The Domain of the Lexical Functions Fact₀, CausFact₀ and Real₁

María A. Barrios Rodríguez
Universidad Complutense de Madrid
Ciudad Universitaria. 28008-Madrid
auxiba@filol.ucm.es

Abstract

The aim of this paper is to show that the study of the domain of meaningful lexical functions by means of glosses and semantic labels can be a great help when listing keywords and values linked to different lexical functions: the results of the study of the domains of Fact₀, Caus₀Fact₀ and Real₁ for CALLEX-ESP—the Spanish version of the computer-assisted lexicon learning program CALLEX—are shown. Following Apresjan (2001) the keywords of this family of lexical functions belong to the ‘artifact’ semantic label. A discussion of the different nature of IncepFact₀ and CausFact₀, the overlapping of lexical functions Caus₁Fact₀ and IncepReal₁, and the prediction of the domain of Fact₀, Caus₀Fact₀ and Real₁ are also included.

Keywords
Domain of lexical function, complex lexical function, links between lexical functions, fulfilment verbs, Meaning-Text Theory.

1 Introduction

The domain of a lexical function is the set of lexical units (LU) for which a Lexical Function (LF) is defined (Mel’čuk, 1996, 76). The domain of a LF can be established more easily by means of glosses and paraphrases (concepts largely explained by Alonso Ramos, 2005, 2006; Popovic, 2003; Polguère and Kahane, 2001). From a theoretical point of view, the domain of a LF allows us: 1) to specify the scope of a lexical function; 2) to establish the existing links between different LFs (Kahane, 2002); 3) to predict some of the possible collocations (Polguère, 2003). From a practical point of view, the study of the domain of LFs allows us to develop more easily: 1) a kind of implementation of the Lexical Inheritance Principle (Mel’čuk and Wanner, 1996; Mel’čuk, 1996, 76-78); 2) some pedagogical applications, like computer-assisted language learning programs (Apresjan et al, 2002, 2003, 2007, in press; Boguslavsky et al, 2006).

In 2003, a computer-assisted language program, CALLEX, was presented by Apresjan and other members of the Institute for Information Transmission Problems of the Russian Academy of Sciences (Apresjan et al: 2003). The program was developed for Russian, English and German (in cooperation with German and Austrian partners within the INTAS program). The system was planned for both learners and natives trying to improve their mother tongue
language. Its Spanish version, CALLEX-ESP is now under development (Boguslavsky et al., 2006), and it has involved the study of more than 2,000 Spanish collocations. The results we present in this paper are a direct consequence of the work carried out for CALLEX-ESP, where almost all the encoding of the LFs follows the CALLEX and the Dicouèbe model.

From a theoretical perspective, scholars claim that some LFs such as Real, Fact and Labreal need a more accurate definition. Thus, modality (-M) LFs (Apresjan: 2001) and notations (see Dicouèbe') are added in order to achieve a better result, but even so, several problems remain. For instance, the Dicouèbe entries for cat and dog include the following LFs, Fact0-mouvement, Fact0-nettoyage, Fact0-gratter, Fact0-casser, Fact0-sentir and Fact0-mordre. The lexical function Fact0 means 'to fulfil the requirement of L' (Mel’čuk, 1996: 68), but it is not clear which the requirements for cats and dogs are. This vagueness explains why Fact0 needs some annotation. Should we understand the translation of these LFs into the natural language as if ‘to fulfil the requirement of movement for a cat’ is to stretch, or ‘to fulfil the requirement of cleaning for a cat’ is to clean, as this dictionary shows? If these keywords cat and dog fulfil one requirement and they are classified as ‘companion animal’ or ‘domestic animals’, why does Fact0 not reflect the sense of ‘to accompany’? Although the importance of the problem would recommend a further study, the lack of space makes it difficult to achieve a solution here, but some suggestions for future works will be included.

Regarding the problem of links between LFs, different papers deals with links, although no specific works are focuses on this issue. Kahane (2002) claims that there are some LFs in the DEFC entry blessure that could be substituted by other complex LFs containing fulfillment or support verbs: for instance, blessure mortelle could be covered by A1Fact0, instead of Magn. Popovic (2003a; 2003b) points out the apparent contradiction in some codifications, like causer une plaie (to cause an injury), coded in DiCo as Caus2Func0, and causer une angoisse (to cause anguish), coded as Oper21. Alonso Ramos (2004:251) mentions the case of to take a photo as a traditional example of a link between Oper1 and CausFunc0. All these problems, however, remain unsolved.

