MULTILINGUALISM AND CONCEPTUAL MODELLING

Ricardo Mairal-Usón and Carlos Periñán-Pascual

Universidad Nacional de Educación a Distancia, Universitat Politècnica de València

rmairal *at* flog uned es, jopepas3 *at* upv es

Abstract

One of the leading motivations behind the multilingual semantic web is to make resources accessible digitally in an online global multilingual context. Consequently, it is fundamental for knowledge bases to find a way to manage multilingualism and thus be equipped with those procedures for its conceptual modelling. In this context, the goal of this paper is to discuss how common-sense knowledge and cultural knowledge are modelled in a multilingual framework. More particularly, multilingualism and conceptual modelling are dealt with from the perspective of FunGramKB, a lexico-conceptual knowledge base for natural language understanding. This project argues for a clear division between the lexical and the conceptual dimensions of knowledge. Moreover, the conceptual layer is organized into three modules, which result from a strong commitment towards capturing semantic knowledge (Ontology), procedural

knowledge (Cognicon) and episodic knowledge (Onomasticon). Cultural mismatches are discussed and formally represented at the three conceptual levels of FunGramKB.

Keywords: lexicon, ontology, culture, multilingualism, conceptual modelling, FunGramKB.

Contents

1. Introduction

Ontological engineering projects focus on the development of knowledge bases (KB) that represent and organize world knowledge. Ontologies take the form of concept taxonomies where universal distinctions and features are applicable to and shared by different languages, although it is clear that not all languages conceptualize reality in the same way and therefore cultural distinctions play an essential part in ontology localization. Hence, Cimiano et al. (2010, pp. 127) defined ontology localization "as the process of adopting a given ontology to the needs of a certain community, which can be

characterized by a common language, a common culture or a certain geopolitical environment". Moreover, in the context of the semantic web, defined by Berners-Lee et al. (2001) as "an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation", the need to make ontological resources multilingual has become an urgent priority in the research agenda. As a matter of fact, while the traditional web is primarily language-specific, one of the leading motivations behind the multilingual semantic web is to make resources accessible regardless of the language or the culture of each particular website (Gracia et al., 2012). Consequently, it is fundamental for KB projects to find a way to manage multilingualism and thus be equipped with those procedures for its conceptual modelling.

As Montiel-Ponsoda (2011, chapter 8) described, ontology localization varies (i) by virtue of the nature of the domain involved, i.e. whether it is an *internationally-standardized* or a *culturally-influenced domain*; (ii) by the function of the ontology, that is, whether it has an instrumental ("the goal of the target ontology is to have the same function in the target community as the original ontology in the source community") or a documentary function ("the purpose is to support the use of the original ontology by members of another linguistic community"); (iii) or by its categorization as interoperable ("how interoperable the new ontology needs to be to the original one") vs. independent. As reported in the literature—cf. Espinoza et al. (2009), Cimiano et al. (2010), or Montiel-Ponsada (2011), among many others—there are three models for representing multilingualism in ontological projects, each depending to a large extent on the type of domain and function involved:

(a)   In RDF and OWL resources,[1] the inclusion of tags containing the translation equivalents in different languages is one of the most popular procedures in

---

[1] RDF stands for Resource Description Framework and OWL is the acronym for Ontology Web Language. These semantic mark-up languages have become two of the standards for linking objects in the multilingual semantic web, where every factual statement can be expressed by means of a triplet consisting of a subject-predicate-object relationship. Thus, the triplet <Shakespeare, wrote, Hamlet> can be represented in RDF as follows (Arms, 2000, pp. 198):

ontological engineering since it guarantees interoperability. However, it is rather limited when it comes to dealing with cases of conceptual mismatches, since these RDF and OWL labels function fairly well for those cases where there is an exact equivalence in the two languages involved. Thus, this procedure could be used successfully for internationally-standardized domains with an instrumental function although its efficiency is more limited for culturally-influenced domains, which require richer and more fine-grained lexical representations.

(b)   Each language can have its own conceptualization model and a mapping mechanism is established in order to account for the different conceptualizations. This has the advantage that it is feasible to provide a fine-grained description of those cultural distinctions in each of the languages involved, although it becomes a complex task to establish the mapping between the different conceptualizations. As a representative of this approach, EuroWordNet[2] (Alonge et al., 1998) can formulate a wordnet, or lexical taxonomy, for each of the languages involved and uses an Inter-Lingual-Index to connect the different conceptualizations as represented in each wordnet. This model can be a good candidate for culturally-influenced domains with an instrumental function, while it is not even relevant for internationally-standardized domains because there is no sense in creating and duplicating wordnets in each language given that the languages involved share the same conceptualization.

(c)   The ontology model can be linked to an external linguistic model, which has the capacity to include a rich amount of linguistic information. Hence, there is a clear separation of the conceptual and the linguistic layers, each with different

---

```
<rdf:rdf>
        <rdf:description rdf:about =  "http://hamlet.org/" >
                <dc:creator>Shakespeare</dc:creator>
        </rdf:description>
</rdf:rdf>
```
[2] http://www.illc.uva.nl/EuroWordNet/

functions: while most of the localization activity falls primarily on the lexicon, there is just one conceptual layer that is claimed to be language-independent and therefore shared by the languages involved. The fact of having an external lexicon permits the inclusion of as much linguistic information as required by each of the languages. Thus, this model is a good candidate for dealing with culturally-influenced domains with either an instrumental or a documentary function since the relevant cultural information can be encoded in the external model to which the ontology is linked. There are several projects illustrative of this line of research, e.g. OntoTerm (Moreno Ortiz, 2000), LexInfo (Buitelaar et al., 2009), LIR (Montiel-Ponsoda et al., 2008) and Lemmon (McCrae et al., 2010).

We believe that the third approach has a number of advantages over the other two since the great load of cultural distinctions is attributed to the external lexical model without having to modify the ontology. Therefore, this clear demarcation of the lexicon and the ontology provides for a flexible and interoperable framework. This move has also the added advantage that it is no longer necessary to build a new ontology every time a new language is incorporated, which is a time-consuming task. Moreover, these projects follow standardization models, which guarantee interoperability and reuse of other resources, if needed, and aligns especially well with the Linked Open Data Cloud initiative. However, as shown in section 3, in a conceptualist model like the one we argue for in this paper, some cultural mismatches can only be resolved at the cognitive level.

With this in mind, the aim of this paper is to discuss how multilingualism is dealt with in FunGramKB (Mairal-Usón and Periñán-Pascual, 2009; Periñán-Pascual, 2013; Periñán-Pascual and Arcas-Túnez, 2004, 2007, 2010; Periñán-Pascual and Mairal-Usón, 2009),[3] a multipurpose lexico-conceptual knowledge base to be implemented in natural language processing systems, and more particularly for natural language understanding.

---

[3] http://www.fungramkb.com

Hence, the organization of this paper is as follows. In section 2, we offer an overview of FunGramKB with the aim of contextualizing this project within the larger framework of ontology localization as described in this introductory section. Indeed, one of the distinctive features of our proposal is that the conceptual layer goes beyond the ontology. In this regard, section 3 describes the different types of knowledge schemata that are part of this knowledge engineering project. Finally, sections 4 and 5 deal with the way that cultural mismatches in common-sense and cultural knowledge respectively are conceptually modelled in FunGramKB.

2. FunGramKB

As in the case of the third approach to multilingualism representation, FunGramKB postulates a clear-cut difference between the ontology and the lexicon, each pertaining to its own level of representation, i.e. the cognitive and the linguistic levels respectively (cf. Figure 1).

