MAPPING MEANING ONTO FORM: A CORPUS-BASED CONTRASTIVE STUDY OF NOMINAL MODIFICATION IN ENGLISH AND SPANISH

Noelia Ramón
University of León, Spain

Abstract
This paper is a corpus-based contrastive study of one particular semantic area – the modification of nouns – and the grammatical structures available in English and Spanish to actualise this particular meaning within the scope of the NP. Two large monolingual reference corpora, one in each language, have been used to extract the empirical data on which to base the analysis. A functional approach has been followed, starting from perceived similarities in meaning and trying to reveal the subtle links between particular semantic functions and the structures most commonly employed by native speakers of the two languages to convey those semantic functions. The working procedure will be described in detail, from the data selection process, to the description, juxtaposition and final contrast. This paper presents part of the contrastive results found in a larger study (Ramón 2003), and show the appropriateness of the methodology employed in order to obtain contrastive data useful in applied areas such as foreign language teaching, and, especially, in the field of translation.

Keywords: contrastive analysis, nominal modification, English-Spanish translation.

1. INTRODUCTION

The main aim of any theory of language is to describe how meaning is mapped onto formal structures, in an attempt to account for the use of human language in general. This paper uses real language data from computerised corpora in order to reveal how one particular domain of meaning – the modification of nouns - is mapped onto two different languages, English and Spanish. This particular semantic area was chosen because it is the source of numerous problems in cross-linguistic activities involving the two languages, mainly foreign language learning and translation. The reason is that nominal modification is patterned in a very different way in English and Spanish, mainly for typological reasons. We understand by modification, or characterisation, (both terms will be used here as synonyms), the means available in a language to describe nouns, on the one hand, and verbs, on the other hand (Darbelnet 1969:41). This study will focus exclusively on the modification of nouns within the scope of the noun phrase (NP).

A wide range of different modifying semantic functions may be active in the NP, and characterising relationships may be conveyed by a wide range of linguistic items in attributive position, comprising adjectives and adjectival phrases, prepositional phrases (PPs), relative clauses, etc. Bearing in mind the similarity of formal means employed by the two languages to actualise characterising functions, the main working hypothesis is that the overall distribution of these grammatical resources differs greatly when considered from the perspective of meaning and use. This paper presents the results of an English-Spanish corpus-based contrastive study of nominal characterisation, which set out to establish the functional and semantic patterns that govern the distribution and realisation of this particular function in both languages. The findings are mainly geared towards their application in the field of translation.

This is a descriptive study that follows a functional approach to language (Bondarko 1991) and is based on data extracted from two large monolingual reference corpora: the CobuildDirect corpus for the English language and the CREA (Contemporary Spanish Reference Corpus) for the Spanish language[1]. Starting from a particular semantic area, nominal modification, the formal structures used for actualising this meaning will be identified and classified semantically. The resulting semantic subcategories will subsequently be contrasted in the two languages in order to

reveal the subtle relationships that exist between meaning and form. The results for the four most common functions identified in our corpus – description, classification, time and place - will be presented in order to prove the viability of the research method employed. The final aim is to be able to provide functional-semantic correspondences in the two languages that could be viewed as potential translation options available for the professional translator to choose from. Not only is the list of options provided, but also statistical information about the typicality of occurrence of the different structures as found in our two corpora. The fact that two monolingual corpora were used for the study makes the results highly reliable, as they are based on authentic instances of language in use produced by native speakers of English and Spanish in a variety of communicative settings.

2. THEORETICAL BACKGROUND

Contrasting languages from a strictly linguistic and scientific point of view started in the second half of the 20th century with the development of contrastive linguistics, which "could at present be broadly defined as the systematic synchronic study of similarities and differences in the structure and use of two or more language varieties, carried out for theoretical or practical purposes" (Bugarski 1991:77). Contrastive studies may be carried out at any level of language, from phonology to discourse, and contrastive data have traditionally been considered useful for cross-linguistic activities such as language teaching and translation, thus making the discipline part of an applied branch of language studies. In more recent times, however, contrastive linguists have tried to provide explanations for the differences found in order to achieve greater theoretical relevance.

Contrastive linguistics is not a theoretical approach to language in itself, nor does it deal with a particular area of language, but may deal with any of them, as mentioned above. Contrastive linguistics is a methodological approach for systematically comparing two or more different languages. Consequently, a particular contrastive study must necessarily draw on some theoretical model. Language contrast has attracted scholars from many different linguistic backgrounds. Since the comparison of two languages inevitably includes a descriptive part, any theoretical model may *a priori* be used in carrying out a cross-linguistic study. From the mid-20th century onwards, contrastive studies have been carried out following structuralist approaches (Fries 1945, Lado 1957), case grammar approaches (Di Pietro 1971), generative approaches (Krzeszowski 1979, 1990), various types of functional and translation-oriented approaches (Ivir 1981; Rabadán 1991; Chesterman 1998), corpus-based approaches (Aijmer et al. 1996; Jaszczolt and Turner 1996a, 1996b; Johansson 1998; Johansson and Oksefjell 1998; Altenberg and Granger 2002), and combinations of these, among others.

Whatever the perspective, all these models seek to account for how similar meanings are actually realised by formally different means in two or more languages. One of the essential concepts related to contrastive studies is the *tertium comparationis*, that is, the criterion for comparison. Structural approaches relied heavily on formal aspects in defining their *tertium comparationis*, whereas in the case of generative approaches identical deep structures were postulated for formally differing sentences in two different languages. It was soon noticed, though, that the type of *tertium comparationis* needed would have to account for all the levels of interlinguistic shared ground, and translation equivalence was considered to be an appropriate candidate (Halliday et al. 1964; James 1980; Ivir 1981). The relationship between translation and contrastive studies is one of mutual benefits (Ramón 2002), as translation is often used

as input for cross-linguistic studies, and on the the other hand, the results of such a study can be fruitfully applied in the field of translation, as we will argue throughout this paper. Finally, more recently Chesterman suggests that "the starting point for a given CFA-type [Contrastive Functional Analysis] is a perception, made by a linguist, a translator, a language learner. This is a perception of a similarity of some kind." (Chesterman 1998:55). Chesterman's functional approach is the model that will be followed in this paper.

I will not deal with all of these models in detail here, but will refer exclusively to some of the theoretical tenets this paper draws on, which combine functional and corpus-related principles. From a linguistic perspective, the contrastive study presented here is of the functional type, mainly following Halliday (1994), as it highlights the communicative function of language and considers that the form a language takes is a result of the use speakers make of it in particular contexts. Particularly interesting for our purposes is the notion of a single level of language, which Halliday calls 'lexicogrammar'. The meaning of words cannot be separated from the grammatical constructions in which they occur. The grammar of a language "can be represented paradigmatically as a network of choices, each choice consisting of a small number of alternatives related by probability; these probabilities appear to form a pattern related to the construing of information." (Halliday 1991:41). The concept of paradigmatic choice in language is one that will turn out to be crucial in presenting our contrastive results, as we aim at identifying linguistic options that could be considered interlinguistic correspondences.

Descriptive corpus-based approaches to language provide the frequency of occurrence of particular linguistic structures, thus revealing empirical data about the paradigmatic organization of language and about the extent to which context is reflected in language. In other words, corpus-based studies have confirmed many of the hypotheses put forward by functional approaches to language. Language has been shown not to be randomly organised at all, but structured in patterns, chunks and bundles of varying types. This underlying organisation of language often remains unnoticed by native speakers, but corpus-based analyses can easily reveal it:

> The majority of text is made of the occurrence of common words in common patterns, or in slight variants of those common patterns. Most everyday words do not have independent meaning, or meanings, but are components of a rich repertoire of multi-word patterns that make up text. This is totally obscured by the procedures of conventional grammar. (Sinclair 1991:108).