In connection with how to establish the domain of a LF, Apresjan (2001: 5) underlines two requirements (transparency and systematisation) working on a computer lexicon in order to show: a) semantic and syntactic similarities and differences among LFs of the same family; and b) their relation with lexemes that belong to the same lexical field (see collocations listed by different LFs in Apresjan et al., 2007). His study about Fact and Real family (2001) reveals two different groups of keywords: artifacts and modal lexemes (see section 2). On the other hand Polguère underlines the relation among semantic labels, paraphrases and glosses (2003:63). He presents the example of the three levels of realization of LFs that can be established for those nouns referring to any kind of ‘container’ (to fill it, to use it, and to empty it), and he adds that it could be interesting to present the links between labels and lexical functions in natural language in the LAF2 because these links could have applications in LFs learning (2003: 64).

---

1 Lexical French Data Base developed by OLST (Observatoire de linguistique Sens-Texte) by Mel’čuk and Polgùre (Steinlin et al., 2004), accessible by an interface via http://olst.ling.umontreal.ca/dicoueb/.

2 The Lexique actif du français (Mel’čuk and Polgùre: 2007) is a dictionary of French where LFs are present in natural language (as glosses).
Following their models and suggestions, the present paper studies the links between some semantic labels and some LFs’ paraphrases and glosses. For some sets of nouns, it is possible to predict some kinds of paraphrases and glosses, such as *to wear clothes*, where *to wear* is the gloss of *Real₁* and *clothes* can be substituted for all the words under the semantic label ‘clothes’. These predictions are made possible by an implementation of the Lexical Inheritance Principle (Mel’čuk and Wanner: 1996; see Barrios and Bernardos, in this volume).

This paper is then organized as follows. Section 2 deals with the notion of the domain of *Fact₀*, *IncepFact₀*, *CausFact₀*, *Real₁* and the conceptual nature of some other relations. Section 3 shows the link between *IncepReal₁* and *Caus₁Fact₀* and proposes a definition for *Caus₀(Incep)Fact₀*. Section 4 points out the different nature of *IncepFact₀* and *CausFact₀*, and suggests the conditions to allow the *Fact₀* domain for any keyword to be included in the *Caus₀(Incep)Fact₀* domain. Section 5 raises the question of whether it is possible to predict values for *Fact₀*, *Caus₀Fact₀* and *Real₁*. Finally, Section 6 summarizes the use of these LFs in the implementation developed for CALLEX-ESP and presents the conclusions.

2 The domain of *Fact₀*, *IncepFact₀*, *CausFact₀* and *Real₁*

Among the mathematical concepts related to *function*, it is interesting to pay attention to the term *domain*: in the mathematical field, it is the set of all the possible values (inputs) that a function is created for. As Mel’čuk (1996:76) proposes, the *domain of the lexical functions* is here understood as the set of LUs for which each LF was defined, in other words the set of arguments of every LF. It must be understood as a potential domain: the set of LU for which the LF has a potential meaning, but not necessarily a value. Thanks to this notion of the domain of a LF, it is possible to apply the Lexical Inheritance Principle in two ways: the values of the LFs can also be shared by the keywords, and the keywords can also be shared by the LF itself (Mel’čuk, 1996: 76-78). The option chosen here is the second one.

Glosses and paraphrases make it easier to understand LFs (Alonso Ramos, 2006, 2005; Popovic, 2003; Polguère and Kahane, 2001), and, consequently, to specify the domain for each of them. Alonso Ramos (2005) distinguishes three kinds of *glosses*, naming the first type, *paraphrase*, with the meaning of a short explanation (2006:74); in CALLEX-ESP this kind of gloss is called *semantic definition* (in order to avoid the ambiguity with the sense of paraphrase that appears, for instance, in paraphrase rules). LFs *Fact* and *Real* present the distinction made by Apresjan between *Factᵢ* and *Factᵢ-M*, *Realᵢ* and *Realᵢ-M*: *Factᵢ-M* is used for arguments such as *advice, demand, or desire* which have a sense of ‘want’, ‘must’, ‘must not’, and *Realᵢ-M* for those with a sense of ‘modality’ (Apresjan, 2001; Apresjan et al: 2003).

