

**COMPARING ORIGINAL AND TRANSLATED SPANISH:
A CORPUS-BASED ANALYSIS OF ADJECTIVE POSITION**

Noelia Ramón

University of León, Spain

Address:

Department of Modern Languages

University of León

Campus de Vegazana s/n

24071 León, SPAIN

Tf. +34 987 291078; Fax. +34 987 291099; E-mail: noelia.ramon@unileon.es

Abstract

It is a well-known fact that translated texts present a number of peculiarities which distinguish its language from the one found in texts produced originally. Many studies have tried to name some of these phenomena, which are usually grouped together under the umbrella term of 'translation universals'. It has been demonstrated that translations do share a number of features irrespective of the source or target languages involved. Other divergences between original and translated texts are due to source language interference and are, therefore, language-dependent. This paper is a corpus-based study of several highly frequent Spanish adjectives in original texts and in texts translated from English. The unmarked position of attributive adjectives is the pre-modifying one in English and the post-modifying one in Spanish, though. Spanish also allows for the pre-modifying position with certain connotations. The aim of this study is to identify differences in behavioral patterns with respect to adjective position in original and translated Spanish and explain these differences in terms of translation universals and/or source language interference. The results have revealed cases of simplification, unique item under-representation and untypical collocations in Spanish translations of English source texts.

Key words: translation universals, corpora, adjective position, simplification, untypical collocations.

Résumé

Les traductions présentent des particularités qui distinguent leur langue de la langue trouvée dans des textes originaux. Beaucoup d'études ont tenté de nommer ces phénomènes, connus habituellement par le terme 'universels de traduction'. Il a été démontré que les traductions se ressemblent entre elles indépendamment des langues source ou cible impliquées. D'autres différences sont dues à l'interférence de la langue source, et sont donc dépendantes de la langue. Ce travail est une étude de corpus de plusieurs adjectifs espagnols très fréquents dans des textes originaux et dans des traductions de l'anglais. En anglais la position non-marquée de l'adjectif attributif est la position de pre-modification du nom, tandis qu'en espagnol c'est la post-modification. Cependant, la pre-modification est possible en espagnol aussi, avec des connotations particulières. Le but de cette étude est l'identification des différences dans les patrons de comportement de certains adjectifs espagnols en ce qui concerne leur position dans des textes originaux et dans des traductions, pour ainsi expliquer ces différences en termes d'universels de traduction et/ou interférence de la langue source. Les résultats ont dévoilé des cas de simplification, sous-représentation de termes uniques et collocations atypiques dans les traductions espagnols de textes originaux écrits en anglais.

Mots clés: universels de traduction, corpus, position adjectivale, simplification, collocations atypiques.

1. INTRODUCTION

When reading a Spanish translation of a text written originally in English one of the first things that strikes the reader if he/she is a native speaker of Spanish is the abundance of adjectives in the pre-modifying position. This abundance does not make the target text illegible, but it certainly detracts from its idiomaticity and results in an accumulation of untypical word combinations in the translation. This fact has prompted the current study.

Adjective position is one of the typological features where English and Spanish differ more clearly, since the unmarked position of adjectives is the pre-modifying one in English and the post-modifying one in Spanish. This dissimilarity is related to the different origin of these two languages, English being a Germanic language and Spanish a Romance language. However, the pre-modifying position of descriptive adjectives is possible in Spanish too, with a number of semantic connotations, involving mainly affectivity. Spanish translations of English texts may therefore be expected to present a higher rate of pre-modifying adjectives than texts produced originally in Spanish due to the influence of the source language English, where the pre-modifying position is dominant. Previous studies have shown that this is one important divergence between naturally occurring Spanish and Spanish translations from English (Rabadán *et alii* 2009). Adjective position can therefore be considered one of the main problems in translations between these two languages, and consequently a key point to take into account in translator training as well as in translation quality assessment.