Figure 1: The FunGramKB architecture (Periñán-Pascual and Arcas-Túnez, 2011, pp. 3).
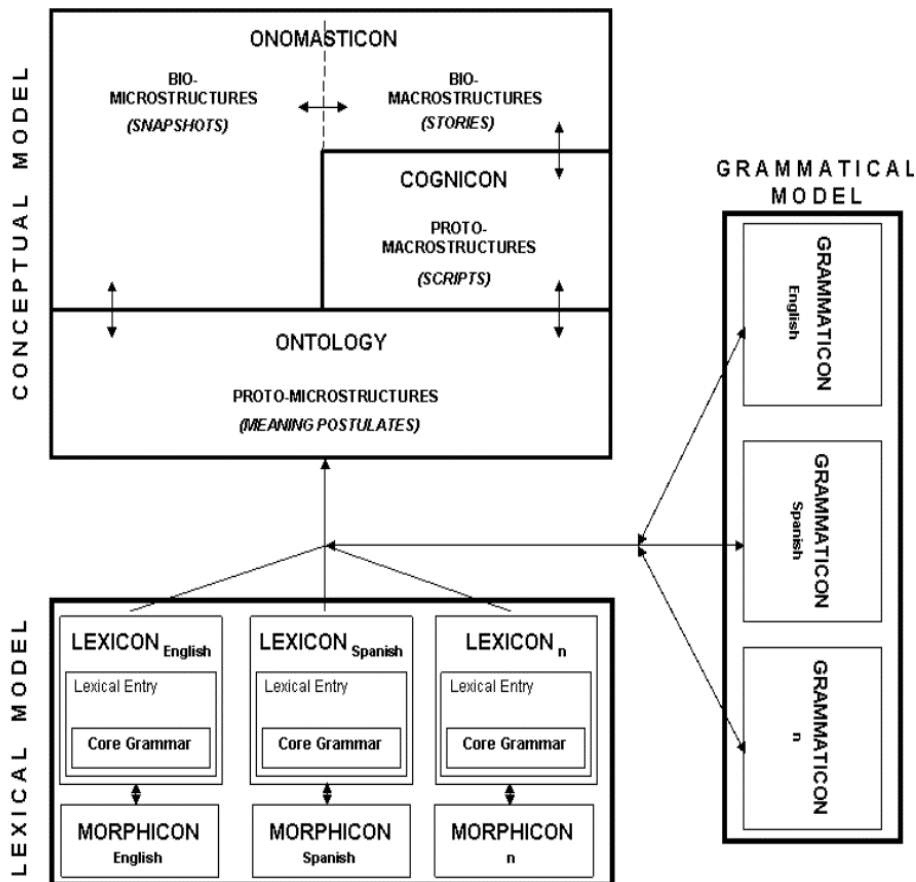
However, one of the differentiating features with respect to the external models of lexicon is that the cognitive level in FunGramKB goes beyond the construction of an ontology and offers two additional modules. The three conceptual modules are described as follows:

(i)   The Ontology stores semantic knowledge and has the format of a hierarchical catalogue of concepts whose properties are expressed in terms of meaning postulates. The Ontology consists of a general-purpose module (i.e. Core Ontology) and several domain-specific terminological modules (i.e. Satellite Ontologies). Satellite ontologies are specially suited to representing internationally-standardized domains.

(ii)   The Cognicon stores procedural knowledge by means of scripts, i.e. schemata in which a sequence of stereotypical actions is organised on the basis of temporal continuity, e.g. how to celebrate a wedding in Spain, how to sue a tenant in Germany, etc.

(iii)   The Onomasticon stores information about instances of entities and events, such as Beethoven or 25/12. This module stores two different types of schemata (i.e. snapshots and stories), since instances can be portrayed synchronically or diachronically. This distinction is particularly relevant to dealing with the documentary function of ontology localization, since it is possible to extend the ontology with culturally-specific information, as shown in the discussion of the lexical units referring to the head of government and the head of state (cf. section 5).

The representation of the internal properties of each conceptual unit in the three modules is done using COREL (Conceptual Representation Language), a machine-tractable metalanguage which permits a fine-grained semantic description of conceptual units (Periñán-Pascual and Mairal-Usón, 2010). This is in fact another differentiating feature with respect to external models, since RDF and OWL triplets do not allow such semantic depth as COREL. For example, triplets cannot easily represent the meaning of *forgive*, i.e. you stop being angry with someone you blamed, where phenomena such as aspectuality and temporality should be stated.

On the other hand, the linguistic level includes three different language-dependent modules (i.e. Lexicon, Grammaticon and Morphicon)[4], although only the Lexicon is relevant for the scope of this paper. The FunGramKB Lexicon includes the following types of information (cf. Mairal-Usón and Periñán-Pascual, 2009):

    a. Basic: headword, index, and language.

    b. Morphosyntax: graphical variants, abbreviation, phrase constituents, category, number, gender, countability, degree, adjectival position, verb paradigm and constraints, and pronominalization.

    c. Core grammar: Aktionsart, lexical template and constructions[5].

    d. Miscellaneous: dialect, style, domain, example and translation[6].

While there is one Lexicon for each of the languages involved, that is, one for Spanish, another one for English, German etc., there is only one common Ontology, Cognicon and Onomasticon, which are shared by the different languages involved.

FunGramKB is able to account for both internationally-standardized and culturally-influenced domains with an instrumental function, although a further step is needed to work on interoperability and represent the data according to the Linked Open Data standards. Hence, FunGramKB is an instance of a conceptually-oriented external model, where lexical entries are linked to conceptual units in the Ontology. Although the activity of localization primarily falls on the Lexicon, some conceptual discrepancies need to be taken care of at other modules in the cognitive level. This is precisely the central concern of the following sections, an aspect which is usually ignored by knowledge engineers.

---

[4] See Mairal-Usón and Periñán-Pascual (2015) and Van Valin and Mairal-Usón (2014) for a discussion of the format and role of the FunGramKB Grammaticon.

[5] The core grammar section follows Role and Reference Grammar (Van Valin, 2005), a functionally-oriented model of linguistic description.

[6] The feature *dialect* clearly serves to capture those cultural distinctions that occur within different cultures that speak the same language, e.g. those variants between European Spanish and Latin American Spanish, not to mention those different instances within different regions in Spain.

3. A typology of cognitive knowledge

We claim that cognitive knowledge can be modelled according to two parameters: prototypicality and temporality (cf. Periñán-Pascual, 2012). On the one hand, conceptual representations can store prototypical knowledge (i.e. proto-structures) or can serve to describe a particular instance of a given entity or event (i.e. bio-structures). For example, the meaning of *song* can be represented in terms of a proto-structure, i.e. it refers to our common-sense knowledge about this particular instance. However, if we were asked to represent the meaning of the song *Imagine*, this would require a bio-structure, that is, a construct invoking our cultural knowledge. This parameter of prototypicality extends over a continuum that includes different types of knowledge: at one end, common-sense knowledge, while, at the other, personal knowledge. Figure 2 illustrates the different types of knowledge involved along the scale of prototypicality.

Figure 2: Types of knowledge according to prototypicality (Periñán-Pascual, 2012, pp. 188).
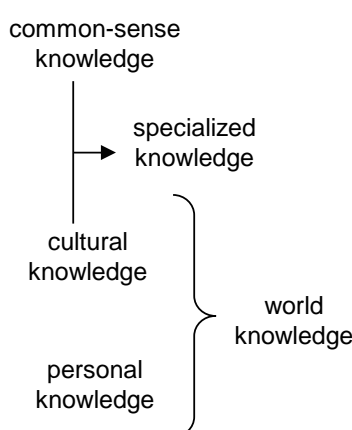
It is commonly agreed that common-sense knowledge has been associated with "the knowledge that every person assumes his neighbors also possess" (Panton et al., 2006, pp. 1). It refers to those pre-scientific beliefs that guide our everyday cognitive activities (Smith, 1995), as shown in (1).

(1)   a. Cats and dogs are mammals.

        b. A table is a piece of furniture.

Specialized knowledge consists of a set of beliefs that form part of scholarly, scientific or technical domains (e.g. medicine, chemistry, law, economics, linguistics, etc.), which are accepted to be true by the competent members of an expert community, as can be seen in (2).

(2)   a. A batholith is a large geological formation, made of granite, quartz monzonite, or diorite.

        b. The operating income margin is the financial ratio calculated as a company's operating profit divided by sales, also referred to as the EBIT or EBITDA margin (Fuertes et al., 2010, pp. 661).

Cultural knowledge refers to factual information about the past, present or future model of the world, i.e. it is the encyclopedic knowledge every educated speaker possesses. This knowledge can be shared to a higher or lesser degree, depending on the individuals' education, as shown in (3).

(3)   a. Vienna is the capital of Austria.

        b. Mozart composed the opera *Die Zauberflöte*.

Finally, personal knowledge refers to that set of beliefs and assumptions that are accepted to be true by the individual who holds them, and consequently it is not shared by other people unless they are told about it (Van Dijk, 2003). The examples in (4) illustrate this type of knowledge.

(4)   a. My friend Paco is a bit of a hypochondriac.

        b. People in our group enjoy eating fish.

On the other hand, this classification can be further enriched by looking at the role of situational context in knowledge representation. In this regard, Barsalou (1991) affirmed that cognitive processing is subject to two types of properties: (i) context-independent properties, which are activated in all background situations, and (ii) context-dependent properties, which are activated just by the relevant situations in which a given conceptual category is encoded. This twofold distinction coincides to a certain extent with our "microstructures" and "macrostructures" respectively. Moreover, he devised a model of what he called "situated conceptualization", whereby people conceptualize the same category in different ways depending on the background situation. For example, my common-sense knowledge about *chair* implies features such as "furniture", "to sit on" or "with back support", which are applicable to every context. Accordingly, the conceptual representation is not linked to a given situation but instead the representation of this conceptual entity is a product of abstraction over different background contexts. However, the attributes of this entity might acquire different values if it is perceived

within a given contextual setting. For example, the *chair* in my university office also has the following set of context-dependent parameters: grey leather material, two black arms, three wheels, etc.