This paper focuses on the following semantic definitions in CALLEX³: *Fact₀*, *IncepFact₀*, *CausFact₀* and *Real₁* (in order to appreciate the differences with (-M) LFs, semantic definitions of *Fact₀-M*, *CausFact₀-M* and *Real₁-M* will also be shown). In the semantic definitions, L is the keyword and a verb collocate takes L as its grammatical subject. *Fact₀* is defined as “L functions or is functioning according to its destination” as in *an airplane flies or a factory works*. *Fact₀-M* is defined as “L is in the state corresponding to normal expectations in the respective situation”, as in *an attack succeeds or a hope comes true*. *IncepFact₀* is defined as “L starts to function according to its destination” as in *a bus starts, a restaurant...*  

³ All the paraphrases and the examples given are taken from the English version of CALLEX.
CausFact₀ is defined as “to cause L to function according to its destination” as someone explodes a bomb or drives a car. CausFact₀-M is defined as “to cause L to be in the state corresponding to the normal expectations in the respective situation”, as in settle the argument, or execute the sentence. In both CausFact₀ and CausFact₀-M, a collocate verb takes an argument (it could be or not an actant of L) as its grammatical subject and L as its primary object¹. Real₁ is defined as “to use L according to its destination” as in ride a bicycle, or read a book. Real₁-M is defined as “to do with regard to L that which is normally expected of X” as in prove the accusation, or realize one’s hope. In both Real₁ and a Real₁-M, a collocate verb takes X (the first actant of L) as its grammatical subject and L as its primary object.

According to Apresjan, the keyword sense required for Fact₀, IncepFact₀, CausFact₀ and Real₁ is ‘destination’ or ‘any normal expectation in the respective situation’. Consequently, the domain of these LFs will be provided by the set of LUs which carry out this condition. If we have a look at Dicouèbe’s proposal, most of the LFs Real, Fact and Labreal fulfil this condition, thereby proving that even though the OLST does not take into account modality (-M) LFs, their understanding of the meaning of these LFs is quite similar. However, some codifications do not adjust to the limit above mentioned because they include animals; let’s have a look at the following (1-5) examples:

(1) Factᵣ-mouvement(chat) = s’étirer [Factᵣ-move(cat) = stretch]
(2) Factᵣ-déplacement(serpent) = ramper [Factᵣ-displacement(snake) = crawl]
(3) Factᵣ-chasser X(chat) = chasser (X = souris) [Factᵣ-hunt X(cat) = hunt (X = mouse)]
(4) Factᵣ-mordre(chien) = mordre, mordiller, ronger [Factᵣ-bite(dog) = bite, gnaw]
(5) Factᵣ-nettoyage(chat) = se lustrer, se nettoyer [Factᵣ-cleaning(cat) = clean, polish]

As pointed out in Section 1, the question that arises is: if the destination of a cat is to hunt mice, why is it necessary to add an annotation to the LF? Regarding the movement of a cat, the ‘normal expectation in the respective situation’ is to stretch, or to curl up in a ball, and both verbs must be included in an explanatory and combinatory dictionary, as Dicouèbe proposes. However, it is not clear which of the annotations for Fact₀ is the best to encode this information, especially if the redundancy that this encoding implies is taken into account. Each of the following situations, such as cleaning or scraping, claims a different ‘normal expectation’ (as to lick or to scratch respectively) which are actually synonym values for the annotations.

It is now necessary to come back to the definition of fulfilment verbs where it is clearly mentioned “to fulfil the requirement of L” (Mel’čuk, 1996: 76-78) or its destination (Apresjan: 2001:6), in order to avoid any allusion to the requirement of the LF or the LF annotation (to stretch would be not the ‘normal expectation’ of the L cat, but of the movement of the cat).

