactic behaviour being based on the notion of transitivity and the syntactic function of the arguments (direct object, indirect object, etc.). In contrast, the lexical entry for dire in FunGramKB states the two Aktionsarten encoded by the

8 French distinguishes between transitive (e.g. manger), ditransitive (e.g. envoyer) and intransitive (e.g. obéir) verbs.
predicate\textsuperscript{9} and the semantic function of the arguments, i.e. Theme, Referent and Goal. A Theme, prototypically human, transmits a message, i.e. set of propositions (Referent) to another entity (Goal).

4.3. Syntactic constructions

The comparative analysis of English and French SAY-type predicates yields important differences concerning the argument constructions in which lexical predicates take part.

A higher proportion of English speech predicates take part in argument constructions. As a matter of fact, several constructions are only sensitive to English verbs. It is the case of the dative alternation\textsuperscript{10}, the together- reciprocal alternation, the reaction object alternation, the as-alternation and the caused motion construction. The following instances exemplify them:

(30) I recommended the camera to all my friends. / I recommended all my friends the camera.
(31) I never spoke with him. / We never spoke together.
(32) They groaned their envy.
(33) She whispered thanks and left.
(34) I strongly recommend Wilt as a good book.
(35) He shouted her into the room.
(36) The principal called me into his office.
(37) My friends talked me into going on a diet.

The dative alternation is not feasible in French since the indirect object is syntactically introduced by the preposition à, e.g. J'ai recommandé l'appareil à tous mes amis; Il a expliqué le poème aux élèves.

Similarly, the together-reciprocal alternation, as illustrated by (31), is impossible in French since the item ensemble ‘together’ is used with verbs denoting physical action or motion (Nous sommes allés ensemble / *Nous avons parlé ensemble).

Several French and English verbs take part in other constructions such as the reciprocal alternation and the cognate object construction.

\textsuperscript{9} Interestingly enough, both \textit{Aktionarten} have the same number of variables.

\textsuperscript{10} Syntactic alternations are regarded as formal variants in the expression of arguments that verbs may participate in (Levin, 1993: 2).
• Reciprocal alternation (intransitive)
  \[ S/\text{NP1} + v + \text{OP-from, into, to, with (NP2)} \rightarrow S/(\text{NP} & \text{NP2}) + v \]
The predicates *speak-parler* and *agree-être d’accord*, which have a prepositional object, can occur without it. This being the case, the subject must be a collective noun phrase. Subjects are typically animate and volitional (Levin, 1993: 59).

Consider the following examples:
(38) The Republicans didn’t agree with the Democrats about the health reform.
(39) The Republicans and the Democrats didn’t agree about the health reform.
(40) Jean a parlé à Marie des vacances à la montagne.
(41) Jean et Marie ont parlé des vacances à la montagne.

• Cognate object construction
  \[ S/\text{NP1} + v \rightarrow S/\text{NP1} + v + O/\text{NP2} \]
The intransitive verbs *pray* and *sing-chanter* conveying a non-verbal expression take as their object a noun that is zero-related to the verb and so called *cognate object*.

(42) Have you ever prayed a pray like Solomon prayed that day?
(43) Sing us a song then. Just to pass the time.
(44) Il nous a chanté une belle chanson.

4.4. Pragmatic features
As regards pragmatic information, the feature <dialect> is only pertinent to the meaning description of four English lexical units employed in the American variety (e.g. verbalise, characterize, kvetch and holler).

The feature <style> is codified by a group of English and French lexical units. On the one hand, twenty-nine English predicates share the value <formal>: utter, verbalise, praise, flatter, counsel, concur, assent, accede, accord, acquiesce, consent, rejoin, retort, riposte, reproach, fault, tax, impute, explicate, expound, iterate, recapitulate, reiterate, elucidate, debar, permit, entreat, solicit, request.

In contrast, in the French language this value is only relevant to the meaning description of one lexical unit (i.e. complimenter).
On the other hand, a considerable number of lexical units linked to the concept +COMPLAIN_00 share the value <informal>. The set of predicates includes English grizzle, holler, moan, bitch, beef, belly-ache, kvetch and grouse, and French bougonner, clabauder, geindre, maronner, ronchonner, rouscailler and rouspéter. There is also a slang predicate, gueuler, linked to the concept +SHOUT_00.

1. Conclusion
This paper has offered an overview of FunGramKB, a knowledge base designed as a part of the LCM which integrates rich semantic and syntactic information and will eventually serve for the development of NLP applications. The discussion of the anatomy of the FunGramKB lexicon through the lexical entries for English and French speech act predicates shows how fully-fledged and fine-grained the semantic representation of lexical units in FunGramKB is since lexical entries specify the morphosyntactic, pragmatic and collocational information of predicates. We sustain that the description of the various kinds of data that form part of the FunGramKB lexical entries shows that the lexical module is a major achievement of the LCM. On the one hand, the lexical entries have been enriched with relevant conceptual, semantic, syntactic and pragmatic information. On the other hand, lexical representation has been provided with a formal metalanguage.
2. Bibliographic references


—, 2009b: “Role and Reference Grammar and ontological engineering” in *Volumen Homenaje a Enrique Alcaraz*, Alicante: Universidad de Alicante,