Hence, we maintain that knowledge can be presented atemporally (i.e. microstructures) or in a temporal framework (i.e. macrostructures). For example, the biography of *John Lennon* requires a macrostructure, but a microstructure is sufficient to describe the profession of a pop singer. In this respect, the availability of a spatio-temporal framework causes macrostructures to be developed from a situated approach to conceptualization. In sum, prototypicality and temporality are the focal axes which provide the following schemata, as illustrated in Figure 3:

a)   Meaning postulates—i.e. proto-microstructures, or atemporal prototypical knowledge schemata—include our common-sense knowledge about a given conceptual entity without any reference to potential contextual features. For example, the entity *cat* has the following properties: animal, feline, has whiskers, four legs, mews, etc.

b)   Snapshots—i.e. bio-microstructures, or atemporal non-prototypical knowledge schemata—include one's beliefs and assumptions about the knowledge of a particular entity. For example, my cat *Rufus* has the following properties: white, brown eyes, friendly and kind, wears a sweater in winter, etc.

c)   Scripts—i.e. proto-macrostructures, or temporal prototypical knowledge schemata—refer to a type of situated generalized knowledge which is constructed from those past and present experiences stored in our minds. Future experiences are then interpreted on the basis of the patterns constructed from past experiences, and so scripts facilitate the inference of information from our perception of the world and thus serve to alleviate the cognitive overload. For example, when an American English speaker says that *Laura usually goes to the hospital early in the morning*, we may infer that Laura is a doctor or a nurse according to our standardized model of the world. However, if we later hear that Laura needs to take her medicine before going to the hospital, then we will infer that she is a patient under medical treatment.

d)   Stories—i.e. bio-macrostructures, or temporal non-prototypical knowledge schemata—contain our knowledge about a particular entity when perceived in a time interval. In other words, a story is the description of one named entity as if it were a biography, e.g. when and where Mozart was born, where he studied music, when he composed his first opera, etc.[7]

Figure 3: Extended typology of knowledge schemata (Periñán-Pascual, 2012, pp. 188).
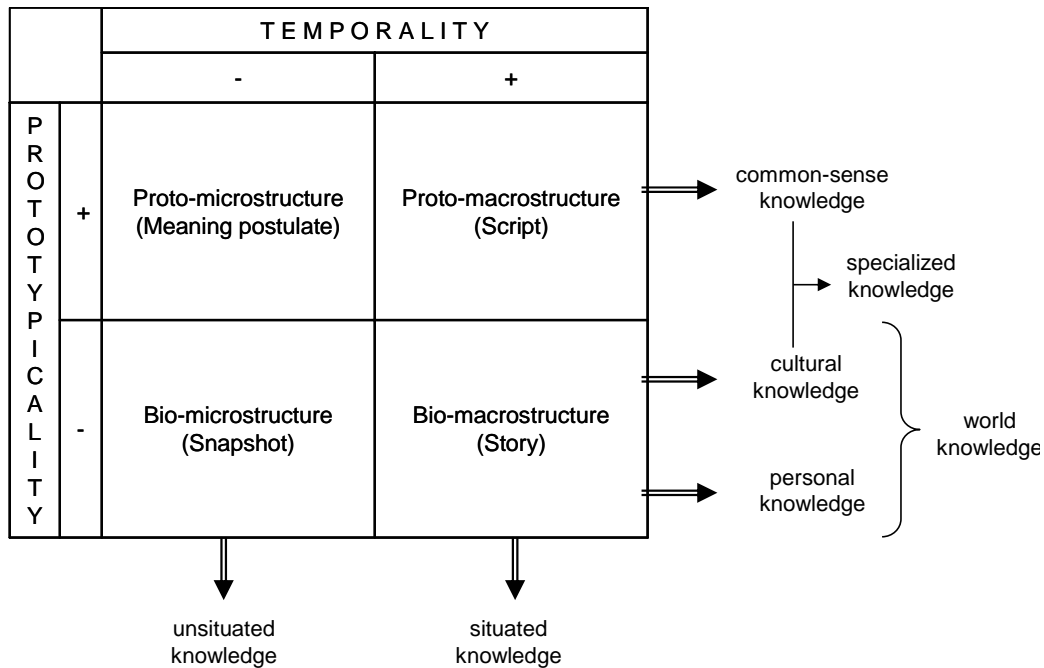
Drawing on this classification, the following section has a twofold objective: (i) we discuss the way cultural distinctions are present in this typology, and (ii) how these distinctions are actually represented in COREL. It can be easily noted that, unlike specialized and personal knowledge, the analysis of common-sense and cultural knowledge raises more fruitful discussion on how cross-cultural discrepancies

_____

[7] It is interesting to point out that a given named entity could be presented as a snapshot or as a story. For example, if we take into account the conceptualization of the Taj Majal, this is subject to two types of features: (i) those that come from our synchronic view of the entity (i.e. snapshot), e.g. it is made of marble, two domes etc, and (ii) those features which are subject to time constraints (i.e. story), e.g. when the Taj Mahal was built, when it was declared a World Heritage Site, etc.

determine knowledge representations together with their linguistic realizations. Consequently, the next section focuses on the latter two types of knowledge.

4. Common-Sense Knowledge

4.1. Unsituated Common-Sense Knowledge

As regards the lexico-conceptual linkage with unsituated common-sense knowledge, languages clearly show contrasting differences that might be situated on a scale ranging from cases of full isomorphism, where two or more languages share exactly the same conceptual unit, to cases of full anisomorphism, where a conceptual unit is lexicalized in one language while it is absent in the rest. In between these two extreme poles, instances of partial anisomorphism complete the scale[8].

4.1.1. Full isomorphism

Full isomorphism is perhaps the less controversial case: two words in two different languages point towards the same cognitive entity, that is, the two share the same conceptual category. Linguistically, this phenomenon will result in potential translation equivalents, as shown in Figure 4.[9]

We can think of the words *parque, park, parco* and *Park* in Spanish, English, Italian and German respectively, with the meaning of a kind of "recreational place". As shown in Figure 4, the word w_2 in the Lexicon$_1$ (e.g. *parque* in the Spanish Lexicon) and the

---

[8] For an alternative modelling of multilingualism in ontologies, we refer the reader to Espinoza et al. (2009, Figure 1).

[9] With the aim of representing cultural distinctiveness diagrammatically, we use circles, boxes and arrows in this paper. Circles represent conceptual units, which consist of predications (e_1, e_2, ….. e_n) representing different meaning features. Moreover, each box illustrates the lexical repository in a given language. Finally, arrows link lexical units to a set of predications. Conceptual units are marked in capital letters followed by a suffix number, e.g. +CLEAN_00.

COREL (COnceptual REpresentation Language) is used for the formal representation of the meaning of conceptual units, but a natural language gloss is also provided. COREL-formatted representations consist of logically-connected predications, which can be computationally considered as a combination of conceptual graphs and frames. We refer the reader to Periñán-Pascual and Arcas (2004) and Periñán-Pascual and Mairal (2010) for a full description of its technicalities.

word w_3 in the Lexicon$_2$ (e.g. *park* in the English Lexicon) are both linked to the conceptual unit B. In this case, this conceptual unit is represented as +PARK_00 and has the meaning postulate (5).

(5)   +(e1: +BE_00 (x1: +PARK_00)Theme (x2: +ARTIFICIAL_AREA_00)Referent)

   *(e2: +BE_02 (x1)Theme (x3: +TOWN_00)Location)

   *(e3: +COMPRISE_00 (x1)Theme (x4: +PLANT_00)Referent)

   *(e4: +PLAY_00 (x5: +CHILD_00)Theme (x6)Referent (f1: x1)Location)

   A park is an artificial area; a park is in a town; a park has plants; children play in the park.

Hence, this category involves those cases where there are no cultural discrepancies and therefore is a frequent procedure for localizing internationally-standardized domains with an instrumental function.

Figure 4: Representation of full isomorphism.
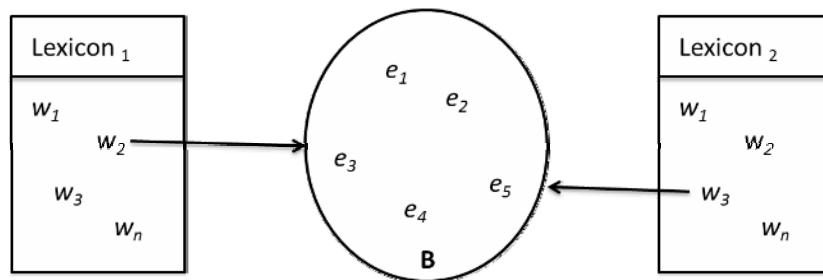
4.1.2. Full anisomorphism

This phenomenon includes those cases which have been traditionally treated as lexical gaps in contrastive linguistics, that is, one language lexicalizes a conceptual distinction which other languages do not. This contrast is represented in Figure 5.