In the case of contrastive studies in particular, corpora have been used from the early 1970s: "No contrastive project can be regarded as complete before its results are verified and completed by means of a corpus." (Filipovič 1974:62), and the corpus-based approach is often considered to be responsible for the revival of contrastive linguistics in more recent times. Different types of corpora may be employed to carry out this sort of research and section 3.1 provides a detailed description of the corpora used for this paper.

Another functional description of language that has been used for the present study is the one put forward by Bondarko (1991). This Russian author has developed the concept of 'functional-semantic field', which is a very interesting way to describe the relationship between syntax and semantics. Every functional-semantic field is formed by one particular meaning plus the different structural resources employed in a particular language to actualise that meaning. These resources will not all have the same frequency of occurrence, but rather, native speakers will tend to identify each meaning

with one or two central formal structures, whereas the various other possibilities available will be more or less peripheral depending on their typicality for actualising that semantic function: "A functional-semantic field is a bilateral unity of form and content which embraces the concrete means of a given language with all the characteristics of their form and content." (Bondarko 1991:22). The concepts of 'central' versus 'peripheral' are often used in language studies and have also been applied to other disciplines such as the *polysystem theory* in descriptive translation studies (Toury 1995). Bondarko's model is meaning-based, and can be conveniently applied to cross-linguistic studies, since meaning can be considered a safe common ground between two languages when forms differ. This paper will try to sketch the functional-semantic fields for the various meanings included in the broad semantic area of nominal modification.

3. METHOD

Traditionally, a contrastive study is said to be structured in three different stages. There are "three steps in CS: description, juxtaposition and comparison." (Krzeszowski 1990:35). However, the approach suggested here is somewhat special in that it makes use of computer corpora, but has meaning as a starting point. Computer corpora are very useful resources for language description, but large reference corpora are not commonly tagged semantically, and formal patterns are needed to carry out a particular search. This peculiarity has made it necessary to devise a specific strategy to obtain corpus material that would be relevant for our purposes. An initial data selection phase is therefore included before the actual contrastive study can be carried out. On the other hand, I have also decided to include the traditionally termed 'juxtaposition' and 'comparison' stages under one single heading in the results section (see section 4 below) because of the close relationship between the two. This is also the approach taken in James (1980), where the author distinguishes only two steps – description and comparison – in any contrastive study.

3.1 Data selection process

The two corpora that were employed to extract the empirical data for this study are Cobuild Direct (a 56-million-word selection of the much larger Cobuild/Bank of English corpus[2]), and CREA (the 156-million-word Reference Corpus of Contemporary Spanish[3]). These two monolingual corpora were chosen among those available because of similarities in their internal structure, scope and wide representativeness of the respective languages in their contemporary stage. Several subcorpora within these two large reference corpora were selected in order to ensure comparability. The subcorpora used included only written texts (fiction, non-fiction, press and ephemera), texts produced between 1990 and 2000, and texts produced in the European varieties of these two world languages, i.e. British English and peninsular Spanish. No register differences are considered here, since the aim of the study is the description of the general language in the written mode. Comparable corpora "consist of original texts in each language, matched as far as possible in terms of text type, subject matter and communicative function" (Altenberg and Granger 2002:7-8). The two resulting subcorpora were similar, and, in a broad sense, 'comparable' in most aspects, including text mode, text type, date, origin, and size - slightly over 30 million words each. I acknowledge the fact that total comparability is difficult to achieve (Laviosa 1997), but the degree of comparability in this case was considered sufficient for our purposes. One factor "which speaks in favour of working with comparable corpora, rather than with parallel corpora, is the relative scarcity of parallel text material, which is hard to come

by, compared to monolingual corpus material in various languages." (Borin 2002:5). Although some small-scale parallel corpora of specialised texts do exist, no sufficiently large English-Spanish parallel corpus of the general language is available as yet to carry out contrastive studies of the type presented here.

Considering that this contrastive study has taken a semantic function as the starting point, and considering the fact that the two corpora employed for the study are totally unrelated to each other, a specific search strategy had to be devised in order to obtain relevant data for the analysis of nominal modification from the two corpora. The solution taken was the use of a number of very common nouns in the two languages as entries for our corpora and thus analyse their syntactic environment in the search for instances of modification. This option is supported by the consideration that "semantically, (...) the noun appears to play the leading role and the predication, whether adjectival or verbal, is subordinated to it." (Aarts and Calbert 1979:137). Frequency lists in the two languages were used in order to find the most common nouns in English and Spanish. The Cobuild corpus provides frequency lists of all parts of speech, and the ten most common nouns in English were selected for the study. The Spanish corpus CREA does not provide this type information, which had to be gathered from other corpus-based sources for this language (Alameda and Cuetos 1995). The ten most common nouns in Spanish were also chosen for the analysis, irrespective of the fact that not all of them were referential equivalents of the English nouns. There are various additional frequency lists available for both English (British National Corpus[4]) and Spanish (Juilland and Chang-Rodríguez 1964, Corpus del Español[5]) with slight differences in the ten most frequent nouns. The list obtained by Alameda and Cuetos (1995) was selected because it was based on a large corpus of mainly peninsular contemporary Spanish including a register distribution similar to the one present in CREA. The aim was not to carry out a lexical contrastive study, but rather reveal the links between syntax and semantics wherever a particular semantic function occurs, no matter what head noun is affected by the modification. As a result, the conclusions that may be derived from this analysis would not differ substantially if some of the nouns had been substituted by other very frequent nouns in the two working languages.

Due to the large size of our two comparable corpora – over 30 million words in each corpus – the number of instances of each of our head nouns was very large (between 20,000 and 50,000 occurrences), a number which is beyond the scope of this analysis. The following statistical formula was employed in order to reduce the number of concordance lines to a manageable sample and determine exactly how many instances of each noun were needed to allow statistical significance and representativeness of the whole sample:

$$n = \frac{N}{(N-1) \, E^2 + 1}$$

The element n is the final sample we will analyse, N is the whole sample of occurrences, and E is the estimated error, in this case 0.05 for a 95% confidence margin. A total of 7,882 concordance lines of the most common nouns in the two languages were collected: 3,939 in English and 3,943 in Spanish. Tables 1 and 2 show the exact number corresponding to each head noun.

Table 1. Data selection process of English head nouns.

Total number of occurrences	Sample analysed
time: 49,940	397
year: 44,769	396

world: 31,091	395
way: 26,279	394
day: 25,532	394
man: 23,330	393
home: 23,246	393
life: 22,428	393
night: 19,276	392
week: 18,888	392

Table 2. Data selection process of Spanish head nouns.

Total number of occurrences	Sample analysed
vez: 39,752	396
parte: 38,542	396
tiempo: 30,417	395
vida: 28,902	395
caso: 28,098	394
día: 27,487	394
año: 27,440	394
forma: 25,226	394
mundo: 23,383	393
momento: 20,817	392

All these instances were analysed, but approximately 75% in each language were cases where the head noun did not contain qualitative modification at all, so the study focused exclusively on the cases where the head was actually modified: 1,136 concordances in English and 1,086 in Spanish. Only single modifiers were taken into account for this study, in order to make the subsequent semantic analysis of the data more straightforward.

