Framenet as a corpus tool for the learning of second languages and for the lexical awareness of one´s first language

OLGA BLANCO CARRIÓN
Departamento de Filologías Inglesa y Alemana
Universidad de Córdoba

ABSTRACT: In this paper, I aim to show how the FrameNet database created by the research group lead by Prof. Charles Fillmore and based on the principles of Frame Semantics, can serve as a valuable tool for the learning of second
language lexis as well as for the better understanding and/or acquisition of the learner's first language lexicon.
Keywords: Frame Semantics, FrameNet, frame, frame elements, conceptualization, culture.

RESUMEN: En este artículo se pretende mostrar como la base de datos FrameNet creada por el grupo de investigación dirigido por el profesor Charles Fillmore y basada en los principios de la Semántica de Marcos puede ser una herramienta útil para el aprendizaje del lexicon de una segunda lengua así como para la mejor comprensión de la lengua madre.
Palabras clave: Semántica de Marcos, FrameNet, marco, elementos del marco, conceptualización, cultura.

1. INTRODUCTION

I have chosen FrameNet as point of departure, not only because its approach is innovative but also because it entrenches with an approach to the study of language, called Frame Semantics, which turns out to be extremely beneficial for the learning of a foreign language. Frame Semantics is, in a way, a development of Fillmore´s Case Grammar (1968). Frame Semantics holds that the lexicon and grammar of a language are not separate fields of study but are interdependent. One cannot study grammar without taking into account meaning, and one cannot study the lexicon of a language without taking into account how words behave within a syntactic context. Of course, this aspect has been strongly discussed in second language teaching. I argue that it is also beneficiary for mastering one´s own mother tongue.

For any approach on language learning, there are some factors that function as prerequisites for the successful language acquisition to take place. The factors presented here are designed along the lines of those developed by Celce-Murcia (1985) in relation to the importance of focus on form in second language teaching. I believe that they are equally applicable to an

approach which focuses on meaning. In the present paper, meaning is to be understood not only as the semantic 'load' of a lexical item, but also in terms of its relation to its conceptual analogue as well. The aforementioned pre-requisites are as follows:

(i) Age: Adult learners. The information which is the source material for the study of a language is intended for adult learners of that language.
(ii) Proficiency level: the information presented in the FrameNet database is too complex for a learner in an early stage of development in his own language or in the foreign language. Thus, the approach to language learning based on Frame Semantics seems to be adequate for someone who has developed an adequate cognitive semantic web of concepts and lexical items and who is fairly active in the usage of that web at least in his mother tongue. This way, the learner could, by comparison, learn about the concepts and lexical items of the foreign language and take full advantage of the information presented by the FrameNet database.
(iii) Learning style: The aforementioned level of conceptual development runs parallel to someone with an analytic capability. The information in the database requires some form of introduction on the part of a tutor or teacher in order for the learner to understand its functioning. Once this has been achieved, the learner should be prepared for independent learning. In order to make this possible, he must possess some analytic skills so that he can extract the information himself from the database. The data are presented in terms of lexical entries (as in dictionaries) and in terms of frames (which consist of an interrelated set of concepts present in the minds of the native speakers of the foreign language).
(iv) Educational background: The ideal learner for this linguistic resource would be literate or well-educated so that he can utilize all the types of information contained in the database.

2. FRAME SEMANTICS AND LANGUAGE LEARNING

Frame Semantics as an approach to language and, in this paper, to language learning aims at always taking into account what may be called the cognitive-semantic-cultural frame of a lexical item. By cognitive, I mean that lexical items (be it single words, compounds, idioms, constructions) are treated as the linguistic constructs which refer to some concept the speaker of a language has in his mind. Therefore, it would be reasonable to defend the view of Frame Semantics as an interesting tool not only when studying a second language, but also when discovering one's own language; and, this way, what concepts or pre-conceptual structures language users have in mind, consciously or subconsciously, when they think and express thoughts in their mother tongue. This may be considered the main aspect which makes this approach original as well as effective when one is learning a foreign language or delving into one's own language.