In this paper I will provide a detailed analysis of some of the most frequent adjectives in English and Spanish, with reference in particular to their position in the noun phrase. The working hypothesis is that there will be a strong trend to locate

adjectives in Spanish translations in pre-modifying positions more often than if these same adjectives occurred in original Spanish texts. The empirical material used for this analysis will be extracted from the English-Spanish parallel corpus known as P-ACTRES compiled at the University of León, Spain. This corpus contains 2.5 million words of contemporary English texts and their corresponding Spanish translations. A variety of registers are represented: fiction, non-fiction, press, etc. The CREA corpus (*Corpus de Referencia del Español Actual*), a large reference corpus of Spanish, will be used for the comparison with original Spanish texts.

The aim of the paper is to illustrate how quantitative and qualitative factors such as the overuse of the pre-modifying position of adjectives, among others (Ramón 2009; Ramón and Labrador 2009), may be used for rating the quality of Spanish translations and should therefore be considered major aspects in translator training and in translation quality assessment. The analysis of empirical data extracted from the translation corpus may also provide a more general insight into the linguistic features of Spanish translations and translation universals, from a broader perspective.

2. THEORETICAL BACKGROUND

2.1. Trends in translated language

It is generally assumed that translations in any particular language read somewhat different from texts written originally in that same language. Many of these differences are due to the interference of the source language in the target language and are therefore, language-dependent; but many other of these peculiarities have been

found to occur in translated texts in general, irrespective of the languages involved in the translation process. All translated texts seem to share an aura of ‘secondaryness’, of being somehow different from original texts. Many scholars have given specific names to this phenomenon calling translated language ‘the third code’ (Frawley 1984) or using the terms ‘translationese’ (Gellerstam 1986) or ‘hybrid language’ (Trosborg 1997; Schäffner and Adab 2001).

In the past 20 years or so, the availability of large corpora (monolingual, bilingual, and multilingual) and the development of powerful tools for language analysis have enabled scholars to investigate translated language in great detail and from many different perspectives (Laviosa 2011). Numerous studies have since been devoted to identifying these common linguistic characteristics shared by all translated texts simply because they are translations, features called by Mona Baker ‘universals of translation’ in her seminal paper published in 1993. Some of these characteristics typically associated to translated texts include issues such as explicitation, simplification, and normalization. Translated texts have been found to include additional information and linguistic material to the one present in source texts, thus making translations more explicit (Olohan and Baker 2000). A number of studies have also found that translated texts simplify the target language to some extent, showing a lower lexical density and a higher frequency of occurrence of already common items in the target language (Laviosa 1996, 1998, 2002). Overlapping with this trend towards simplification, we find a tendency of translations towards linguistic conventionality or normalization, trying to normalize grammar, punctuation, dialect or other exaggerated target-language features, thus avoiding peripheral issues and remaining within the mainstream conventions of the target language (Toury 1995).

More recently, Mauranen (2008) has described further candidates for becoming universals of translation, as the result of extended corpus-based research using translation corpora from the early 2000s, and these include the under-representation of unique-items in the target language (Tirkkonen-Condit 2004), source language interference, and untypical collocations in the target language (Mauranen 2000, 2004).

In our study on the different usage patterns of adjectives in original and translated Spanish, we will contribute to this line of research by describing the collocational behavior of these particular linguistic items in translations with respect to original texts. Source language interference and untypical collocations are both issues that will have to be considered here and could appear in an analysis of the behavior of adjectives in original and translated Spanish.

2.2. Adjective position in English and Spanish

Typological differences between Germanic and Romance languages are many and varied, and word order is one of the areas where these differences are most obvious. The differences relating to adjective position in these two language types are particularly important and have clear implications in the translation process between English and Spanish, as we will see below. From a morphological perspective, adjectives are invariable lexemes in English, whereas in Spanish all adjectives vary morphologically to express number, and many of them also to express gender. But most importantly, as for adjective position, the unmarked position for attributive adjectives is the pre-modifying one in English and the post-modifying one in Spanish. This radical difference is one important source of error in foreign language learning as well as in the translation process involving these two languages.