Figure 5: Representation of full anisomorphism.
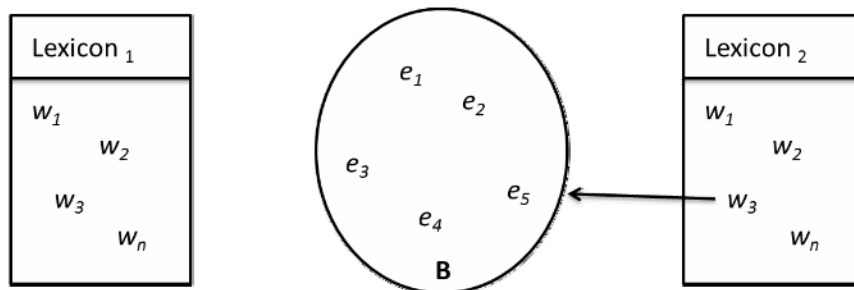
As illustrated in the above Figure, there is one lexical unit, e.g. w_3 in Lexicon$_2$, which is linked to the conceptual unit B. However, no lexical unit in Lexicon$_1$ lexicalizes this semantic feature. For example, consider the word *gazump* in British English, which has no lexical counterpart in Spanish. This word is linked to the conceptual unit $GAZUMP_00, whose meaning postulate is (6).

(6)   *((e1: +INCREASE_00 (x1: +HUMAN_00)Theme (x2: +PRICE_00)Referent (f1: (e2: ing past +SELL_00 (x1)Agent (x3: +HOUSE_00)Theme (x4)Origin (x5: +HUMAN_00)Goal))Scene (f2: (e3: ing fut +SELL_00 (x1)Agent (x3)Theme (x4)Origin (x6: +HUMAN_00)Goal))Reason) (e4: +COST_00 (x3)Theme (x2)Attribute))

When someone (x5) was about to purchase a house (x3), the seller (x1) raises the price of x3 because x1 is going to sell x3 to someone else (x6).

A similar case is *romería*, which is a Spanish culturally-bounded word that refers to a particular religious (and also social) festivity held typically in the south of Spain. To the best of our knowledge, there is no counterpart in any of the European languages, except in those spoken in the Iberian Peninsula (i.e. Basque, Catalan, Galician, Portuguese and Spanish). This is another case of full anisomorphism.

4.1.3. Partial anisomorphism

One of the most complex issues is to deal with those cases of partial anisomorphism. In order to deal with this phenomenon, we anticipate two possible scenarios:

(i) **Scenario A**: a language presents a lexicalization which is finer-grained than its lexical counterpart in another language. For instance, while English has *finger* and *toe*, Spanish only has *dedo* to refer to both.[10] A further question is to decide how the contrasting features will be accounted for in the Ontology. In this regard, let us consider

---

[10] English also has *digit*, which is cognate to *dedo* in Spanish. However, the use of the former is closely related to the anatomical or medical discourse. Therefore, *digit*, as well as other lexical units such as *carpal, metacarpal* or *phalanx*, should be linked to a specialized-domain ontology.

the case of *finger, toe* and *dedo*, whose lexico-conceptual linkage could be graphically represented as in Figure 6.

Figure 6: Representation of *finger, toe* and *dedo* in an ontological framework.
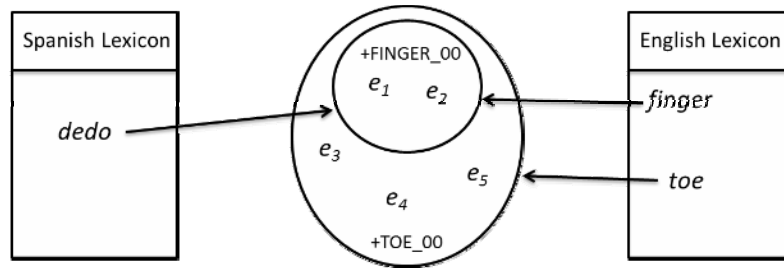
Here, we do find two conceptual units, a superordinate unit (i.e. +FINGER_00) which is defined by the predications *e1* and *e2* and a subordinate unit (i.e. +TOE_00) which, apart from the conceptual properties inherited from the superordinate, adds a few more predications, i.e. *e3, e4* and *e5*. Then, if we now look at the two lexica, we note that while the English predicates *finger* and *toe* are linked to two different conceptual units, i.e. +FINGER_00 and +TOE_00, the Spanish word *dedo* is linked just to the former, given that Spanish does not lexicalize the distinction of "fingers on the foot". Obviously, this scenario implies a semantic loss in Spanish. The examples (7) and (8) illustrate these distinctions by means of the conceptual representations of +FINGER_00 and +TOE_00 respectively.

(7) *(e_1: +BE_00 (x_1: +FINGER_00)Theme (x_2: +EXTERNAL_ORGAN_00)Referent)

  *(e_2: +COMPRISE_00 (x_1)Theme (x_3: 1 +NAIL_00)Referent)

  *(e_3: +COMPRISE_00 (x_4: +HAND_00 & +FOOT_00)Theme (x_1)Referent)

  A finger is an external organ; it has one nail; the hand has fingers.

(8) +(e_1: +BE_00 (x_1: +TOE_00)Theme (x_2: +FINGER_00)Referent)

  *(e_2: +BE_01 (x_1)Theme (x_3: +SHORT_00 & +THICK_00)Attribute)

  *(e_3: +COMPRISE_00 (x_4: +FOOT_00)Theme (x_1)Referent)

  A toe is a finger; a toe is relatively short and thick; the foot has toes.

At first sight, the predication *e3* in (7) could come into conflict with the predication *e3* in (8) when the semantic features of +FINGER_00 are inherited by +TOE_00. However, the FunGramKB inheritance mechanism allows non-monotonicity:

> In FunGramKB, each predication taking part in a meaning postulate is preceded by a reasoning operator in order to state if the predication is strict (+) or defeasible (*). Our inference engine handles predications as rules, allowing monotonic reasoning with strict predications, and non-monotonic with defeasible predications. (Periñán-Pascual and Arcas-Túnez, 2005, pp. 240)

More particularly, when two defeasible predications turn out to be contradictory, as with the above examples, the system overrides that predication being at the highest level of the Ontology. Therefore, the predication *e3* in the meaning postulate of +FINGER_00 is cancelled out by the predication *e3* in the meaning postulate of +TOE_00.

An oft-cited example which also illustrates this phenomenon of anisomorphism is the French distinction between *fleuve* ('river flowing into the sea') and *rivière* ('river flowing into another river'). In this case, FunGramKB has an ontological unit +RIVER_00, to which the English and the Spanish words *river* and *río* are linked, while the two French words would be linked to two different conceptual units $RIVER_00 and $RIVER_01, which contain different predications expressing the specific distinctive features, that is, the place where they flow into. Thus, the conceptual unit +RIVER_00 subsumes $RIVER_00 and $RIVER_01. It is interesting to note that, while probably English has lexical gaps for these two conceptual units, the Spanish word *afluente* can be linked to $RIVER_01. Moreover, if we consider the English word *stream* (i.e. a small narrow river) or its Spanish equivalent *riachuelo*, they will be linked to the conceptual unit $STREAM_00, which would be a subordinate of +RIVER_00 and a sibling of $RIVER_00 and $RIVER_01. In essence, although each language has a particular way to express this particular movement of water, it is possible to express these conceptual mismatches under the same ontology.

There are abundant examples of this type of partial anisomorphism, e.g. the French *voyage* and the two English distinctions *trip* and *journey*; the Spanish word *convento* where religious people live and the English *convent* (inhabited by nuns) and *monastery* (inhabited by monks), among many other examples. Figure 7 serves as a general illustration for this type of scenario.

Figure 7: Representation of partial anisomorphism: scenario A.
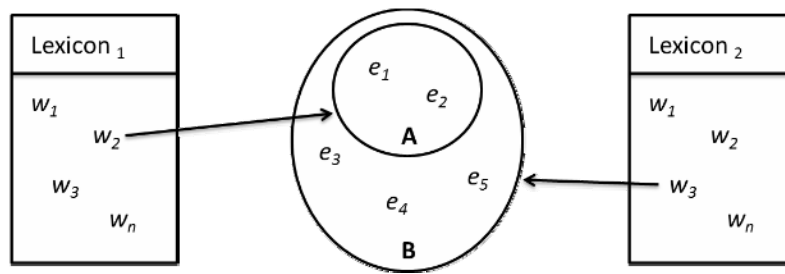
(ii) **Scenario B**: a language has a superordinate term whose referential scope involves a conjunctive or disjunctive relation between two other lexical items. For example, if we consider the lexical items *flora* and *fauna* in English and Spanish, the former refers to plants while the latter to animals. However, the English Lexicon has the word *wildlife*, which refers to both plants and animals, that is, its referential scope involves the sum of the two categories of referents. Figure 8 serves to represent this contrast. On the left hand side, we have the Lexicon of one language, where two words refer to two different conceptual units, A and C; on the right hand side, the Lexicon of a second language has one word which is linked to the conceptual unit B, which, in turn, involves the conceptual units A and C.