The concept of frame as defined above is related to the issue of categorization. It is well known that concepts are grounded in experience (be it bodily-physical experience and/or socio-cultural experience). For instance, let us take the classic example "bachelor" (Fillmore, 1982). This word does not have a simple meaning but it consists of a net of interrelated

meanings, because the definition of "bachelor" cannot be shown from a limited set of characteristics. Thus, the definition of *bachelor* includes many of the following aspects:

A bachelor is:

– an adult male
– someone who lives in single hood, that is, someone who is not married.
– someone who has never married
– someone who can marry but decides not to (i.e. a priest cannot be said to be single)
– etc.

These aspects of the definition show, in a way, how a definition may turn to be complex and how it cannot be reduced to what the checklist theories of meaning argued (Fillmore, 1975), i.e. a neat and complete set of characteristics. However, this approach as such is not new. It was first used by some psychologists and cognitive scientists that studied the issue of the perception of colours (Berlin and Kay, 1969), and which in a sense coincides with the origin of what has been called "cognitive linguistics".

As concepts, frames are human created structures organized around human experience. So, it makes sense to argue that the teaching of a second language should take into account that the activation of cognitive structures occurs from the use of linguistic forms which are not exactly equivalent from language to language and which may lead to the common misunderstandings when one tries to analyze the foreign language (and by extension its culture) from the point of view of a different language (e.g. one's mother tongue). This is when the knowledge about foreign language frames becomes useful.

By semantic, I mean the terms used to code the cognitive reality the speakers of a language have adopted as relevant for their functioning in the world. That is, while some realities are apparently relevant for some speech communities, they are of no use for some others (Lakoff, 1987). This aspect is especially advantageous when studying languages which belong to a culture far removed from one´s own, and turns out to be essential to understand the system of communication of the foreign speech community and the foreign language.

By cultural, I refer to the frames which do not refer to the external realities out there in the world but to realities which have been created by society (norms, rules, constructs) and which differ from culture to culture. Hence, taking the point of view of Frame Semantics, and Framenet, as a tool for learning the intricacies typical of a certain culture seems reasonable as the database provides most of such information (although it is still work in progress). In fact, it is also widely known that in order to master a foreign language one has to be aware of the culture that, in some way or another, governs the behaviour of the speech community. In this sense, one understands language as well of other types of interaction as part of that behaviour.

In short, learning a foreign language integrates the idea of knowing and/or learning about its culture. In the present paper, I am focusing on English language and culture but, as I previously pointed out, this approach is more useful the further the foreign culture is from one´s own culture.

3. FRAMENET: DATA PRESENTATION

As it can be observed the features which are considered determinant for foreign language learning and for language understanding are not independent; they all belong to a complex structure which is introduced above under the name of cognitive-semantic-cultural frame. It is true, however, that one of the aspects might be more prominent in a certain linguistic item than another- be it a single word or a construction[1]. But if we consider the example "bachelor" again, it seems clear that this word is relevant for a culture in which the concept of "marriage" exists, and where there is a difference between celibacy and being a widow or divorced.

In addition, every lexical item codes a complex frame which the speaker of a language has in mind when he uses the term and which is not always completely explicit in the chosen linguistic expression. However, the awareness of the existence of a frame as well as the knowledge about its components and characteristics helps the learners of a language to understand the cultural system encrypted, so to speak, in the lexical item and the concept the native speakers of a language have in their minds when using that term.

For a better understanding of the approach to language learning presented above, a practical instance of a frame is going to be shown along the following lines. To begin with, I shall point out that the FrameNet database (henceforth FN), is a lexicographic tool which can be used as a dictionary and as a thesaurus. When used as a dictionary, it contains several types of information, which are to be introduced below. When viewed as a thesaurus, it relates some lexical items to the conceptually related lexical items, which as expected belong to the same frame. There are also relationships among different frames.

As an illustration, I have chosen the frame Grinding[2], which in the FN database is defined as follows: "In this frame a Grinder or Grinding cause causes an Undergoer to be broken into smaller pieces. A Result or Goal can be present." The items introduced in capitals are labels for what the Frame Semantics approach calls Frame Elements. These are the cognitive semantic entities that are present in a frame (mind of speakers) as part of the frame-scenario (real situation in which the lexical item is used). These Frame Elements may be syntactically realized within the clause or omitted, although present in the minds of the speakers of a linguistic society or culture. For instance[3]:

(1) Crush the biscuits to bits.
(2) Crush the biscuits with a rolling pin.