Typologically, this difference is clearly related to the relatively strict word order we find in English as the result of a poor inflectional morphology. Except for a few fossilized expressions, the pre-modifying position of adjectives in attributive function is the only option available in English. In contrast, a very rich morphology in all content words in Spanish allows for a rather freer word order, including the possibility to locate descriptive adjectives in pre- or post-modifying positions:

(1) *un hombre pobre* vs. *un pobre hombre*; ‘a poor man’ vs. ‘a poor (unhappy) man’

Descriptive adjectives may occupy both positions within the noun phrase, though generally with different meanings or connotations, mainly involving affectivity in the pre-modifying position and a neutral meaning in the unmarked post-modifying position. A few very common adjectives even have morphologically distinct forms for the pre-modifying position, such as *gran* for *grande*, or *buen* for *bueno*. Classifying adjectives, on the other hand, are restricted to occur in the unmarked post-modifying position in Spanish:

(2) *un oso polar* vs. **un polar oso*; ‘a polar bear’

When is it idiomatic to place a particular descriptive adjective in front of a noun in Spanish and when is it not so idiomatic, but results in an untypical collocation? There is no clear answer to that question. No hard and fast rules exist in Spanish with respect to the use of the pre-modifying positions of descriptive adjectives. Phonemics seems to play a role, since mainly short adjectives may pre-modify longer nouns. It is also common to find pre-modifying adjectives in original Spanish when the head noun also has a post-modifier, especially if it is a classifying adjective, which cannot occupy the pre-modifying position, as in (3):

(3) *un gran oso polar*; ‘a big polar bear’

When multiple modification is involved, Spanish prefers locating one adjective before and one after the head noun, rather than coordinating both after the head noun. Despite these general trends, each particular adjective presents its own pattern with respect to dominant pre- or post-modifying positions and this pattern may vary depending on the head nouns. I will investigate in this paper the patterns of several of the most common Spanish adjectives in texts written originally in Spanish and, subsequently, compare the use of these same adjectives in texts translated into Spanish from English source texts to try and shed some light on this issue.

3. METHODOLOGY AND DATA

The aim of this paper is to reveal the collocational patterns of some of the most common Spanish adjectives with respect to their position in the NP, pre- or post-modifying position, in original and in translated texts. An overuse of the pre-modifying position in translated Spanish might be attributed to the influence of the source language English, where the pre-modifying position is the only one available. The empirical data for this study have been extracted from two electronic corpora: a monolingual reference corpus of Spanish and a parallel corpus of texts written originally in English and their corresponding translations into Spanish.

CREA (*Corpus de Referencia del Español Actual*) is a large monolingual reference corpus of the Spanish language available online. It contains slightly over 154 million words of texts in all registers, oral and written, and in all geographic varieties of contemporary Spanish, from 1975 until our days. From this large reference corpus we

have selected a subcorpus of texts from the year 2000 on, published in Spain only (Peninsular variety), and from the written section only. All in all there are 18,500,104 million words in our selection of original Spanish texts. This previous selection is done to restrict the analysis to the contemporary stage of the language and also to make the results comparable with the ones extracted from our corpus of translated Spanish, which includes only translations published in Spain from the year 2000 on.

P-ACTRES is an English-Spanish parallel corpus compiled at the University of León, Spain, and which currently contains nearly 2.5 million words. The corpus includes texts written in English and belonging to various registers with the corresponding translations into the European variety of Spanish. All original texts were produced in the year 2000 or later. For more details on the compilation process of this corpus, see Izquierdo *et alii* (2008). Table 1 shows the number of words in each subcorpus by register and language.

Table 1: Register distribution of P-ACTRES.

	ENGLISH	SPANISH	TOTAL
Books – fiction	396,462	421,065	817,527
Books – non-fiction	494,358	553,067	1,047,425
Newspapers	115,502	137,202	252,704
Magazines	119,604	126,989	246,593
Miscellanea	40,178	49,026	89,204
TOTAL	1,166,104	1,287,349	2,453,453

For this paper I will use the whole of the Spanish translation corpus, so all in all 1,287,349 words of translated Spanish. The analysis will consist of two stages:

1. First, I will carry out a quantitative study of some of the most frequent adjectives in Spanish to determine whether there is a clear under-use or overuse of any of these adjectives in the translations when compared to the use in original texts. This is done to see whether our translations show untypical item frequencies, as found in other studies (Gellerstam 1996; Laviosa 1996).
2. Second, I will carry out a syntactic analysis of a representative number of cases of some of these adjectives in both corpora to reveal the tendencies in use with respect to the position of these adjectives within the NP. Differences in the distribution of the pre-modifying position in particular could be attributed to the influence of the source language English.