3.2 Data description process

This contrastive study combines two types of approaches in the course of its descriptive stage:

    a. first, an onomasiological phase, from meaning to form, that is from the common semantic function of nominal modification, to the grammatical means used by English and Spanish speakers to actualise it;

    b. and second, a semasiological phase, from form to meaning, where the different formal structures are classified semantically in an attempt to describe the syntax-semantics interface in this area of meaning.

The final stage compares the results of the semasiological phase in the two languages. Descriptive data for three parameters have been obtained: the grammatical structures, the semantic functions, and the actual head nouns chosen for the analysis, since their lexical nature proved to influence the semantic type of modification they may carry. This paper focuses exclusively on contrasting the semantic functions, which have provided the most interesting results from a translational point of view.

Traditional grammatical categories have been used in the syntactic classification of the means employed to modify nouns in English and Spanish. As for the semantic classification of these formal structures, a particularly broad model had to be found to account for the wide range of structures that may convey modifying meanings within the boundaries of the NP, from single words, to phrases and clauses. Considering the important role of adjectives as noun modifiers, the criteria used are based on the dichotomy between description and classification as two basic adjectival functions.

Numerous approaches support this theory, sometimes with changes in the terminology employed. Descriptive and classifying functions (Teyssier 1968; Warren 1984; Biber et al. 1999) are termed by some authors ascription and association (Ferris 1993), or epithetic and classifying functions (Tucker 1998).

This broad two-fold classification was used as the starting point for the semantic typology, and it was applied to all noun modifying structures alike. In the course of the analysis it became clear that in some cases neither of the two categories applied, so additional semantic categories were added, taking care to keep the classification as broad as possible and using well-established traditional terms to describe the semantic nuances found. Most of these additional functions were of the adverbial type, i.e. they indicated place, time, manner, purpose, etc. A total of eight semantic functions were contrasted in the two languages, although this paper will only present the results for the four most common ones: description, classification, time and place.

Semantic classifications are always problematic. Meaning is elusive and no clear boundaries can be easily drawn between different types of semantic categories, especially in the case of very broad categories such as the ones used in this study. The analysis revealed a considerable overlap of semantic functions in many cases in the two languages, and more than one semantic function often applied to the same instance of modification. However, a choice had to be made, and a number of guidelines were established in order to assign one particular structure to the descriptive group, the classifying group, or to some other category. I assume, firstly, that all the instances of nominal modification are inherently descriptive, since that is their defining nature as characterising resources. In many cases, though, that basic semantic function remains in the background and is overlapped by one or more additional functions, one of which can be considered as the dominant one in each case. On the basis of the strong interrelation between form and function, formal features were also used to determine whether a particular structure would be classified into one or the other group. This was the procedure employed for the classification of relative clauses, for example: when the relative clause indicated an adverbial function by means of an explicit formal marker, be it adverbial or prepositional (*where, when, in which, donde, en el cual*, etc.), it was classified as belonging to the corresponding adverbial function, but if no marker was present, it was assigned to the descriptive function, which was used as the default category in many borderline cases.

As an example, table 3 shows the different grammatical structures found modifying the English noun *day*, with the corresponding semantic classification and the number of instances found in each case. Similar tables were drawn for all the nouns selected for the study, and the data then added up and all the percentages calculated to obtain the final results, summarised in section 4 below.

Table 3. Description of resources and functions for the head noun *day*.

Grammatical resources	Semantic functions			
	Descr.	*Class.*	*Time*	*Place*
of-phrase	8	9	4	-
prem. Adjective	17	1	1	-
rel. clause	13	-	3	-
postm. Adverb	-	2	12	-
prem. Noun	-	12	-	-
PP *in*	2	-	1	1
PP *at*	-	-	-	2
-ed clause	1	-	-	-
prem. adverb	-	1	-	-
PP *on*	-	-	-	1

4. RESULTS: SEMANTIC FUNCTIONS IN CONTRAST

This section gives a brief summary of some of the results found, focusing exclusively on the juxtaposition and contrast of the four most common semantic functions found in English and Spanish and the corresponding formal means employed by native speakers of those languages to actualise those meanings. These four semantic functions amount to 93% of all cases of nominal modification in the English material, and to 86% of all cases in the Spanish material.

4. 1 The descriptive function in contrast.

The descriptive function is the most common of all the semantic functions of nominal characterisation found in English and Spanish. This function is the one that is intrinsically related to the notion of characterisation, and somehow all characterising resources include this function to a greater or lesser extent. In the English sample there were 965 concordances where the head noun was characterised by one single grammatical resource. Of these, 434 corresponded to the semantic function of description, about 45% of the total. In the Spanish sample there were 975 cases of single element characterisation and the descriptive function was the predominant one in 465 examples, about 48% of the total. The descriptive function is thus the dominant one in both languages, with nearly half of the occurrences, besides being present in all other cases too, overlapping with other functions. Figure 1 shows the percentage of occurrence of the most common structural resources found that were used to convey the descriptive function in English and Spanish[6].
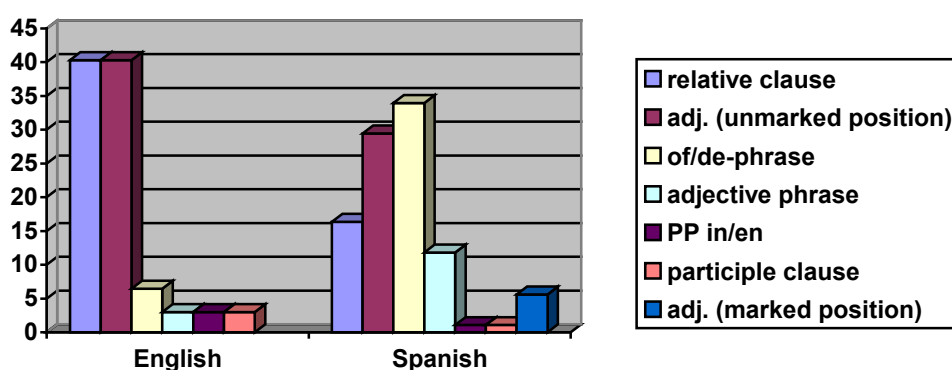

Figure 1: Juxtaposition of English-Spanish means used to express the descriptive function. (The unmarked position of adjectives is the premodifying one in English and the postmodifying one in Spanish.)

The most central resources for conveying this function in English are the relative clause (1) and the single premodifying adjective, i.e. the adjective in its unmarked position in the language (2).
   (1) a life which can repay our efforts; a man who is hit in the face
   (2) a stressful time; an exciting day

Both these resources show a similar frequency of about 40% each. All other descriptive structures are much less common and can therefore be considered rather peripheral, since they occur in only 6% of the cases or less. This semantic function is thus clearly marked grammatically in English and limited to a great extent to two dominant grammatical structures.

Figure 1 shows that the more peripheral descriptive resources in English are *of*-phrases (3), with 6.45% of cases, premodifying adjective phrases (4), PPs headed by the preposition *in* (5), and participle clauses, including both *–ed* and *–ing* clauses (6), the last three with about 3% of cases each:

(3) a man of great dignity; a night of moonlit romance
(4) a rapidly changing world; a thoroughly shameless man
(5) another year in wet, miserable conditions; The Man in the Iron Mask
(6) the man charged with sex murder; a man drawing money from a cash machine

In contrast, the grammatical realisation of the descriptive function in Spanish appears to be distributed much more evenly, with four resources showing a frequency of occurrence of over 11%. The two most central structures in this language are *de*-phrases (7), with about 34% of occurrences, and adjectives in their unmarked postmodifying position (8), with about 30% of occurrences:

(7) *el día de su fundación; el tiempo del bloqueo marítimo*
    "the day it was founded"; "the time of the sea block"
(8) *aquel día memorable; forma irregular.*
    "that memorable day"; "irregular form"

Relative clauses (9) and postmodifying adjective phrases (10) can be seen to be less typical, but also important for expressing this semantic function, with about 16% of cases and 12% of cases respectively.