---

¹ We will use Caus₀ᵣFactᵣ, even though CALLEX only uses CausFact₀, because it contains the sense we are focusing on. Discussion about the convenience of Caus₀ᵣFactᵣ or CausFact₀ exceeds the possibilities of this study.
On the other hand, even if human beings find any kind of ‘usefulness’ in some animals, they do not have only one destination and, usually, there is more than one expectation: hunt for cats but also for some kinds of dogs; bite for dogs but also defend. If we consider this criterion, we should include other verbs like steal for monkeys because of their ability to steal food from humans as it is frequently seen in many African countries. Although there are idioms that reflect the links between an animal and something that we expect from it, such as the Spanish expression estar como el perro y el gato (to be like the cat and the dog) it does not imply that the linguistic expression of this action would be a value of Fact0. Actually, this idiom is used to describe very different situations such as a fight, an argument or even after two people searching one another without success.

For the family of the LFs studied here, more examples of ‘usefulness of’ could be found in which they are used in the same way. Moreover, examples of (1-2), ‘movement typical of’ is somewhat similar to ‘sound typical of’. Actually the LF Son is equal to Fact0 when the keyword is a musical instrument (ref. Dicouèbe). If there are several values for ‘movement typical of’, a non standard or semi-standard LF could be proposed, which would cover examples (1-2), plus Fact0(coq) = picorer (Fact0(cock) = peck at, Dicouèbe), and other relations such as peacock shows off, waddle, etc.; other expressions should be reviewed and, when there is not a linguistic but a conceptual relation (as in kangaroo jumps, dog bites and the like), it could be interesting to encode them in the same way as the values of Mero (‘part of’) and Hypo (‘kind of’) in Dicouèbe (encoding these relations is more conceptual than linguistic, as the heterogeneous list of values proves). There is nothing special in jump, nor in bite that can justify its attachment to these keywords.

The conclusion is that the meaning ‘destination’ or ‘any normal expectation in the respective situation’ is an essential requirement for the arguments of Fact0, IncepFact0, Caus0Fact0 and Real1 (and consequently IncepReal1) domains, where ‘any normal expectation’ must be understood as an expectation of the argument, not of the LF annotation. ‘Animals’ would be left out because the nature of the concept ‘animal’ is too complex to be attached to ‘any normal expectation in the respective situation’, and we should keep to ‘artifacts’, ‘instruments’, and so on (see lists on Apresjan 2001, 2007), because they were created for some purpose.

3 Links between Caus1Fact0 and IncepReal1

Many collocations can be covered simultaneously by Caus1Fact0 and IncepReal1. Examples from Dicouèbe (6-9) can illustrate the link among these LFs. Because of the non-uniqueness of linguistic descriptions, links are possible and correct. However, some examples (6-7) reveal paraphrastic relations between formulas but no links; the more the MTT model is free of ambiguity, the more it can help to solve the problem of the ambiguity in natural language.

(6) \textit{Caus1Fact0}(foulard) = mettre \textit{Caus1Fact0}(scarf) = to put on

(7) \textit{IncepReal1}(bracelet) = mettre \textit{IncepReal1}(bracelet) = to put on

(8) \textit{Caus1Fact0}(frein) = mettre \textit{Caus1Fact0}(brake) = to put on

(9) \textit{Real1}(frein) = freiner \textit{Real1}(brake) = to brake

It is quite obvious that expressions (6-7) such as to put on a scarf and to put on a bracelet do not reveal enough differences to justify two different LFs. However, the fact that they have
been used does not indicate the existence of a link, but rather the proximity of semantic definitions of formulas, which could be understood as two paraphrases of the same meaning: ‘to cause L to function according to its destination’ (semantic definition of $\text{Caus}_0\text{Fact}_0$) and ‘to start using L according to its destination’ (semantic definition of $\text{Incep}_1\text{Real}_1$). So, the problem that arises is whether these two meanings are one and the same meaning or not.

As $\text{Caus}_0\text{Fact}_0$ and $\text{Incep}_1\text{Real}_1$ have the same syntactic structure, it would be interesting to see if they have the same or a different meaning (Polguère and Kahane, 2001). From our point of view, the LFs conveyors of meaning (as causatives and phasal) should impose a semantic restriction on the base they select.