The starting point for our analysis was the frequency list in the CREA corpus for Spanish, which was used to identify the most frequent adjectives in this language. The list of the 25 most frequent adjectives runs as follows: *gran, general, mayor, nacional, mejor, nuevo, pasado, nueva, social, grandes, posible, importante, final, unidos, cierto, largo, claro, español, buena, internacional, igual, española, interior, buen, especial*. The 25 most commonly employed adjectives in English, according to the Cobuild Wordbanks Online are: *new, good, old, long, little, great, high, best, big, national, small, full, young, free, public, important, white, local, black, able, early, political, real, hard, available*. We can see that many of them are semantic equivalents, some even cognates.

As for the Spanish list, some of the adjectives included are the short forms of adjectives for the pre-modifying position (*gran, buen*), so they are excluded from our qualitative analysis, as they may never occupy the post-modifying position. Other adjectives are of the classifying type (*nacional, social, internacional*), and must always

appear after the noun they modify, so they are not useful in this study either. I have selected 5 adjectives from this list – *grande*, *bueno*, *nuevo*, *importante*, and *largo* ('big, good, new, important, and long') - to determine the following issues in comparing original and translated Spanish: the overall frequency of use of these particular lexical items in original texts and in translations, and the statistical significance of the differences identified. The qualitative part of the analysis consists in the extraction of concordances of 3 common adjectives to analyze their particular patterns with respect to position in the NP in original and translated texts.

4. RESULTS AND DISCUSSION

4.1. Frequency analysis

The frequency analysis of the most common adjectives in Spanish has revealed interesting differences between original and translated texts. In this exclusively quantitative part of the study, all the morphological forms (masculine singular and plural, and feminine singular and plural, as well as short forms and irregular comparatives) of the 5 Spanish adjectives selected (*grande*, *bueno*, *nuevo*, *importante*, *largo*) were searched for in CREA and P-ACTRES for raw figures. The aim here was to identify divergences in the distribution of particular lexical items. It was found that in all cases there were statistically significant differences in at least one of the morphological forms.

Table 2 shows the number of cases in the two corpora used. Because the corpora vary greatly in size (over 18 million words in original Spanish and 1.5 million words in

translated Spanish), it was necessary to apply a hypothesis test for two independent proportions to determine the statistical significance of the differences found. This test measures the likelihood that the differences found between the two corpora are due to chance. Here we have considered statistically significant only p-values under 0.01, thus allowing for a confidence interval of 99%. The differences identified in this study can therefore be confidently attributed to reasons other than chance, in our case to the fact that one set of data was collected from original texts and the other from translations. Table 2 also shows the p-value for each lexical item and whether the difference is an overuse or an under-use in the translations.

Table 2: Number of cases and p-value of Spanish adjectives in CREA and P-ACTRES.

	CASES CREA	CASES P-ACTRES	p-value	TYPE OF DIVERGENCE
GRAN	11.697	1.047	0	Overuse
GRANDE	1.674	188	0	Overuse
GRANDES	6.135	477	0.019	Not significant
MAYOR	10.941	712	0.08	Not significant
MAYORES	2.508	162	0.35	Not significant
BUEN	3.819	198	0.00005	Overuse
BUENO	1.701	264	0	Overuse
BUENA	4.058	274	0.62	Not significant
BUENOS	1.117	106	0.002	Overuse
BUENAS	962	108	7e-8	Overuse
MEJOR	9.141	650	0.59	Not significant
MEJORES	1.939	142	0.55	Not significant
NUEVO	8.863	360	0	Under-use
NUEVA	6.037	402	0.39	Not significant
NUEVOS	3.683	177	0.000001	Under-use