(9) *un mundo que los arrincona; la vida que nos ha tocado vivir*
    "a world that corners them"; "the life we have to live"
(10) *una vida muy difícil; un mundo no siempre cómodo*
    "a very difficult life"; "a not always comfortable world"

Premodifying adjectives were found to convey a descriptive function in Spanish in only 5.5% of cases as in (11):

(11) *su turbulenta vida; un mal momento.*
    "his/her turbulent life"; " a bad moment"

PPs headed by the preposition *en* occurred in 1% of cases with a descriptive function (12), and participle clauses had a similar frequency of occurrence (13):

(12) *la vida en movimiento; el mundo en general*
    "life in motion"; "the world in general"
(13) *el mundo creado por Mozart; el día fijado para la inauguración*
    "the world created by Mozart"; "the day set for the inauguration"

The same grammatical structures can be identified for both languages, although they can be seen to be distributed quite differently. A horizontal bar chart (Figure 2), illustrates this distributional difference more clearly. The different grammatical means have been ordered here according to typicality of occurrence in the two languages.
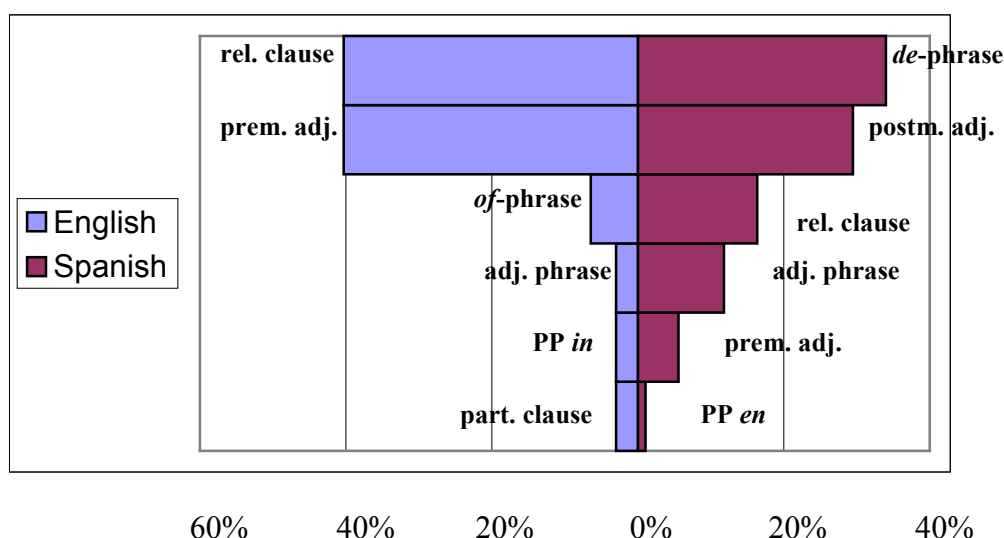
Figure 2: Contrast of English-Spanish means used to express the descriptive function.

The English half of the inverted pyramid is much less balanced than the Spanish half, clearly favouring the use of one of the two most common grammatical resources and leaving little space for the remaining options, which are all very peripheral. The Spanish half of the pyramid provides not only the list of formal means employed by native speakers of Spanish for actualising descriptive meanings in the NP, but also information about their typicality. That is, the results can be considered potential translation equivalents for the structures found with the same meaning in English texts, and vice versa.

A direct consequence of the concentration of the descriptive function in English on the two most central resources and of the more even distribution in Spanish, will be the fact that when English speakers employ one of these two structures the functional and semantic equivalents in Spanish will be more varied, including not only *de*-phrases or adjectives in their unmarked postmodifying position, but also relative clauses, adjective phrases and even adjectives in their marked premodifying position. The descriptive function should thus not be limited exclusively to the two formal equivalents of the most frequent structures in English. Relative clauses and adjectives do appear among the most common resources in Spanish, but neither of them is the most central in this language, where *de*-phrases are most commonly used to express this particular function. In addition, adjective phrases also show a relatively high frequency of occurrence in Spanish, close to the frequency of relative clauses, thus implying an important option for translators to take into account when translating English relative clauses or descriptive adjectives, especially in the case of highly meaning-specific adjectives or compound adjectives in English.

4.2 The classifying function in contrast

The classifying function is the second most common noun modifying function in both English and Spanish. Together with the descriptive function, the classifying function is considered one of the two most typical adjectival meanings intrinsically associated with the characterisation of nouns. A total of 238 of the English concordances analysed included the classifying function, about 25% of the total of nominal characterisation. The figures are very similar in Spanish: 236 concordances included a classifying resource, about 24% of the total. Figure 3 shows the most common resources employed in English and Spanish to actualise this semantic function.
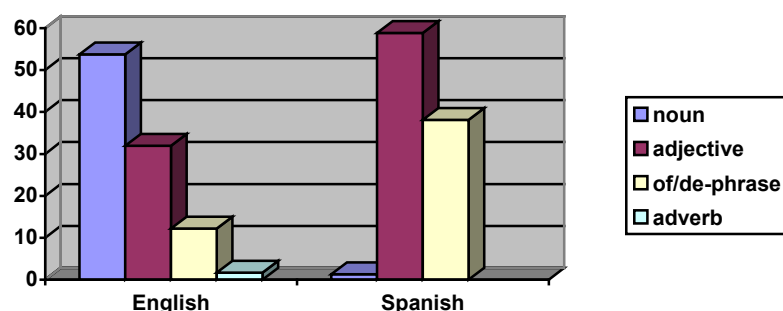
10

Figure 3: Juxtaposition of English-Spanish means used to express the classifying function.

In the case of this particular semantic function it may be said that there is no great dispersion of possible grammatical resources that may actualise it. Rather the opposite is true, since the function appears concentrated on only a few formal structures. The figure shows that there is one clearly discernible central resource in each of the two languages, and it is not the same one in English and Spanish. The premodifying noun is the means most typically expressing the classifying function in English, with about 54% of the cases:

(14)   lunch time; fashion world; luxury home; community life

In Spanish, in contrast, the adjective in its unmarked position, the postmodifying position, is the most central formal structure used, with almost 59% of the total:

(15)   *la vida política; el mundo empresarial; este año académico*
         "political life"; "the business world"; "this academic year"

Adjectives are also quite central resources in English in their unmarked premodifying function, with about 32% of the occurrences (16), and *de*-phrases in Spanish occur in 38% of the cases (17):

(16)   the natural world; social life
(17)   *tiempo de guerra; un mundo de fantasía*
         "time of war"; "a world of fantasy"

In English *of*-phrases occur in about 12% of cases (18), thus being relatively peripheral in comparison with the two most common resources, and postmodifying adverbs occur in only 1.5% of cases (19). Postmodifying nouns in Spanish occur in only about 1% of cases, and the examples found suggest borrowings from English (20):

(18)   *the world of showbiz; a day of mourning*
(19)   *a day off; a night out*
(20)   *el mundo Macintosh; el momento D*
         "the Macintosh world"; "D moment"