Figure 8: Representation of partial anisomorphism: scenario B.
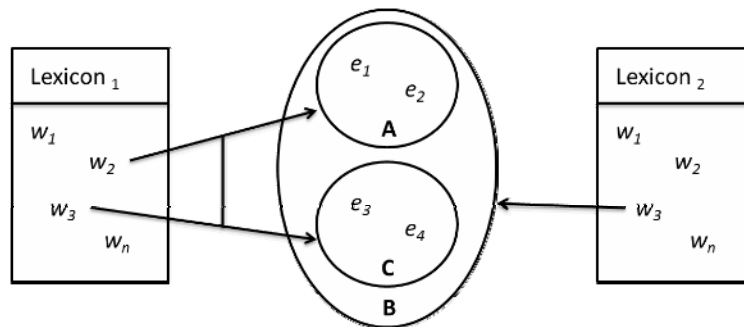
For example, the lexical units *fauna, flora* and *wildlife* are linked to the conceptual units +FAUNA_00, +FLORA_00 and +WILDLIFE_00, which have the conceptual representations (9), (10) and (11) respectively.

(9)   +((e1: +BE_00 (x1: +FAUNA_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: m +ANIMAL_00)Referent))

+(e2: +LIVE_00 (x1)Theme (f1: +NATURAL_AREA_00)Location)

Fauna is a group of many animals that live in a natural area.

(10)   +((e1: +BE_00 (x1: +FLORA_00)Theme (x2: +GROUP_00)Referent)(e2: +COMPRISE_00 (x2)Theme (x3: m +PLANT_00)Referent))

+(e2: +LIVE_00 (x1)Theme (f1: +NATURAL_AREA_00)Location)

Flora is a group of many plants that live in a natural area.

(11)   +(e1: +BE_00 (x1: +WILDLIFE_00)Theme (x2: +FAUNA_00 & +FLORA_00)Referent)

Wildlife is fauna and flora.

4.2. Situated Common-Sense Knowledge

As mentioned in Section 2, situated common-sense knowledge is the actual focus of the Cognicon. The representation of this type of knowledge is not new since it dates back to the work of Schank and Abelson's scripts (1977), Lakoff's scenarios (1987) or Ruiz de Mendoza's low-level situational models (2014). Indeed, this type of procedural knowledge is a suitable means for the description of social protocols, which form part of the daily routines of a given society. Hence, it is not surprising that these schemata are very sensitive to cultural distinctions. In fact, most cultural divergences take place in social protocols, and more particularly in three different types of scenario, i.e. everyday, bureaucratic and ritual situations, as can be seen in examples (a), (b) and (c) respectively:

(a)   In the United Kingdom and Ireland, people drive on the left, while in the rest of Europe it is on the right.

(b)   In those countries where ID cards are not operative, either the driving licence or the passport serves as a proof of identity. In Ireland, for instance, if a citizen on an election day wants to vote, his working card or student card will be sufficient as long as these have a photograph. In Spain, this would not be possible.

(c)   In a traditional wedding banquet in Spain, the bride and groom offer the guests a present, e.g. a fan for the ladies and a cigar for the men. In some other countries, e.g. Germany, this tradition does not exist.

These examples show a set of non-written instructions, a sort of "culturally-biased behaviour", that someone in a society should follow if he/she wants to behave in a socially acceptable manner; otherwise, that person would suffer a "cultural shock"

(Oberg, 2006). In order to avoid this psychological stress, it is essential to know those norms and protocols that form part of a society. Accordingly, Lee (2009) affirms

> […] societies ought to come with an instruction manual. Such a manual would be especially useful to immigrants and other longer term visitors. Indeed, the instruction manual might also be useful for training children born within the society. It may also have value for the socially marginalized within the society.

These social mechanics have to be learned, and in the case that concerns us in this paper, we should have the analytical tools at our disposal to represent this culturally-biased behaviour in a knowledge engineering framework, so that a computer can have access to it during text comprehension. In this regard, these social protocols have the format of scripts. As shown in Periñán-Pascual (2012), the FunGramKB script can be formed by a number of predications $(e_1 \ldots e_n)$ within a linear temporal framework—in particular, Allen's temporal model (Allen, 1983; Allen and Ferguson, 1994). In these scripts, every predication represents a particular semantic feature of a given event and is represented by a node in the propositional network. Let us consider the example of "greeting habits when visiting people at home in a non-business context". The example (12) presents the first COREL-formatted predications of the script @VISITING_SOMEBODY_00, where x1 is the guest and x6 is the host.

(12)   @VISITING_SOMEBODY_00

+((e1: +MOVE_00 (x1)Agent (x1)Theme (x2)Location (x3)Origin (x4: $ENTRANCE_00)Goal)(e2: +BE_02 (x4)Theme (x5: +DWELLING_00)Location)(e3: +LIVE_01 (x6)Theme (x5)Location))

+(e4: @REQUESTING_ENTRY_AT_THE_DOORWAY_00 (x1: x1, x5: x7; x9: x6))

+(e5: +MOVE_00 (x6)Agent (x6)Theme (x7)Location (x8)Origin (x9)Goal)

*((e6: +SEE_00 (x6)Theme (x1)Referent (f1: $PEEPHOLE_00)Instrument)(e7: +COMPRISE_00 (x9)Theme (x10: f1)Referent))

+(e8: +OPEN_00 (x6)Agent (x9)Theme (x11)Location (x12)Origin (x13)Goal)

+(e9: @GREETING_00 (x1: x1, x6: x3))

Somebody (x1) goes to the entrance of some type of dwelling place in which another person (x6) lives, and x1 requests to enter at the doorway; then, x6 goes

to the door and can decide to see x1 through the peephole (if there is one); finally, x6 opens the door and the two people greet each other.

As in the case of meaning postulates, the script predications are also preceded by a reasoning operator stating if they are strict (+) or defeasible (*). For example, suppose that John wants to visit his friend Mary. When John knocks on the front door of Mary's home, she has the choice of using the door peephole to check the caller's identity (defeasible predication). However, regardless of the expectations in the script to be fulfilled, Mary will undoubtedly open the door (strict predication). Otherwise, the execution of the script will be halted.

Moreover, some of the nodes in the propositional network can represent script activators, enabling us to invoke other scripts within a given script. For example, the predication (e4) in (12) is used to trigger the embedded script (13), which many other scripts will in turn reuse.

(13)   @REQUESTING_ENTRY_AT_THE_DOORWAY_00

*((e1: +PUSH_00 (x1)Agent (x2: +BELL_00)Theme (x3)Location (x4)Origin (x5)Goal)(e2: +BE_02 (x2)Theme (x6: +DOOR_00)Location (f1: +NEAR_00)Position)(e3: +COMPRISE_00 (x7: +DWELLING_00)Theme (x6)Referent))

*(e4: +KNOCK_00 (x1)Agent (x8)Theme (x9)Origin (x6)Goal)

Somebody rings a bell, which is close to a door, and/or that person knocks on the door.

A script activator takes the form of the name of a script together with a set of parameters which serve to integrate some of the entities involved in one given script (source script) into a new conceptual framework (target script).[11] Thus, the parameters in the script activator (14), which has been extracted from (12), are interpreted as the

---

[11] In Periñán-Pascual (2012), the terms "source" and "target" scripts correspond to "host" and "guest" scripts respectively. In this paper, we have chosen to use different names in order to avoid the ambiguity which might have resulted from analyzing a script such as @VISITING_SOMEBODY_00, i.e. a setting about a guest and a host.

mapping of values between some of the slots in @VISITING_SOMEBODY_00 (source script) and in @REQUESTING_ENTRY_AT_THE_DOORWAY_00 (target script). More particularly, the guest (x1) and the host's dwelling (x5) in the source script will be mapped to (x1) and (x7) respectively in the target script, whereas the door of the dwelling (x6) in the target script will be mapped to (x9) in the source script.

(14)   @REQUESTING_ENTRY_AT_THE_DOORWAY_00 (x1: x1, x5: x7; x9: x6)

It may also be noted that the predication (9) in @VISITING_SOMEBODY_00 serves to introduce the script @GREETING_00, where the two people involved in the salutation are actually used as the links between the source and the target scripts, that is, the guest (x1) and the host (x6) in @VISITING_SOMEBODY_00 will be mapped to (x1) and (x3) respectively in @GREETING_00, as shown in (15).