At this point, we must take into account the notion of granularity (the degree of detailedness: Polguère and Kahane, 2001). In this sense $\text{Incep}_1\text{Real}_1$ could be enough for all these collocations. However, our proposal is the following: firstly, to underline the sense of $\text{Incep}$ (‘to begin’) that is usually included in the meaning ‘to cause’ (Mel’čuk, 1996:68) by using $\text{Caus}_0\text{Incep}_0\text{Fact}_0$, whose semantic definition would be “to cause X to start to function according to its destination”; in this way it would be easier to restrict the domain of $\text{Caus}_0\text{Incep}_0\text{Fact}_0$ to the keywords containing the meaning ‘possible to be activated by someone’ (as *to turn on the television*), and consequently to distinguish another group of artifacts like a bus or a plane (*to drive a bus, to pilot a plane*), which are encoded by $\text{Caus}_0\text{Fact}_0$ (Apresjan: 2001; 2007); and secondly, to formulate the hypothesis that the keywords of the $\text{Caus}_0\text{Fact}_0$ domain (and consequently of $\text{Caus}_0\text{Incep}_0\text{Fact}_0$) must also belong to the $\text{Fact}_0$ domain.

Section 4 focuses on the hypothesis that belonging to $\text{Fact}_0$ domain is necessary in order to be codified as $\text{Caus}_0\text{Fact}_0$.

### 4 The nature of LFs $\text{Caus}_0\text{Fact}_0$ and $\text{Incep}_0\text{Fact}_0$

It is interesting to note that in the *Dicouèbe* none of the keywords in the $\text{Incep}_0\text{Fact}_0$ domain is included in the $\text{Fact}_0$ domain. However, most of the *Dicouèbe* keywords in the $\text{Caus}_0\text{Fact}_0$ domain belong to $\text{Fact}_0$ domain.

As the term *function* is used in a mathematical sense (Mel’čuk: 1996), making use of the algebraic properties of mathematical functions for a complex LF as $\text{Caus}_0\text{Fact}_0$, the *Caus* domain would fit into the image of the $\text{Fact}_0$ domain. This would mean that keywords of $\text{Caus}$ should be values of $\text{Fact}_0$. Consequently, the $\text{Caus}_0\text{Fact}_0$ domain would be a subdomain of the $\text{Fact}_0$ domain (the numerical difference of keywords in each one in *Dicouèbe* would corroborate it). This would not occur in the case of $\text{Incep}_0\text{Fact}_0$, which does not have this property (because its behaviour is different, see note 5).

Figure 1 shows the hypothetical correspondence of some *Dicouèbe* collocations for these LFs:

---

[5] Only *bateau* is in $\text{Incep}_0\text{Fact}_0$ domain (*prendre la mer*) and in $\text{Fact}_0$ domain (*naviguer*); the rest of the keywords for $\text{Incep}_0\text{Fact}_0$ are: *bourgeon* (*éclater*), *ciment* (*durcir*), *clinique* (*ouvrir*), *colle* (*sécher*), *gaieté* (*éclater*), *maison* (*ouvrir ses portes*), *musée* (*ouvrir ses portes*), *parachute* (*s’ouvrir*). On the other hand, there are collocations for both $\text{Caus}_0\text{Fact}_0$ and $\text{Fact}_0$ for *ampoule* (*Caus}_0\text{Fact}_0* *allumer*, $\text{Fact}_0$ *être allumée*, *bougie* (*Caus}_0\text{Fact}_0* *allumer*, $\text{Fact}_0$ *brûler*, *cigarette* (*Caus}_0\text{Fact}_0* *allumer*, $\text{Fact}_0$ *se consumer*, *désir* (*Caus}_0\text{Fact}_0* *satisfaire*, $\text{Fact}_0$ *se réaliser*).
Domain of lexical functions $\text{Fact}_0$, $\text{CausFact}_0$ and $\text{Real}_1$

As it is shown, it is not clear that a keyword of domaine like $\text{se réaliser}$ can take $\text{satisfaire}$ as a value, and so on. However, the empirical data in $\text{Dicouèbe}$ and $\text{CALLEX}$ point out that the nature of $\text{IncepFact}_0$ and $\text{CausFact}_0$ could be different, so we decided to presuppose that the existence of $\text{Caus(i)Fact}_0$ implies the existence of $\text{Fact}_0$.