It can be noticed that no relative clauses are found to express a classifying function in either language. Apart from adverbs, the resources are identical in the two languages. However, the distribution of these resources across English and Spanish is radically different, as the most central resource in English, the premodifying noun, is actually the most peripheral of all the resources in Spanish. On the other hand, the most central resource in Spanish is only the second most central resource in English, and the *de*-phrase is far more frequent with the classifying function than its formal counterpart

in English, the *of*-phrase. Figure 4 illustrates the cross-linguistic contrast in the form of an inverted pyramid.
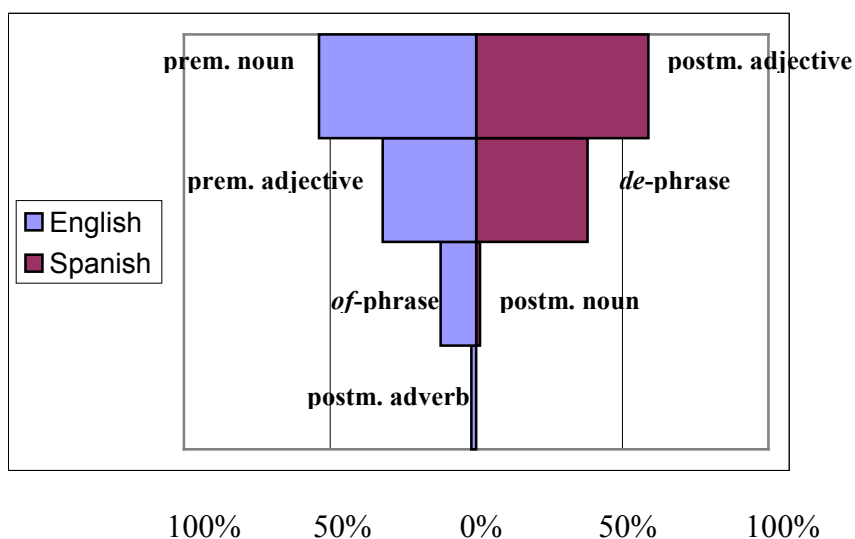

Figure 4: Contrast of English-Spanish means used to express the classifying function.

Figure 4 shows that the Spanish half is shorter and less balanced than the English half, thus representing graphically the fact that Spanish speakers rely more heavily on only two very central grammatical resources for actualising the classifying function in the NP, with few other options available. In contrast, the English half appears to be a bit more balanced, with a few more options available than in Spanish and with intermediate structures, neither too central nor too peripheral, as is the case with *of*-phrases.

One particularly striking fact is that about 85% of the modifiers corresponding to the classifying function occur in a premodifying position in English. In Spanish, on the other hand, this particular meaning is closely linked to the postmodifying position, the unmarked position for adjectives. Less than 1% of the cases correspond to premodifiers in the case of the classifying function in Spanish. These data confirm the hypothesis that classifying means do not normally occupy a premodifying position in Spanish. The unmarked position of noun modifiers is closely linked to issues in language typology, and the classifying function seems to be a clear example of one of the most striking typological features of Germanic and Romance languages, respectively.

In English, premodifying nouns are intrinsically associated with the classifying function. When a translator encounters this type of resource with this particular meaning, the semantic and functional equivalents in Spanish will be adjectives and *de*-phrases, which are the most central structures in this language to actualise classifying meanings, and practically the only ones available. The other two possible constructions in English cannot be considered problematic from the point of view of translation, as there are formal equivalents in Spanish that may also be used to actualise classifying meanings.

The data have shown that the classifying function is rather different in nature from the descriptive function, and this is reflected syntactically in the two languages. The descriptive and classifying functions together make up most of the cases of nominal modification in both languages, reaching about 70% of the total in English and about 72% in Spanish. Consequently, a line may be drawn between these two typical

adjectival functions and all the remaining functions, which are primarily of the adverbial type, including temporal or locative meanings, among others.

4.3 The temporal function in contrast

The semantic function of time is the most frequent one of the adverbially oriented functions in the NP in both languages, with 14.5% of the total in English (141 instances) and 12% of the total in Spanish (117 instances). This relatively high frequency of occurrence is due to the lexical nature of the head nouns analysed, several of them having temporal referents in the two languages: *time, year, day, week, año, tiempo, día, momento*. "Syntax is driven by lexis: lexis is communicatively prior" (Francis 1993:142), and for this reason the meanings and the structures employed here are directly influenced by the lexical nature of the noun carrying the modifier. Figure 5 summarises the main grammatical structures employed by native speakers of English and native speakers of Spanish in order to actualise the temporal function in the NP.
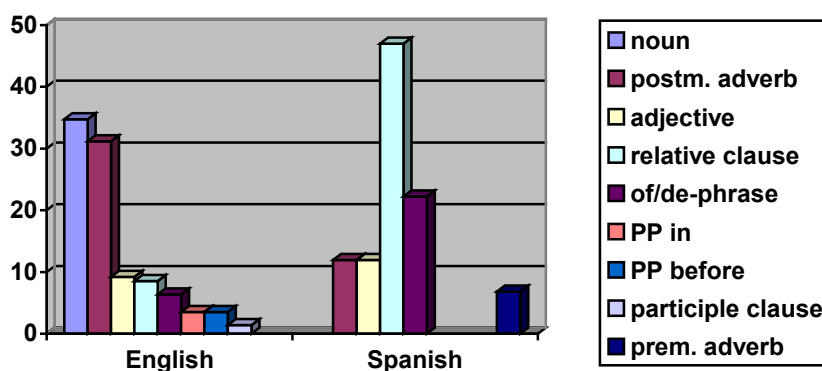

Figure 5: Juxtaposition of English-Spanish means used to express the temporal function.

The two most central grammatical resources used by native speakers of English to express this function are the premodifying noun (21) and the postmodifying adverb (22), with a very similar frequency of occurrence, 35% and 31%, respectively.
    (21)   Saturday night; winter night
    (22)   the day before; a week later

The remaining structures available in English are adjectives in about 9% of cases (23), relative clauses in 8.5% of cases (24) and *of*-phrases in about 6% of cases (25), all rather peripheral and occurring in less than 10% of cases each. PPs headed by the prepositions *in* and *before* occur in 3.5% of cases each (26), and participle clauses are found in about 1% of cases (27).
    (23)   the long day; late night
    (24)   the day when I can relax
    (25)   this time of year; the last day of the season
    (26)   a day in September 1952; the year before his death
    (27)   the week beginning June 19

Figure 5 suggests that in Spanish the expression of the temporal function is distributed in a completely different way. The most central resource in Spanish is the relative clause, with 47% of the total:

(28)   *el día en que murió su hijo mayor, el tiempo en que se produce*
        "the day his/her first child died", "the time when it happens"

De-phrases are next with about 22% of the cases:
(29)   *el día de hoy, el día de ayer*
        "today", "yesterday"

Postmodifying adjectives (30) and adverbs (31) in their unmarked positions show a similar frequency of occurrence, about 12% each:
(30)   *el momento justo, un momento determinado*
        "the exact moment", "a precise moment"
(31)   *poco tiempo después, un día antes*
        "a short time later", "one day before"

Finally, the Spanish corpus included a number of instances of a highly lexicalised structure in which the premodifying adverb *hoy* modifies the head noun *día*:
(32)   *hoy día*
        "nowadays"