(15)   @GREETING_00 (x1: x1, x6: x3)

Indeed, (14) and (15) illustrate that embedded parameterized scripts allow us not only to minimize informative redundancy but also to incorporate culturally-bound cognitive schemata in the procedural knowledge repository. In particular, @GREETING_00 can be used as an umbrella script comprising varying socio-cultural customs from many countries around the world. For example, as shown in (16), @GREETING_SKO_00 can be used to store knowledge about the social etiquette on expressions of greeting in South Korea.

(16)   @GREETING_00

    +(e1: ….. | @GREETING_SKO_00 (x1: x1, x6: x3) | …..)

Within the script @GREETING_SKO_00, we will in turn describe various settings in which different profiles of participants could be involved, as shown in (17).

(17)   @GREETING_SKO_00

    +(e1: @GREETING_SKO_01 (x1: x1, x6: x3) ^ @GREETING_SKO_02 (x1: x1, x6: x3) ^ …)

For example, according to South Korean social conventions, @GREETING_SKO_01 can be employed to describe the men-greeting-men scenario (18), whereas @GREETING_SKO_02 would be used to describe the women-greeting-women scenario (19).

(18)   @GREETING_SKO_01

+(e1: +GREET_00 (x1)Theme (x2)Referent (x3)Goal ((f1: (e2: +BE_00 (x1)Theme (x4: +MAN_00)Referent))Condition & (f2: (e3: +BE_00 (x3)Theme (x5: +MAN_00)Referent))Condition))

+(e4: @BOWING_00 (x1: x1, x3: x3))

*(e5: @HAND_SHAKING_00 (x1: x1, x3: x3))

When greeting, two men bow to each other; they can also shake hands.

(19)   @GREETING_SKO_02

+(e1: +GREET_00 (x1)Theme (x2)Referent (x3)Goal ((f1: (e2: +BE_00 (x1)Theme (x4: +WOMAN_00)Referent))Condition & (f2: (e3: +BE_00 (x3)Theme (x5: +WOMAN_00)Referent))Condition))

+(e4: @BOWING_00 (x1: x1, x3: x3))

When greeting, two women bow to each other.

At first sight, the event +GREET_00 which is present in the predication *e1* in the above two scripts might seem redundant. Indeed, any other script included in @GREETING_SKO_00, or even in @GREETING_00, should also begin in the same way. Therefore, we could assume that if this cognitive component were placed in the top source script, redundancy could be minimized, since it is shared by all its target scripts. However, there are three problems with this approach. First, the actants in the event +GREET_00 are portrayed through a set of conditions that are to be met—e.g. (f1) and (f2) in (18) and (19)—if the script is intended to be triggered. These profile-based conditions should be linked to the behaviour of the actants, so the +GREET_00 predication should be placed at the beginning of each script. Second, these same conditions can be used to describe features not only of the actants but also of the settings. Thus, we can contextualize social behaviours according to parameters such as geography, sex, age or religion. Third, and most importantly, embedding +GREET_00 in the target scripts also enables verbal greetings (i.e. what you say) to be linked to physical greetings (i.e. what you do), as explained at the end of this section.

Moreover, the scripts (18) and (19) also have embedded scripts—e.g. @BOWING_00 and @HAND_SHAKING_00, described in (20) and (21) respectively. In fact, the former has been incorporated into both @GREETING_SKO_01 and @GREETING_SKO_02.

(20)   @BOWING_00

+(e1: $BOW_00 (x1)Theme (x2)Referent (f1: x3)Goal)

+(e2: $BOW_00 (x3)Theme (x4)Referent (f2: x1)Goal)

Somebody (x1) bows to another person (x3), and x3 bows to x1.

(21)   @HAND_SHAKING_00

+((e1: $SHAKE_HANDS_00 (x1)Theme (x2)Referent) (e2: +BE_02 (x2)Theme (x3)Location))

+((e3: $SHAKE_HANDS_00 (x3)Theme (x4)Referent) (e4: +BE_02 (x4)Theme (x1)Location))

Somebody (x1) shakes hands with another person (x3); and x3 shakes hands with x1.

But what if the socio-cultural customs of Spaniards, for example, were taken into account? First, the script @GREETING_00 (16) should have to be modified to (22).

(22)   @GREETING_00

+(e1: ….. | @GREETING_SKO_00 (x1: x1, x6: x3) | @GREETING_SPA_00 (x1: x1, x6: x3) | …..)

Thereafter, the procedure is similar to the one described above, since @GREETING_SPA_00 would comprise settings based on the different profiles of the participants. According to the algorithmic description in Figure 8, @GREETING_SPA_00 would be implemented as (23).[12]

(23)   +(e1: @GREETING_SPA_01 (x1: x1, x6: x3) ^ @GREETING_SPA_02 (x1: x1, x6: x3) ^ @GREETING_SPA_03 (x1: x1, x6: x3))

---

[12] Figure 8 does not intend to present a comprehensive view of the Spanish social etiquette in this domain, but our aim is to demonstrate how different variants of such a culturally-biased knowledge can be formally modelled.

Figure 9: Algorithmic representation of Spanish greeting conventions.
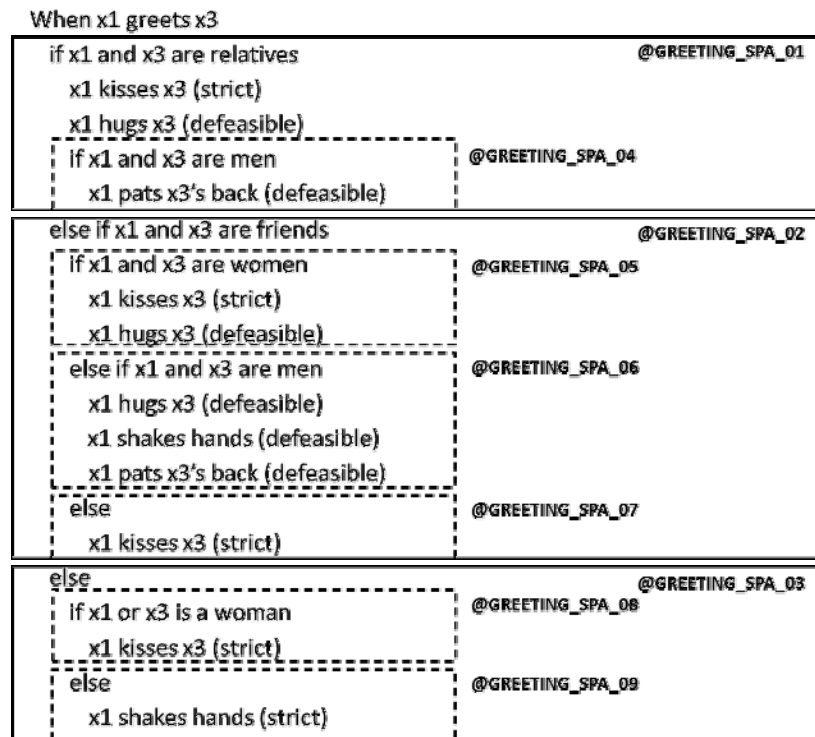
Furthermore, and similarly to the cases of (20) and (21), each physical expression of greeting in Figure 9 can be treated as a reciprocal action, so they would also take the form of scripts. Therefore, if we take into account all the scripts described in this section, we could model the network shown in Figure 10, where the arrows point to the source scripts.

Figure 10: A script-based network for @GREETING_00.
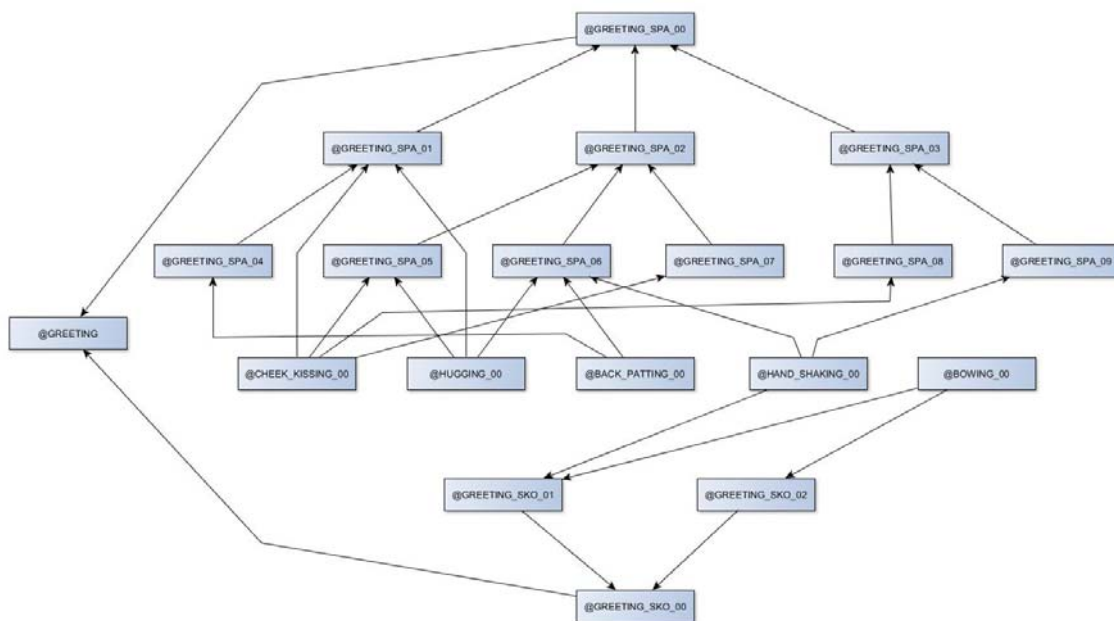
For the sake of reusability, when any of the scripts in Figure 10 is formally represented in the Cognicon, the parameters are not specified. Only in this way will we be able to reuse scripts in different contexts. For example, the various scripts described in this section have been ultimately connected to the script @VISITING_SOMEBODY_00, but other scripts such as @MEETING_SOMEBODY_ON_THE_STREET_00 can also draw on those greeting-related scripts. As a result, scripts are parameterized only when they are contextually integrated into a source script.