In example (1) the Frame elements which are specified are the Undergoer (the biscuits) and Result (to bits), whereas in example (2) the Frame elements are Undergoer (the biscuits)

---

[1] Since single words, compounds, idioms and constructions have been implicitly dealt with as elements of the same nature, I shall clarify this issue. Each linguistic resource is considered to be of the same nature and value no matter how complex it may appear to be to the speaker or learner's eye. Its existence is motivated by a linguistic need and, in the present paper, and in the opinion of Construction Grammarians (Fillmore and Kay, 1995; Golberg, 1995) they should be given the same importance, not only regarding their complexity but also regarding whether they have been considered by traditional grammars as central or peripheral instances of language.

[2] http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper& Itemid=118&frame= Grinding&

[3] Examples taken from the FN database.

and Instrument (with a rolling pin). Nevertheless, we know that in sentence (1) there must be an Instrument, or a Manner to crush the biscuits although it is not specified; and similarly, in sentence (2) we know the biscuits end up as something different from what they originally were although this "resulting state" is not specified. The frame elements in both sentences tend to coincide or, at least, there are the same potential frame elements although they do not appear in every situation or utterance. It depends on the speaker to specify them or not, and this decision to specify some elements or not is usually due to an information need or prominence reasons. However, a native speaker of English knows of the existence of all those potential frame elements, because he is aware and knows the frame Grinding.

There are two major types of Frame Elements for every frame. Core Frame Elements are those which are specific to the frame in question whereas non-core are the frame elements which may appear as part of that frame but which may also occur in other frames. The core Frame Elements of the Grinding frame are Grinder or Grinder cause, which are mutually exclusive, and Undergoer. In other words, this frame codes a prototypical transitive state of affairs. The definition of these core Frame Elements is as follows[4]:

Grinder: The Grinder breaks the Undergoer into smaller pieces.
Grinding cause: An animate or inanimate entity, a force, or event that grinds. Volitionality is not a necessary characteristic.
Undergoer: The Undergoer is the entity which undergoes the change brought about by the Grinder or Grinding_cause.

One may wonder how this approach to language study differs from a semantic one, such as Dik´s Functional Grammar. The difference lies in the fact that the frame elements are of a purely conceptual nature whereas in a functional approach the participants of the sentence are semantic roles or functions which do not refer to a cognitive correlate. That is, a functional approach considers language as a system of communication but not as a system of reflection or conceptualization.

Among the non-core frame elements for the Grinding frame there are the following:

Duration: The length of time for which the Grinder grinds the Undergoer.
Goal: This is where the Undergoer might end up, after undergoing the grinding process.
Instrument: This FE identifies the Instrument with which the Grinder grinds the Undergoer.
Locus: The Locus is the locus of the grinding action.
Manner: This FE identifies the Manner in which the Grinder grinds the Undergoer.
Means: This FE identifies the Means by which the Grinder grinds the Undergoer.
Place: This FE identifies the Place where the Grinder grinds the Undergoer.
Purpose: This FE identifies the Purpose for which the Grinder grinds the Undergoer.
Result: The Result is the shape the Undergoer becomes.
Time: This FE identifies the Time when the Grinder grinds the Undergoer.

The FN database includes a set of annotated sentences extracted from the British National Corpus, which helps the learner understand the meanings of the Frame Elements just mentioned

[4] Taken from the FN database.

as well as learn some of their syntactic realizations. To illustrate this point, a selection of annotated sentences is presented below:

1. CRUSH [the amaretti biscuits or almond macaroons]Undergoer [with a rolling pin]Instrument and mix with the ricotta cheese, chopped apricots, Sweetex Granulated, Cointreau or Grand Mariner and ground cinnamon.

2. [Finely]Result CRUSH [the biscuits]Undergoer .

3. Of a sudden Lexandro's free hand gripped Valence 's wrist with [a power which]Grinding cause would have CRUSHED [any ordinary bones] Undergoer .

4. A farmer is using a shire horse to pull [the millstone which]Grinding cause CRUSHES [the fruit]Undergoer

5. [I]Grinder CRUSHED [some of the flowers]Undergoer and rubbed them on my forehead as I 'm told it has an invigorating effect on cabbages and thought it might do the same for me .

6. The lorries came and went on a service road near the town; on some days [the bulldozer]Grinding cause clanked back and forth, CRUSHING [the rubbish]Undergoer [into the soft, orange earth]Goal .