This is a reduced alternative of the more formal expression *hoy en día*, with no occurrences in our corpus. These cases amount to about 7% of the total of the temporal function in Spanish. Figure 6 contains the same information ordered according to typicality criteria.
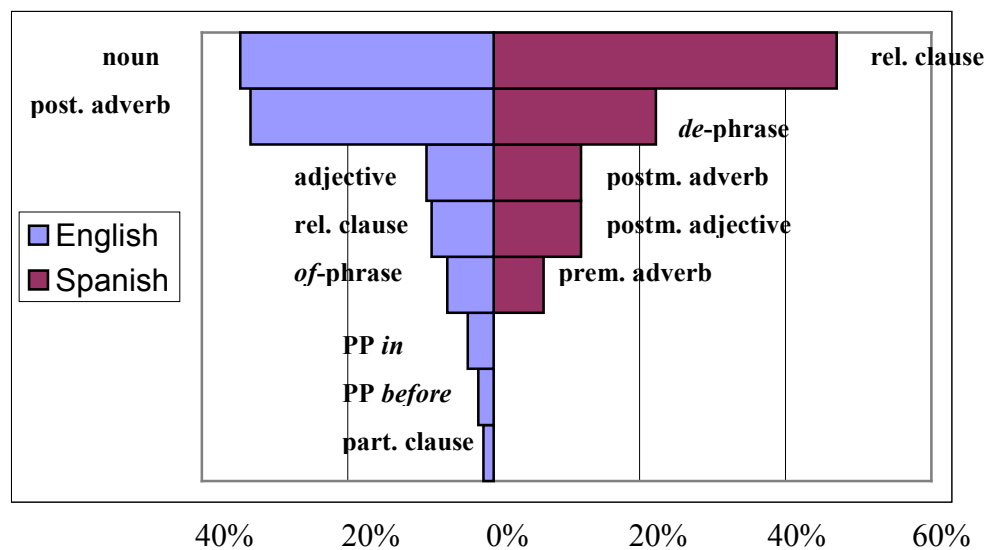

Figure 6: Contrast of English-Spanish means used to express the temporal function.

It should be noticed that, even though most of the grammatical means do occur in the two languages, their cross-linguistic distribution is very different. Premodifying nouns are the most central temporal resource in English, whereas there was not one single case of a formally similar structure in Spanish. Relative clauses also show important differences in their distribution across the two languages, being by far the most central resource in Spanish, but reaching only 8.5% of the total in English. This difference is due to the criteria employed for the semantic classification of relative clauses, which were only assigned an adverbial function if there was an explicit

adverbial marker for that function. In Spanish, a temporal particle is generally required if a temporal relationship is to be expressed by means of an adverbial relative clause. In contrast, in English the temporal relationship may be implied in the meaning of the modified head noun with no explicit marker indicating that relationship, since relative pronouns may be omitted in this language when they do not occur in subject position. Furthermore, "relative clauses with the head nouns *time* and *way* are exceptional in that they commonly occur with the preposition omitted" (Biber et al. 1999:626), and those two general nouns are among our head nouns. In all of these relative clauses in English, the adverbial function overlaps with the descriptive function, but if no adverbial particle is found in the context, the descriptive function was automatically assigned to the relative clause as the main function.

In contrast with this phenomenon, adverbs clearly show a much higher frequency of occurrence in English for expressing temporal meanings than in Spanish, representing a more synthetic way of expressing an adverbial function compared to the more analytic resource of relative clauses.

Taking these data into account, it may be inferred that when a noun is modified by one of the two most central resources in English, noun or adverb, the means available in Spanish that are equivalent from a functional-semantic point of view are the relative clause and the *de*-phrase: *the day before – el día de antes*. A lower degree of typicality applies to the remaining options in Spanish.

4.4 The locative function in contrast

The locative function was isolated in 85 instances of the 965 concordances analysed in the English corpus, amounting to about 9% of the total. In Spanish, however, only 20 instances of this semantic function were found in the 975 concordances analysed, just about 2% of the total of nominal characterisation. Languages have long been known to represent the particular view of the world held by their native speakers (Humboldt 1836). The data tie in with the general trends observed in the modification of nouns in English and Spanish. In English there is a clear trend towards a very graphic and precise description of the real world, a type of approach where spatial relationships play an important role, whereas in Spanish descriptions of entities tend to be less graphic and much more abstract (Vázquez Ayora 1977:44), and the locative function is therefore much less relevant in that language in the NP, and consequently less frequent too. Another reason that could explain this low frequency of the locative function in the Spanish corpus could be the fact that spatial relationships tend to be expressed not in the NP itself but rather after copula verbs, for example. This goes beyond the scope of this study which is exclusively concerned with modification within the scope of the NP.

Figure 7 shows the functional-semantic field for locative meanings in English and Spanish with the grammatical resources found to be used by native speakers of each language to express that particular meaning.
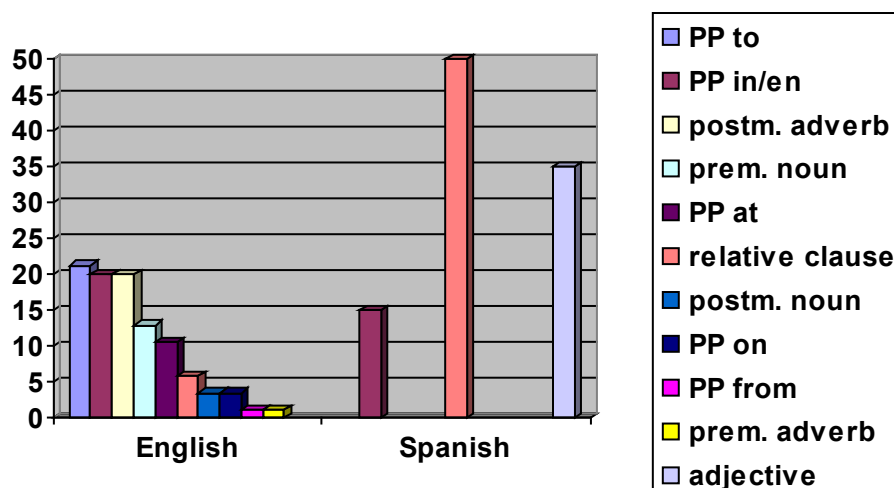
Figure 7: Juxtaposition of English-Spanish means used to express the locative function.

The most central structures used in English to actualise the locative function are PPs headed by the preposition *to* in about 21% of cases, all modifying the head noun *way*:

(33)   on the way to Havana; on the way to Nice airport; on the way to Dundalk

The next two resources occur in 20% of cases each and are PPs headed by the preposition *in* (34) and adverbs in their unmarked postmodifying position (35), all of which again applied to the noun *way*:

(34)   her home in America; life in Brixton; that night in Turin
(35)   on the way up; on her way back; your way down

Each of these three grammatical resources has a frequency of occurrence of approximately 20% of the total of the locative function, thus being the most typical structures employed by native speakers to express this meaning. The three structures are of a typically adverbial nature, thus logically related to the adverbial type of meaning they represent. The next two most common grammatical resources in English are somewhat more peripheral: the premodifying noun in about 13% of instances and having place names as their referents (36), and the PP headed by the preposition *at*, which also has a clearly locative meaning and occurs in 10.5% of cases (37).

(36)   her Chelsea home; his New York home
(37)   the time at the Carnaby plant; her home at 25 Cromwell Street

The remaining structures employed are relatively peripheral, with less than 6% of occurrences, and they include adverbial relative clauses in 5.5% of cases (38), the noun *home* as a postmodifier in 3.5% of cases (39), PPs headed by the preposition *on*, 3.5%, and the preposition *from*, 1% (40) and one case of a premodifying adverb, 1% (41).