NUEVAS	3.602	186	0.00006	Under-use
IMPORTANTE	6.272	368	0.001	Under-use
IMPORTANTES	2.550	169	0.53	Not significant
LARGO	5.194	321	0.03	Not significant
LARGA	1.436	126	0.012	Not significant
LARGOS	506	48	0.03	Not significant
LARGAS	472	49	0.007	Overuse

We will now see each case in greater detail and try to explain the differences identified in terms of the possible translation universals underlying the patterns found. As for the adjective *grande* and all its morphological variants, the short form for the pre-modifying position and the singular form of the post-modifying position are both overused in translations to a degree which is statistically significant. The plural and the comparative forms show small differences which are not statistically significant. The general high frequency of this adjective, which is the most frequent one in Spanish, makes us think that the strong trend to overuse it in translations is a feature of the lexical normalization of translated texts in general. This phenomenon is known as the ‘simplification hypothesis’ (Baker 1993; Toury 1995; Laviosa 1996) and suggests that translations tend to boost the use of typical features of the target language, showing a proportional over-representation of the most frequent lexical items in that language. However, this adjective will not be a good candidate for the second part of the analysis, as there are two separate morphological forms for the pre- and the post-modifying positions in Spanish.

A similar situation is the one observed in the case of the adjective *bueno* and all its variants. The short form, the masculine singular, and both masculine and feminine plural forms are all overused significantly in Spanish translations when compared with

texts written originally in Spanish. Again this indicates a trend towards the normalization or simplification of translated language in general, i.e., high-frequency lexical items or grammatical or syntactic resources tend to be preferred as translation solutions at the expense of other target language options, which would make the translations less flat and more lexically varied. This adjective presents two different morphological forms for the pre- and the post-modifying positions, so it will not be considered in the qualitative part of the analysis, where we will try to identify cases of overuse of the pre-modifying position in Spanish translations through the influence of the source language English.

In contrast to these two cases of overuse, the adjective *nuevo* and several of its morphological variants appear to be underused in Spanish translations. The adjective *nuevo* is very frequent in Spanish and has a very general meaning applicable to virtually any noun, so the under-use in translations may be due to the fact that other more precise and less general adjectives are being used instead by professional translators to avoid homogeneity, repetition or too much simplicity. The adjective *new* is the most frequent one in English texts, so the under-use of the formal and functional equivalent in Spanish may not be attributed to the influence of the source language, but rather to an attempt to achieve lexical variation by using other more specific adjectives with a similar meaning to convey the lexical content carried by the adjective *new*. This result seems to go against the simplification hypothesis observed in the previous two cases, but it must be said that many studies have also found tendencies that oppose simplification, particularly when collocational patterns are considered (Mauranen 2000). Contrary to other translation universals, in particular explicitation, simplification has not yet been completely supported or refuted. The analyses of patterns of lexical combinations most often indicate that translations present untypical trends, suggesting a wider rather than a

narrower variety in the use of the resources available in the target language. This seems to be the tendency observed in the case of the adjective *nuevo*.

On the other hand, because the same morphological forms are possible for the pre- and the post-modifying positions, this adjective will be a good candidate for the qualitative part of the analysis in this paper. A detailed observation of a sufficiently representative number of cases of this lexical item in original and in translated texts will reveal whether the pre-modifying position is actually overused in translations because of the influence of the source language English or not.

As for the adjective *importante*, only two forms have been considered, since this item varies only in number in Spanish, but not in gender. The singular form appears to be underused in translations when compared to original Spanish texts and this under-use is statistically significant. The formal equivalent in English, *important*, is also one of the most frequent adjectives in that language, so the under-use in the Spanish translations of this source language can only be attributed, as in the case of the adjective *nuevo*, to an attempt on the part of translators to avoid repetition and simplification by using other less general adjectives. Again we find evidence against the simplification hypothesis, and again this may be related to the more varied combinatory patterns in which this particular adjective seems to be involved.