7. [I]Grinder tried to CRUSH up [a salt tablet]Undergoer [in a cup]Goal and dissolve it in the water but she moaned when I held the cup to her lips and tried to turn her head away.

8. [Jack]Grinder plucked the flower out of his morning coat and closed his fist slowly over it, CRUSHING [it]Undergoer [into a pulp]Result.

9. CRUSH [them]Undergoer [to a powder]Result and boil them, then soak the ankle when the water 's cooled enough to bear.

10. They knew that [anything, even Alice]Grinding cause, could advance over them like bulldozers, CRUSH [them]Undergoer [to bits]Result.

11. [He]Grinder spoiled twice as much fruit as he ate by CRUSHING [the berries]Undergoer [beneath his clumsy feet]Place in a mad rush from bush to bush.

12. [Children]Grinder can either take remedies as a powder by CRUSHING [the pill]Undergoer [between two clean teaspoons]Place, or dissolved in a little clean, fresh water in a glass.

13. The Mickey Mouse girl would be sweeping it all into the bin, along with [any cockroaches]Undergoer [she]Grinder 'd CRUSHED [underfoot]Place.

14. Fresh or frozen prawns and cockles are good — even [prawns in their shells, which]Undergoer are CRUSHED [by the throat teeth]Grinding cause and provide roughage.

15. The noise of splintering wood became more commonplace than the singing of birds as [saplings and lesser trees]Undergoer were CRUSHED [beneath the weight of falling forest giants]Place.

16. She looked across the stream, through the leaves at the distant field; at the nettles and the meadow-sweet and the wild roses ; down at [the camomile daisies]Undergoer CRUSHED [under her feet]Place.

In the sentences above only the non core frame elements and Goal, Result, Place and Instrument appear. However, the others may appear in the other lemmas or lexical units which belong to the frame Grinding and which are specified in the database. This way, the learner finds the information not only defined but also shown in examples in which he can

judge for himself the more common syntactic realizations of the different frame elements. For instance, Instrument is introduced by a prepositional phrase with the preposition "with"; Goal is introduced by a prepositional phrase with the prepositions "into" and "in"; Result is expressed by means of a prepositional phrase with "into", "to"; and finally Place is expressed by means of noun, prepositional or adverbial phrases.

The learner of English can learn the lexical items from real and commonly used sentences extracted from a corpus. The observation of a large set of examples, in my opinion, may help him to gain a better idea of the meaning of that lexical item for English speakers as well as in which contexts that lexical item is used. In addition to this observation task on the part the learner, the teacher may help the learner to focus on the certain aspects by giving him the following guidelines for search:

1) What kind of substances codes the Undergoer. What do they have in common? Are there any collocations shown in the examples?
2) What is the difference between Grinder and Grinding cause? Is it a question of animacy or of volition?
3) Can you foresee the kind of frame element called Result if you focus on the preposition chosen to introduce it? What prepositions introduce this frame element? What noun phrases follow those prepositions? Do the noun phrases following the prepositions have anything in common?
4) If you have to rank the most important non core frame elements which order would you follow. Which ones are more specific of this frame and which ones less specific?
5) Do the non-core frame elements introduce any piece of meaning that help you understand better the meaning of the lexical item "crush"?
6) Search for common issues in other lemmas that belong to the frame "Grinding".
7) Search for differences between "crush" and other lemmas in the frame "Grinding"; etc.

The database also provides us with a definition of each lemma taken from the Collins dictionary[5] and shows the Frame Elements for the word sense under study in a table with their syntactic realizations.

Table 1. Core Frame Elements and their syntactic realizations

Frame Element	Realizations(s)
Grinder	CNI. DNI. NP.Ext
Grinding_cause	NP.Ext PP[beneath].Dep PP[by].Dep
Undergoer	CNI. NP.Obj NP.Ext

[5] Crush.v: pulverize by compressing forcefully.

The FN project only considers three types of grammatical functions which in turn facilitate the learning process for a learner of English who is not so much interested in acquiring a detailed account of English grammar, but needs to learn enough to be efficient in communication and to understand the functioning of the language. The grammatical functions are marked as the extension of each of the syntactic realizations of the frame elements. They are *external*, which refers to any phrase which is outside the main verb phrase; *object*, which corresponds to the traditional idea of object; and *dependent* which is assigned to adverbs, prepositional phrases, verb phrases and clauses which are part of the main verb phrase. It includes the idea of argument and adjunct[6].