(38)   a world in which camp manners have been …
(39)   their way home
(40)   life on Orkney; a man from Northern Ireland
(41)   the outside world

In Spanish, the distribution of the functional-semantic field of space may be seen to be radically different from the way this meaning is actualised in the English language. Considering the low frequency of occurrence of the locative function in Spanish NPs, it does not come as a surprise that only three grammatical structures have been found to convey this meaning. Locative connotations do not seem to be common in Spanish NPs in general. The most central means found is the relative clause with 50% of the total (42). The head nouns modified by these resources clearly have a locative referent, thus leading to a higher likelihood of locative modification.:

(42)   *un mundo donde siempre se juega; el mundo en el que le tocó vivir.*
        "a world where one is always playing"; "the world in which he had to live"

Next come single adjectives in their unmarked position in 35% of cases and locative referents too (43), and the PP headed by the preposition *en* – the most typical locative preposition in Spanish - in 15% of instances (44):

(43)   *la parte final; la parte central*
        "the final part"; "the central part"
(44)   *la vida en dicha península; un día en nuestra emisora.*
        "life in that peninsula"; "one day at our broadcast station"

Figure 8 shows the data corresponding to the locative function in the NPs ordered according to the degree of typicality in the two languages.
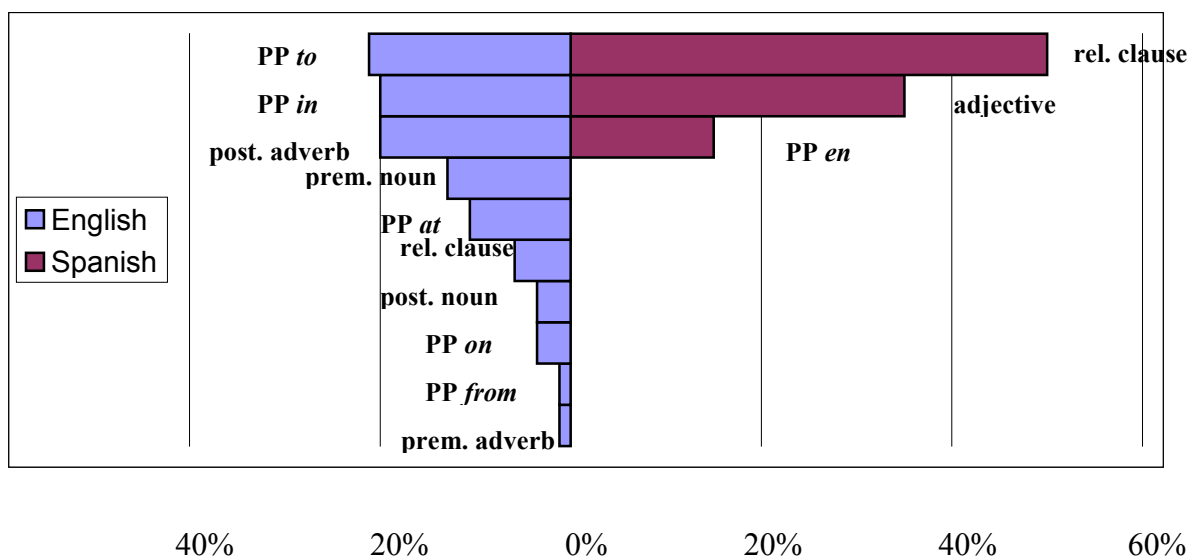

Figure 8: Contrast of English-Spanish means used to express the locative function.

The first thing that catches the eye when looking at Figure 8 is the far wider range of grammatical structures associated with this semantic function in English. Ten different resources were found in English, and only three in Spanish. Besides, some of the resources are closely linked to the lexical nature of the nouns chosen as head nouns, for example in the case of the PP headed by the preposition *to* following the head noun *way*.

It can also be seen that relative clauses and the PP headed by *in/en* appear in both languages, although with a somewhat different distribution. We find up to five

different prepositions heading PPs that modify nouns in English, each with their respective space-related function: *to, in, at, on* and *from*. Different meaning nuances are represented, from 'direction towards', to 'point in space' and 'provenance'. In contrast, there is no variation at all in Spanish, as the only type of PP found was the one headed by *en*, the preposition most clearly linked to locative meanings. Again, as in the case of the temporal semantic function, there is a considerable difference between the frequency of occurrence of relative clauses in English and in Spanish due to the criteria used for the semantic classification. Spatial meaning is mostly made explicit by some adverbial marker in the case of Spanish relative clauses, whereas in English this semantic relationship often remains implicit with no external marker, leading us to assign the relative clause to the descriptive type, even though a locative nuance overlaps with it.

The data show that when native speakers of English employ one of the three resources for expressing the locative function in that language, the equivalent functional-semantic options for expressing the same meaning in Spanish are basically relative clauses with an explicit locative marker, adjectives or PP headed by *en*. There is a great likelihood that English relative clauses and other more peripheral English resources will all correspond to relative clauses in Spanish, since this is the most central resource in that language. Chances are that PPs headed by *in* in English will at least in some cases not be translated by their formal equivalent in Spanish, but by any of the other more common options available in the language. At any rate, we are able to provide translators with the range of translational options available for the particular semantic function of space with the different degrees of typicality that correspond to each, represented by their frequency of occurrence.

5. CONCLUSIONS

This paper has given a general overview of the underlying rules that organise the functional-semantic field of nominal modification in English and Spanish. The study was carried out on the basis of empirical data extracted from two large monolingual reference corpora. The main advantage of this type of study based on comparable corpora is precisely the fact that all the data correspond to authentic instances of language in use as produced by native speakers of English and native speakers of Spanish. Consequently, the usefulness in foreign language learning and translation activities involving the two languages including translation practice, translator training and translation quality assessment, among others, is self-evident.

Using perceived similarity in meaning as a starting point (nominal modification), the contrastive functional method suggested by Chesterman (1998) has been followed to construct a descriptive framework of the structural patterns associated to this function in English and Spanish.

> It starts from perceived similarities of meaning across two or more languages, and seeks to determine the various ways in which these similar or shared meanings are expressed in different languages. [...] The perspective is from meaning to form. (Chesterman 1998:1).

The contrastive stage of our analysis (see section 4 above) has provided a number of parallel correspondences that include the structural options available in the two languages for actualising a particular semantic function. The data have been obtained independently in the two languages, and thus constitute potential translational options or translation equivalents in either direction (Labrador de la Cruz 2004). The inverted pyramids that summarise the results contain not only possible options for translators to choose from when they need to convey a particular semantic function, but

they also contain information about the degree of typicality of the different grammatical means in the two languages represented by their frequency of occurrence in our corpora. Real language use by native speakers is considered to be the most reliable source for these data, thus ensuring the naturalness of the target language in any translation process.

The results reveal the underlying trends that link form and meaning in one particular area of language that is known to be problematic in English-Spanish cross-linguistic activities. It was found that particular meanings tend to be actualised by particular structures, as shown in the figures in section 4. In addition, it was found that the reverse could be stated, since particular grammatical structures were found to be more closely linked to the expression of some meanings than to others, as in the case of premodifying nouns in English: "Every grammatical construction encodes a certain meaning, which can be revealed and rigorously stated, so that the meanings of different constructions can be compared in a precise and illuminating fashion, both within one language and across language boundaries." (Wierzbicka 1988:3). Furthermore, the third factor influencing the selection of a particular structure to express a particular meaning in the NP is the lexical nature of the head noun: "a word's syntactic behaviour reflects, and depends on, its meaning. (…) The grammatical behaviour of a word is governed by subtle 'semantic' rules." (Wierzbicka 1996:379). Overall, it can be said that the results found here would not have been very different if a different list of nouns had been used for the analysis, since it is not the actual nouns that have been studied, but the syntactic-semantic relationships of modification whenever they occurred, irrespective of the head noun being modified in each case.