Finally, I have checked the raw frequencies of the adjective *largo* and all its morphological variants in Spanish. The only form with a statistically significant overuse is the feminine plural form *largas*, but it is so infrequent in both corpora that this finding cannot be considered relevant. The findings indicate that there is a slight trend towards simplification in the case of the feminine plural form *largas*, but in the remaining forms, which are all far more frequent, the figures are very similar in originals and translations. One possible reason for this may be the fact that the adjective

refers to a physical quality which can be observed and is therefore not susceptible to interpretation in translations. In any case, this adjective does not seem to be one that is very prone to cause problems in translations between English and Spanish with regard to its frequency of occurrence. However, it may be used in both the pre- and the post-modifying position, so it is a good candidate for the qualitative analysis in the next section of this paper.

4.2. *Pre- and post-modification in translations*

The next step in this analysis consists in looking at sufficiently representative numbers of concordance lines of different adjectives to compare the figures of pre- and post-modification in originals and in translations. The adjectives *grande* and *bueno* cannot be used for this purpose, as mentioned above, because they have specific morphological forms for the pre-modifying position. I will therefore choose the other three adjectives for this study: *nuevo*, *importante*, and *largo*, in their masculine singular forms (unmarked gender in Spanish grammar), which are the most numerous ones in all cases. To be sure to study a sufficiently representative number of cases of each adjective in original and translated Spanish, I have used a statistical formula to determine how many occurrences are required:

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

The element n is the final sample we will analyze, N is the whole sample of occurrences, and E is the estimated error, in this case 0.05 for a 95% confidence margin. Table 3 below shows the number of instances analyzed of each of the adjectives

selected in both corpora, CREA for original Spanish and P-ACTRES for translated Spanish.

Table 3: Number of cases selected for the study.

	CREA		P-ACTRES	
	Total	Selected	Total	Selected
<i>Nuevo</i>	8,863	383	360	190
<i>Importante</i>	6,272	376	368	192
<i>Largo</i>	5,194	371	321	178
TOTAL	20,329	1,130	1,049	560

4.2.1. The case of *nuevo*

A careful analysis of the concordance of the adjective *nuevo* has revealed that this lexical item clearly prefers the pre-modifying position in its attributive function, both in original Spanish texts as well as in Spanish translations. Apart from the use of *nuevo* as a single pre- or post-modifier I have also distinguished other syntactic combinations, in particular those that refer to multiple pre- and multiple post-modification (always involving *nuevo*, of course), or a combination of both pre- and post-modification. Table 4 shows the number of cases found in the various positions with the corresponding percentages to make the data comparable in both corpora.

Table 4: Syntactic functions of *nuevo* in original and translated Spanish.

Function in context	Number of cases in	Number of cases in
	CREA	P-ACTRES
Pre- + post-modification	170 – 44.3%	75 – 39.4%
Single pre-modifier	122 – 31.8%	65 – 34.2%
Fixed expression ‘de nuevo’	63 – 16.4%	14 – 7.3%

Single post-modifier	21 – 5.4%	19 – 10%
Adjective used as noun	4 – 1.04%	-
Predicative position	2 – 0.5%	5 – 2.6%
Multiple pre-modification	1 – 0.2%	5 – 2.6%
Multiple post-modification	-	7 – 3.6%
Total	383	190

It can be seen that *nuevo* is a predominantly attributive adjective, as the predicative position is very infrequent (0.5% and 2.6%, in originals and translations, respectively). The two most frequent positions in both corpora are the pre-modifying positions, with or without additional post-modifiers, accounting together for 75% of the cases in original Spanish and for 73% in translated Spanish. Other uses are marginal in both languages, particularly multiple pre- or post-modification. The chi-square test¹ was applied to this set of data where possible² and the p-value obtained was 0.0071, i.e., the differences are statistically significant. Figure 1 shows the 4 most common functions identified in the analysis, those with 10% or more of the cases in at least one of the two subcorpora.

¹ The chi-square test is a statistical tool to determine whether the differences between two samples of different sizes are significant or not. A p-value lower than 0.01 indicates that the difference is not due to chance, but must be due to another reason, in this case the influence of the source language.