With regards to the syntactic realizations, CNI stands for Constructional Null Instantiation and it refers to those cases in which an element in the clause is not syntactically realized because of the type of construction it belongs to (e.g. the subject of imperative verbs.) DNI, on the other hand, stands for Definite Null Instantiation, which refers to those elements which are omitted because they code "something that is already understood in the linguistic or discourse context."[7] Sentences 2 and 16 from the annotated sentences shown above exemplify these two cases:

2. Finely CRUSH the biscuits (Grinder, CNI)

16. She looked across the stream, through the leaves at the distant field; at the nettles and the meadow-sweet and the wild roses; down at the camomile daisies CRUSHED under her feet. (Grinder, DNI)

The database also contains information about the semantic-syntactic valences of the lexical item under study. That is, the different patterns showing the distribution of the core frame elements and their syntactic realizations for the chosen lemma. Although this table is really valuable for the most advanced learners, I personally find the annotated sentences much more helpful as they provide an inductive learning of those patterns through examples. However, both pieces of information are non exclusive but complementary and the learner can choose whichever best suits his needs or enquiries about a particular aspect of the language.

---

[6] For more information on this matter see the FrameNet book, available at http://framenet.icsi.berkeley.edu/index.php?option=com_wrapper&Itemid=126

[7] From the FrameNet book, see footnote 5.

Table 2. Valence Patterns

Patterns	
Grinder	Undergoer
CNI	CNI
CNI	NP.Obj
DNI	NP.Ext
NP.Ext	NP.Obj
Grinding_cause	Undergoer
NP.Ext	NP.Obj
PP[beneath].Dep	NP.Ext
PP[by].Dep	NP.Ext

Therefore, in the process of learning the English lexicon, the learner will be enriched by looking up frames and/or lexical entries in the FN database as the database not only contains the definition of a lexical item but also the frame in which that item is inserted and the frame elements conceptualized by the native speakers of English for that frame. This helps the learner to gain a better understanding of the native speaker's conceptualization of a certain reality.

4. CONCLUSION

In this paper, I have presented an approach to language learning based on the theory of Frame Semantics and on its applied project FrameNet. Both the theoretical background and the lexical tool are shown to be of advantageous use when applied to foreign language learning and understanding one's own language. It is shown that learning a language implies much more than learning the system of communication and that aspects such as a reflection about the mode of conceptualization of the foreign language community, the culture and social constructs created by that community are key factors to be taken into account.

Although the approach is mostly useful for advanced learners I also claim that the input from a teacher or tutor is essential when working with an online tool like FrameNet. The degree of dependence from the tutor, depends on the factors mentioned in the introduction such as literacy level of the learner.

All in all, the learner who uses this approach to learning a foreign language can only benefit from such a detailed presentation of real data and in such a clear and accessible manner.

5. REFERENCES

Berlin, B. and Kay, P. (1969). *Basic Color Terms. Their Universality and Evolution.* Berkeley, Los Angeles: University of California Press.

Celce-Murcia, M. (1985). *Making informed decisions about the role of grammar in language teaching*. TESOL Newsletter 19/1, 4-5.

Dik, S. C. (1997). *The Theory of Functional Grammar: The Structure of the Clause*. Part I. Berlin, New York: Mouton the Gruyter.

Fillmore, C. J. (1968). "The case for case", in Bach and Harms (ed.), *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1-88.

Fillmore, C. J. (1975): "An Alternative to Checklist Theories of Meaning", in *Proceedings of the First Annual Meeting of the Berkeley Linguistic Society*. Berkeley, California: Berkeley Linguistics Society, 123-131.

Fillmore, C. J. (1982). "Frame Semantics", in Linguistic Society of Korea (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin, 111-138.

Fillmore, C. J. (1985). "Frames and the Semantics of Understanding", in *Quaderni di Semantica* 6-2: 222-254.

Fillmore, C. J. & Paul Kay (1995). *Construction Grammar*. CSLI lecture notes. Leland Stanford Junior University.

FrameNet Database http://framenet.icsi.berkeley.edu/

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Lakoff, G. (1987). *Women, Fire and Dangerous Things. What Categories Reveal About the Mind.* Chicago: Chicago University Press.