It is noteworthy that scripts should not only be regarded as a means of storing procedural knowledge, but they also aim to situate some linguistic realizations in the most adequate social setting. In the present case, verbal expressions of greeting would be modelled as idiomatic constructions, which can be instantiated as a conceptual preference of the Referent slot (x2) in the thematic frame of +GREET_00. For example, in Spain women often greet other women with *hola, guapa* ["hi, pretty one"] in a friendly manner; however, it is very unusual for men to show this affection to other men with a greeting such as *hola, guapo* ["hi, handsome"], unless the speaker is homosexual. Supposing that these two linguistic expressions were conceptualized as $SALUTATION_07, this concept would then be included in @GREETING_SPA_05 (24), but not in @GREETING_SPA_06.

(24)   +(e1: +GREET_00 (x1)Theme (x2: $SALUTATION_07)Referent (x3)Goal
((f1: (e2: +BE_00 (x1)Theme (x4: +WOMAN_00)Referent))Condition & (f2:
(e3: +BE_00 (x3)Theme (x5: +WOMAN_00)Referent))Condition))
+(e3: @CHEEK_KISSING_00 ())
*(e4: @HUGGING_00 ())

5. Cultural knowledge

As discussed in section 3, common-sense knowledge can be shaped by our cultural knowledge. This is why FunGramKB, apart from the Ontology and the Cognicon, postulates the Onomasticon, the place where our world knowledge is stored (cf. Figure 1). For example, although most cultural communities do share the same knowledge about what the season *summer* means, this same concept may acquire different semantic shades depending on the geographical location of a country: while in Spain summer

extends from June 23 to September 23, it is usually very hot, most people take a vacation, and it is a non-teaching period at educational institutions, among many other expectations, this is not the case in countries like Argentina and Chile, where summer extends from December 23 to March 23, it is cold, etc. The bottom line is that there are conceptual entities that, although shared by most cultural communities, are subject to our knowledge and/or experiences of the world, affecting the semantic burden of lexical definitions.

This group involves dealing with full isomorphism in the case of proto-structures (i.e. meaning postulates and scripts) but with anisomorphism in the case of their related bio-structures (i.e. snapshots and stories). On the one hand, these are cases of full isomorphism since the same notion exists in different cultures, but on the other hand the knowledge representation of this conceptual entity can be enriched with predications stemming from the cultural knowledge that is deeply grounded in some linguistic communities but not in others, which would be regarded as a case of anisomorphism. For example, we all agree that different cultures share conceptual notions such as *breakfast, lunch* or *dinner*, which would be cases of isomorphic common-sense knowledge. However, we would also agree that the conceptual scope of these notions differs if different countries are involved; for example, while in Spain lunch time goes from 14:00 to 16:00, in most other European countries there is a different schedule, typically from 12:00 to 14:00. This means that we need to resort to our knowledge of the world to understand the full conceptual scope of these notions. In sum, we have a conceptual notion which, although recognized in most linguistic communities, is shaped by specific cultural knowledge.

In keeping with the preceding sections, a further issue is how to represent these instances in a knowledge base. We have two words, e.g. w_2 and w_3, in the lexica of two different languages (e.g. *dinner* in English and *cena* in Spanish) and both are linked to the same conceptual unit in the Ontology, i.e. +DINNER_00. Moreover, this conceptual relation is enriched by information coming from other sources of knowledge, e.g. the Onomasticon, which is symbolized by clouds in Figure 11—e.g. the information that specifies that *dinner* in the United Kingdom typically goes from 18:00 to 20:00, while in Spain it goes from 21:00 to 23:00.
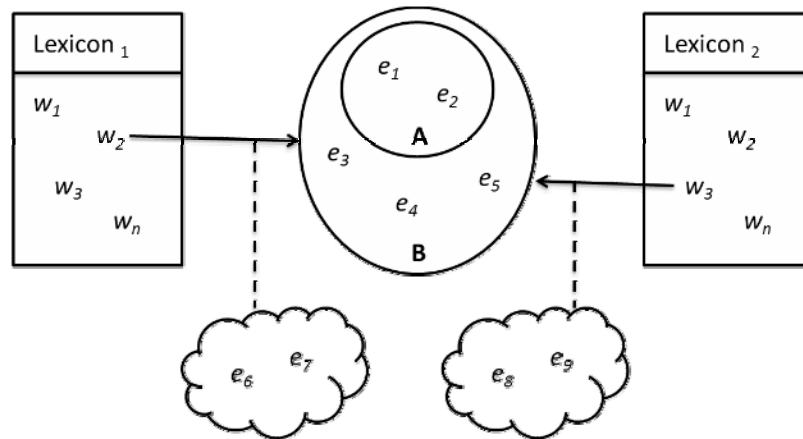
Figure 11: Integrating common-sense and world knowledge.

Let us consider the case of *number/license/registration plate* and *matrícula de coche* in English and Spanish respectively, which are clear instances of the integration of different types of knowledge. Both lexical units are linked to the ontological unit $PLATE_01, whose meaning is codified with the COREL scheme (25).

(25)   +(e1: +BE_00 (x1: $PLATE_01)Theme (x2: (+ARTIFICIAL_OBJECT_00 & +CORPUSCULAR_00 & +SOLID_00) | +WRITING_00)Referent)

+(e2: +BE_01 (x1)Theme (x3: +METAL_00)Attribute)

*(e3: +BE_02 (x1)Theme (x4: +CAR_00)Location (f1: +FRONT_00 & +BACK_01)Position)

A plate is an artificial, corpuscular and solid object and/or a piece of writing; it is made of metal; it is located in the front and at the back of a car.

Then, this knowledge will be expanded by the cultural knowledge stored in the COREL-formatted snapshots of %SPAIN_00 and %UNITED_KINGDOM_00, which are partially represented in (26) and (27) respectively:

(26)   *(e1: +BE_01 (x1: $PLATE_01)Theme (x2: +WHITE_00)Attribute)

*((e2: +COMPRISE_00 (x1)Theme (x3: 4 +NUMBER_00 & 1 +SPACE_00 & 3 +CONSONANT_00)Referent)(e3: +BE_01 (x3)Theme (x4: +BLACK_00)Attribute))

A vehicle plate is white; it has four numbers, one space and three consonants; the numbers and the consonants are black.

(27)　*((e1: +COMPRISE_00 (x1: $PLATE_01)Theme (x2: 2 +LETTER_00 & 2 +NUMBER_00 & 1 +SPACE_00 & 3 +LETTER_00)Referent)(e2: +BE_01 (x2)Theme (x3: +BLACK_00)Attribute))

*((e3: +BE_02 (x4: $PLATE_01)Theme (x5: +CAR_00)Location (f1: +FRONT_00)Position)(e4: +BE_01 (x4)Theme (x6: +WHITE_00)Attribute))

*((e5: +BE_02 (x7: $PLATE_01)Theme (x5)Location (f2: +BACK_01)Position)(e6: +BE_01 (x7)Theme (x8: +YELLOW_00)Attribute))

A vehicle plate consists of a sequence of two letters, two numbers, one space and three letters, being all of the letters and numbers in black; the front plate is white and the rear plate is yellow.