The main aim of this study has been fulfilled, since the data included in the results section can be said to represent the functional-semantic fields (in Bondarko's terms) of the four most common functions found in nominal modification. The study has thus contributed to the consolidation of English-Spanish cross-linguistic studies by providing data on a particular semantic area and proving the validity of the corpus-based method employed here to obtain relevant data. No register differences were taken into account in the collection of data for this paper, but further studies could attempt to unveil register differences in the distribution of the syntactic patterns that actualize the different semantic functions. It would also be interesting to find out whether similar patterns function in spoken registers in the two languages. Finally, more in-depth studies in this semantic area should be based on a parallel corpus and the results found here could then be compared with translated Spanish so that deviations from naturally occurring language can be identified and analysed.

REFERENCES

Aarts, J.M.G. and Calbert, J.P. 1979. *Metaphor and Non-Metaphor. The Semantics of Adjective-Noun Combinations.* Tübingen: Max Niemeyer.

Aijmer, K., Altenberg, B. and Johansson, M. (eds). 1996. *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies.* Lund: Lund University Press.

Alameda, J.R. and Cuetos Vega, F. 1995. *Diccionario de frecuencias de las unidades lingüísticas del castellano.* Oviedo: University of Oviedo.

Altenberg, B. and Granger, S. 2002. "Recent Trends in Cross-Linguistic Lexical Studies". In *Lexis in Contrast: Corpus-based Approaches*, B. Altenberg and S. Granger (eds), 3-48. Amsterdam: John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. London: Longman.

Bondarko, A.V. 1991. *Functional Grammar. A Field Approach*. Amsterdam: John Benjamins.

Borin, L. 2002. " '… and Never the Twain Shall meet?' ". In *Parallel Corpora, Parallel Worlds*, L. Borin (ed), 1-43. Amsterdam: Rodopi.

Bugarski, R. 1991. "Contrastive Analysis of Terminology and the Terminology of Contrastive Analysis". In *Languages in Contact and Contrast. Essays in Contact Linguistics*, V. Ivir and D. Kalogjera (eds), 73-82. Berlin: Mouton de Gruyter.

Chesterman, A. 1998. *Contrastive Functional Analysis*. Amsterdam/Philadelphia: John Benjamins.

Darbelnet, J. 1969. *Pensée et structure*. New York: Charles Scribner's Sons.

Di Pietro, R.J. 1971. *Language Structures in Contrast*. Rowley, Mass.: Newbury Publishers.

Ferris, C. 1993. *The Meaning of Syntax. A Study in the Adjectives of English*. London: Longman.

Filipovič, R. 1974. "The Use of a Corpus in Contrastive Studies". In *Trends in kontrastiver Linguistik*, H. Raabe (ed), 51-66. Tübingen: Gunter Narr.

Francis, G. 1993. "A Corpus-driven Approach to Grammar – Principles, Methods and Examples". In *Text and Technology. In Honour of John Sinclair*, M. Baker, G. Francis and E. Tognini-Bonelli (eds), 137-156. Amsterdam: John Benjamins.

Fries, C.C. 1945. *Teaching and Learning English as a Foreign Language*. Ann Arbor: University of Michigan Press.

Halliday, M.A.K. 1991. "Corpus Studies and Probabilistic Grammar". In *English Corpus Linguistics*, K. Aijmer and B. Altenberg (eds.), 30-43. London: Longman.

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. London: Arnold.

Halliday, M.A.K., McIntosh, A. and Strevens, P.D. 1964: *The Linguistic Sciences and Language Teaching*. London: Longmans.

Humboldt, W. 1968 (1836). *Über die Veschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Bonn: Dümmler.

Ivir, V. 1981. "The Communicative Model of Translation in Relation to Contrastive Analysis". In *Kontrastive Linguistik und Übersetzungswissenschaft*, W. Kühlwein, G. Thome and W. Wills (eds), 209-218. München: Wilhelm Fink.

James, C. 1980. *Contrastive Analysis*. London: Longman.

Jaszczolt, K. and Turner, K. (eds). 1996a. *Contrastive Semantics and Pragmatics,* vol. 1., *Meanings and Representations*. Oxford: Pergamon.

Jaszczolt, K. and Turner, K. (eds). 1996b. *Contrastive Semantics and Pragmatics,* vol. 2., *Discourse Strategies*. Oxford: Pergamon.

Johansson, S. 1998. "On Computer Corpora in Contrastive Linguistics". In *Compare or Contrast? Current Issues in Cross-Language Research*, W.R. Cooper (ed), 269-289. Tampere: University of Tampere.

Johansson, S. and Oksefjell, S. (eds). 1998. *Corpora and Cross-linguistic Research. Theory, Method and Case Study*. Amsterdam: Rodopi.

Juilland, A. and Chang-Rodríguez, E. 1964. *Frequency Dictionary of Spanish Words*. The Hague: Mouton.

Krzeszowski, T.P. 1979. *Contrastive Generative Grammar: Theoretical Foundations*. Tübingen: Gunter Narr.

Krzeszowski, T.P. 1990. *Contrasting Languages. The Scope of Contrastive Linguistics*. Berlin: Mouton.

Labrador, B. 2004. "A Methodological Proposal for the Study of Semantic Functions across Languages". *Meta* 49(2): 360-380.

Lado, R. 1957. *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor: The University of Michigan Press.

Laviosa, S. 1997. "How Comparable Can 'Comparable Corpora' Be?" *Target* 9(2): 289-319.

Rabadán Álvarez, R. 1991. *Equivalencia y traducción*. León: University of León.

Ramón, N. 2002. "Contrastive linguistics and translation studies interconnected: the corpus-based approach". *Linguistica Antverpiensia. New Series* 1: 393-406.

Ramón, N. 2003. *Estudio contrastivo inglés-español de la caracterización de sustantivos*. León: Universidad de León.

Sinclair, J. 1991. *Corpus. Concordance. Collocation*. Oxford: OUP.

Teyssier, J. 1968. "Notes on the Syntax of the Adjective in Modern English". *Lingua* 20: 225-249.

Toury, G. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.

Tucker, G.H. 1998. *The Lexicogrammar of Adjectives. A Systemic Functional Approach to Lexis*. London: Cassell.

Vázquez Ayora, G. 1977. *Introducción a la Traductología. Curso Básico de Traducción*. Washington, D.C.: Georgetown University Press.

Warren, B. 1984. *Classifying Adjectives*. Göteborg: Acta Universitatis Gothoburgensis.

Wierzbicka, A. 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins.

Wierzbicka, A. 1996. *Semantics. Primes and Universals*. Oxford: OUP.

---

[1] This article presents part of the findings of my PhD thesis (Ramón 2003).
[2] Cobuild/Bank of English: www.cobuild.collins.co.uk
[3] CREA: Corpus de Referencia del Español Actual. www.rae.es.
[4] British National Corpus: http://www.natcorp.ox.ac.uk/
[5] Corpus del Español: http://www.corpusdelespanol.org/
[6] All the figures include only those resources that occurred in more than 1% of cases.