5 The prediction of the domains for $\text{Fact}_0$, $\text{Caus(i)Fact}_0$ and $\text{Real}_1$

As shown in section 2, it should be possible to predict the domains of $\text{Fact}_0$, $\text{Caus(i)Fact}_0$ and $\text{Real}_1$ because of their meaning. The predictions of the domains of $\text{Fact}_0$ can be reached looking for keywords which can answer to what $L$ is useful for? Similarly the domain of $\text{Real}_1$ would include those keywords which answer the question what do we do with $L$? Consequently, the questions that would be helpful for prediction are: what things are useful for something? ($\text{Fact}_0$); and what things exist that we can do something with them? ($\text{Real}_1$).

Spanish examples (10) to (14) show some questions and responses which are useful for predicting those domains. Each question is formulated looking for one collocation for every keyword. As examples (12-13, 20-21) show, the more natural question allows us to deduce the collocation for each keyword; example (12) shows that even for some conceptual relations, such as clothes-to dress, there is not necessarily a corresponding linguistic expression ($? \text{the clothes dress}$ is not said), so we must formulate a different question, as in (13), that will be attached to another LF.

(10) ¿Para qué sirve la tetera? Para hervir agua. La tetera hierve el agua [What is a kettle used for? It is used for boiling water. A kettle boils water]

(11) ¿Para qué sirve una llave? Para abrir/cerrar algo. La llave abre/cierra la cerradura de la puerta. Giramos una llave. Abrimos/cerramos la puerta con la llave. [What is a key used for? It is for opening/closing something. The key opens/closes the door. We turn the key. We open/close the door with the key]

(12) ¿Para qué sirve la ropa? Para vestir? La ropa viste [What are clothes used for? For dressing? ?The clothes dresses]

(13) ¿Qué hacemos con la ropa? Vestirnos: nos ponemos la ropa [What do we do with the clothes? We wear the clothes]
Spanish collocations (14) to (21) show the LFs appropriate for examples (10) to (13). We have considered as a prototypical collocation the one which answers the appropriate question (10, 11 and 13). We have encoded this collocation using Fact\textsubscript{a/1} for those answers corresponding to what \textit{L} is useful for? and Real\textsubscript{1} for those answers to what do we do with \textit{L}? Like in Dicouèbe (see note 5 Section 4), only keywords of Fact\textsubscript{a/1} are keywords of CausFact\textsubscript{0/1} (14-17); the other relations are covered by Real\textsubscript{1} and IncepReal\textsubscript{1} (18-21). The word key (16-19) appears in four collocations encoded in a different way; examples (18-19) show that even if there exists one value of Fact\textsubscript{a/1}, such as to open, and there is another one for Real\textsubscript{1}, such as to turn, the verb to put (the key into) must be encoded as a previous phase of to turn (IncepReal\textsubscript{1}):

(14) Fact\textsubscript{1}(kettle) = to boil (the kettle boils the water)

(15) CausFact\textsubscript{0}(kettle) = to light the kettle (someone lights the kettle)

(16) Labreal\textsubscript{1}(key) = to open/close something with the key (someone opens/closes something with the key)

(17) Fact\textsubscript{1}(key) = to open (the key opens something)

(18) IncepReal\textsubscript{1}(key) = to put the key into something (someone puts the key into something)

(19) Real\textsubscript{1}(key) = to turn

(20) Real\textsubscript{1}(clothes) = to wear the clothes (someone wears the clothes)

(21) IncepReal\textsubscript{1}(clothes) = to put the clothes on (someone puts the clothes on)

Groups of LFs as (17-19) are also present in other collocations such as to sit down on a bus (IncepReal\textsubscript{1}), to go by bus (Real\textsubscript{1}), to drive a bus (CausFact\textsubscript{0}) and the bus goes/walks (Fact\textsubscript{0}), (Apresjan: 2007). Almost all the collocations encoded for CALLEX-ESP present a similar encoding to the equivalent entries in Dicouèbe (only some complex lexemes as in examples (1-5) have been encoded taking into account the conditions presented in section 2). However the word freno (brake, see examples 8-9, to put on the brake, to brake) has been encoded by IncepReal\textsubscript{1} and CausFact\textsubscript{0}. This is a real case of overlapping, because the hypothetical Spanish collocation corresponding to Fact\textsubscript{0} (el freno frena, the brake brakes), although redundant, would be possible. The word pipe (encoded by three equivalent formulas in Dicouèbe, Caus\textsubscript{1}Fact\textsubscript{0}, Prepar\textsubscript{1}Fact\textsubscript{0} and IncepReal\textsubscript{1}) is complex: following CALLEX (where smoke a cigarette is encoded by Real\textsubscript{1} and light a cigarette by IncepReal\textsubscript{1}) light a pipe could also be covered by IncepReal\textsubscript{1}. But Spanish words cigarillo and pipa belong to the Fact\textsubscript{0} domain, consumirse un cigarillo/ una pipa (to be consumed). Therefore, it was considered another overlapping case: both LFs (IncepReal\textsubscript{1} and CausFact\textsubscript{0}) are valid.