Cimiano et al. (2010) discussed a very interesting example of culturally-biased notions: the role of head of government—e.g. *Bundeskanzler* (Germany), *presidente* (Spain) and *prime minister* (UK)—and the role of head of state—e.g. *Bundespräsident* (Germany), *monarca* (Spain) and *monarch* (UK). This is represented by means of two different ontological units, that is, $HEAD_OF_GOVERNMENT_00 and $HEAD_OF_STATE_00. However, the meaning postulates of these two conceptual units are underspecified, since their definition can be enriched with information related to the political background of their corresponding geographical zones. Hence, the political functions of the head of government and the head of state in a given country, which differ in each of the three countries, can be described by means of the snapshot which serves to describe the socio-political system of the country. This allows us to differentiate not only the functions between the two conceptual units in the Ontology but also the functions of each in the three different countries in the Onomasticon.

6. Conclusions

Recent research in the area of the multilingual semantic web has stressed the importance of designing and developing multilingual resources. In this regard, this paper is concerned with multilingualism and conceptual modelling within a knowledge engineering project like FunGramKB. After describing some of the most relevant approaches in the field, the paper argues for an external model in which a clear separation of the conceptual level and the lexical level is one of the hallmarks. Moreover, a differentiating feature of our proposal is that it provides a more

comprehensive cognitive layer, which not only consists of semantic knowledge in the Ontology but also makes a strong commitment towards capturing procedural knowledge in the Cognicon and world knowledge in the Onomasticon. Cultural mismatches have been discussed and formally represented in these three modules. A further line of research is to exploit FunGramKB knowledge schemata in the multilingual Linked Open Data community, which requires that these schemata should be mapped to representations based on standards such as RDF, so that they can be linked up with a large number of linguistic resources (cf. Chiarcos et al., 2012; Gracia et al., 2012; Montiel-Ponsoda et al., 2011).

Acknowledgments

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.

Alonge, A., Calzolari, N., Vossen, P., Bloksma, L., Castellón, I., Martí, M.A., Peters, W., 1998. The Linguistic Design of the Eurowordnet Database. Computers and the Humanities 32 (2-3), 91-115.

Arms, W.Y., 2000. Digital Libraries. The MIT Press, Cambridge (Mass.).

Barsalou, L.W., 1991. Deriving Categories to Achieve Goals. In: Bower, G. (Ed.), The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 27. Academic Press, San Diego, 1-64.

Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. Scientific American 284 (5), 34-43.

Buitelaar, P., Cimiano, P., Haase, P., Sintek, M., 2009. Towards Linguistically Grounded Ontologies. In: Proceedings of the 6th European Semantic Web Conference, 111-125.

Chiarcos, Ch., Nordhoff, S., Hellmann, S. (Eds.), 2012. Linked Data in Linguistics. Representing and Connecting Language Data and Language Metadata. Springer: Berlin.

Cimiano, P., Montiel-Ponsoda, E., Buitelaar, P., Espinoza, M., Gómez-Pérez, A., 2010. A Note on Ontology Localization. Applied Ontology 5(2), 127-137.

Espinoza, M., Montiel-Ponsoda, E., Gómez-Pérez, A., 2009. Ontology Localization. In: Proceedings of the 5th International Conference on Knowledge Capture (K-CAP 2009). Redondo Beach, California.

Fuertes, P.A., Gordo Gómez, P., Niño Amo, M., Rios Rodicio, A. de los, 2010. Expert Knowledge and Translation Equivalents in the English Spanish Accounting Dictionary. In: Caballero, R., Pinar, M.J. (Eds.) Ways and Modes of Human Communication. Universidad de Castilla La Mancha, Ciudad Real, 1083-1089.

Gracia, J., Montiel-Ponsoda, E., Cimiano, P., Gómez-Pérez, A., Buitelaar, P., McCrae, J., 2012. Challenges for the Multilingual Web of Data. Journal of Web Semantics 11, 63-71.

Lakoff, G., 1987. Women, Fire, and Dangerous Things. University of Chicago Press, Chicago.

Lee, R.M., 2009. Mitigating Culture Shock: E-Learning Cultural Scripts. In: http://www.fiu.edu/~ciber/ws/Mitigating_culture_shock_RLee.pdf

Mairal-Usón, R., Periñán-Pascual, C., 2009. The Anatomy of the Lexicon within the Framework of an NLP Knowledge Base. Revista Española de Lingüística Aplicada 22, 217-244.

Mairal-Usón, R., Periñán-Pascual, C., 2015. Representing Constructional Schemata in FunGramKB Grammaticon. In: Fleischhauer, J., Latrouite, A., Osswald, R. (Eds.) Exploring the Syntax-Semantics Interface. Düsseldorf University Press, Düsseldorf.

McCrae, J, Aguado-de-Cea, G., Buitelaar, P., Cimiano, P., Declerck, T., Gómez Pérez, A., Gracia, J., Hollink, L., Montiel-Ponsoda, E., Spohr, D., Wunner, T., 2010. The Lemon Cookbook. In: http://lexinfo.net/lemon-cookbook.pdf.

Montiel-Ponsoda, E., 2011. Multilingualism in Ontologies: Multilingual Lexico-Syntactic Patterns for Ontology Modeling and Linguistic Information Repository for Ontology Localization. Unpublished PhD thesis. Universidad Politécnica de Madrid, Madrid.

Montiel-Ponsoda, E., Aguado-de-Cea, G., Gómez-Pérez, A., Peters, W., 2008. Modelling Multilinguality in Ontologies. In: Proceedings of the 21st International Conference on Computational Linguistics.

Montiel-Ponsoda, E., Gracia, J., Aguado-de-Cea, G., Gómez-Pérez, A., 2011. Representing Translations on the Semantic Web. In: Proceedings of the 2nd Workshop on the Multilingual Semantic Web, Bonn, 25-37.

Moreno Ortiz, A., 2000. OntoTerm: Un Sistema Abierto de Representación Conceptual. Procesamiento del Lenguaje Natural 26, 61-62.

Oberg, L., 2006. Culture Shock and the Problem of Adjustment to New Cultural Environments. Practical Anthropology 7, 177-182.

Panton, K., Matuszek, C., Lenat, D., Schneider, D., Witbrock, M., Siegel, N., Shepard, B., 2006. Common Sense Reasoning: From Cyc to Intelligent Assistant. In: Cai, Y., Abascal, J. (Eds.), Ambient Intelligence in Everyday Life. Springer, Berlin, 1-31.

Periñán-Pascual, C., 2012. The Situated Common-Sense Knowledge in FunGramKB. Review of Cognitive Linguistics 10 (1), 184-214.

Periñán-Pascual, C., 2013. A Knowledge-Engineering Approach to the Cognitive Categorization of Lexical Meaning. VIAL: Vigo International Journal of Applied Linguistics 10, 85-104.

Periñán-Pascual, C., Arcas-Túnez, F., 2004. Meaning Postulates in a Lexico-Conceptual Knowledge Base. In: Proceedings of the 15th International Workshop on Databases and Expert Systems Applications. IEEE, Los Alamitos, 38-42.

Periñán-Pascual, C., Arcas-Túnez, F., 2005. Microconceptual-Knowledge Spreading in FunGramKB. In: Proceedings of the 9th IASTED International Conference on Artificial Intelligence and Soft Computing. ACTA Press, Anaheim-Calgary-Zurich, 239-244.

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., 2010. The Architecture of FunGramKB. In: Proceedings of the 7th International Conference on Language Resources and Evaluation. ELRA, Malta, 2667-2674.

Periñán-Pascual, C., Arcas-Túnez, F., 2011. Introduction to FunGramKB. Anglogermánica Online 8, 1-15.

Periñán-Pascual, C., Mairal-Usón, R., 2009. Bringing Role and Reference Grammar to Natural Language Understanding. Procesamiento del Lenguaje Natural 43, 265-273.

Periñán-Pascual, C., Mairal-Usón, R., 2010. La Gramática de COREL: Un Lenguaje de Representación Conceptual. Onomázein 21, 11-45.

Ruiz de Mendoza, F.J., 2014. Low-level Situational Cognitive Models within the Lexical Constructional Model and their Computational Implementation in FunGramKB. In: Nolan, B., Periñán-Pascual, C. (Eds.), Language Processing and Grammars: the Role of Functionally Oriented Computational Models. John Benjamins, Amsterdam/Philadelphia, 367-390.

Schank, R.C., Abelson, R.P., 1977. Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. Lawrence Erlbaum, Hillsdale.

Smith, B., 1995. Formal Ontology, Common Sense and Cognitive Science. International Journal of Human-Computer Studies 43(5-6), 641-666.

Van Dijk, T.A., 2003. Specialized Discourse and Knowledge. A Case Study of the Discourse of Modern Genetics. Cadernos de Estudos Lingüísticos 44, 21-55.

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

Van Valin, R.D. Jr., Mairal-Usón, R., 2014. Interfacing the Lexicon and an Ontology in a Linking Algorithm. In: Gómez González, M.A., Ruiz de Mendoza, F., Gonzálvez-García, F. (Eds.), Theory and Practice in Functional-Cognitive Space. John Benjamins, Amsterdam, 205-228.