6 Results and conclusions

When looked on closely, the Real and Fact keywords in the CALLEX system seem to be applied to two distinct groups of ‘artifacts’: those requiring human intervention to begin working (radio, washing machine, etc.); and those that require human intervention in order to continue working (plane, university, door, etc). In CALLEX-ESP the former corresponds to the semantic labels such as ‘aparato’ (apparatus), ‘dispositivo de almacenamiento’ (storage device), and the like. All of them would be related to CausIncepFact\textsubscript{0} by glosses such as
encender (to turn on) or introducir (to introduce). The latter corresponds to semantic labels such as ‘medio de transporte’ (transport system), ‘entidades’ (entities), among others. They are related to the LF CausFact<sub>0</sub> by verbs as to drive (a car), to pilot (a plane), to open (a university), to close (the door).

So far, a total of 41 lexical units have received the glosses of Caus(Incep)Fact<sub>0</sub> attached to each semantic label. After adding all the values, 90 collocations corresponding to this LF were included in CALLEX-ESP. On the other hand, 473 collocations have been coded as IncepReal<sub>1</sub> and 2371 as Real<sub>1</sub>. If Caus(Incep)Fact<sub>0</sub> is finally applied, this LF could be related to other semantic labels (always ‘artifacts’ in the sense of Apresjan 2001) and cover other collocations such as (someone) starts the car.

In conclusion, semantic definitions and glosses of Fact<sub>0</sub>, CausFact<sub>0</sub> and Real<sub>1</sub> were extremely useful in order to predict their domain and solve some problems concerning the links between CausFact<sub>0</sub> and IncepReal<sub>1</sub>. The meaning ‘destination’ or ‘normal expectations in the respective situation’ is an essential requirement for the keywords (but not for the notation of LFs) of Fact<sub>0</sub>, IncepFact<sub>0</sub>, Caus<sub>0</sub>Fact<sub>0</sub> and Real<sub>1</sub> domains. Consequently, we claim that keywords belonging to some lexical fields as in the case of animals cannot be included inside this domain. We have considered the domain of Caus<sub>0</sub>(Incep)Fact<sub>0</sub> as a subdomain of Fact<sub>0</sub>, and we have formulated Caus<sub>0</sub>IncepFact<sub>0</sub> semantic definition as ‘to cause X to start to function according to its destination’. So far Caus(Incep)Fact<sub>0</sub> has been used for the Spanish keywords grouped around semantic labels such as ‘apparatus’, ‘storage device’, CausFact<sub>0</sub> for ‘transport system’, ‘entities’, among others, and Real<sub>1</sub> or IncepReal<sub>1</sub> for several other semantic labels.

Generally speaking, it is possible to state that the study of the domains of LFs by means of semantic labels and their relations with LFs glosses and semantic definitions is very helpful in order to specify the links to some other LFs and to predict collocations semantically motivated. However, the problem claims a deeper study that lack of space makes it impossible to undertake now.

Acknowledgements

I would like to show my gratitude to Ignacio Bosque, Sylvain Kahane and Inmaculada Álvarez de Mon for their comments on the second version of this paper; to Igor Boguslavsky for his comments on the first version and his support on CALLEX-ESP tasks; to Juri Apresjan because if he had not allowed me to participate in the Spanish version of CALLEX, this paper would not have been written; to Igor Melčuk for his generosity in answering my questions; and especially to Alain Polguère for all his suggestions about glosses and semantic labels. I apologize for the mistakes that the reader can find: I am the only one responsible for them.

Bibliography