² The test can only be applied when the expected frequency is over 5, so only the four most common functions shown in Figure 6 were used. The same restriction was followed in the remaining two cases.

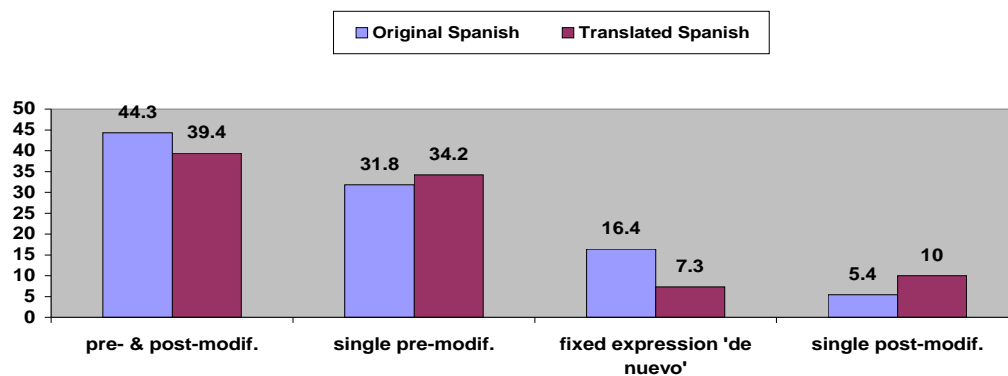


Fig. 1: Percentage of occurrence of the various functions of *nuevo* in original and translated Spanish.

The frequency distribution of the most common syntactic positions of this adjective seems to be quite similar in original and translated texts, with only small differences in most cases. In original Spanish the most common use of *nuevo* – 44% of the cases – is as the pre-modifier of a noun which also carries post-modification of various types: other adjectives, especially of the classifying type as in (4), participle clauses as in (5), and relative clauses as in (6), among others:

(4) *el nuevo planeamiento urbanístico consistiría en ...*; ‘the new housing plan would consist of ...’

(5) *un nuevo caso revelado ayer*; ‘a new case revealed yesterday’

(6) *un nuevo movimiento que sorteaba la censura*; ‘a new movement which escaped censorship’.

The same pattern is also the most frequent one in translations, though slightly less frequent than in original texts, with 39% of the cases:

(7) *un nuevo plan nacional*; ‘a new national plan’

(8) *un nuevo café que acaba de abrir*; ‘a new café which has just opened’.

The second most frequent option, single pre-modification, occurs in 31% of the cases in original texts, as in (9), and in 34% of the cases in translations, as in (10):

(9) *el nuevo sistema*; ‘the new system’

(10) *un nuevo conductor*; ‘a new driver’.

We notice here a slight overuse of this option, which may be influenced by the source language, but the difference is too small to be considered significant.

The use of *nuevo* in the fixed expression *de nuevo*, meaning ‘again’, occurs in 16% of the cases in original Spanish and appears to be much less frequent in translations, with less than half the cases (7%). This is a clear case of under-representation of a typical item in the target language. Because no similar expression exists in English involving the adjective *new* to be the source of this collocation in the target texts, *de nuevo* is dramatically underused. The chi-square test has shown that this function is the one that contributes most to the statistical difference between original and translated Spanish (52.5%).

Finally, single post-modification is an infrequent option in both corpora, although slightly more frequent in translations (10%) than in original texts (5%). Again here we may speak of a deviation from the typicality of source language patterns, but not because of the source language, as the post-modifying position is not possible in most constructions in English. This case is another example of an untypical collocation or structural option in translations. Despite the small number of occurrences, this deviation is relevant here, contributing 39.2% in the chi-square test to making original and translated Spanish different.

In general, the data reveal that there are not many differences between the use of *nuevo* in originals and translations when considering its position in the NP, which is predominantly the pre-modifying one, with or without post-modifiers. What is very frequent in originals is slightly less frequent in translations, and frequent combinations in original texts appear under-represented in the translations. In particular, the low

frequency of occurrence of the common expression *de nuevo* in translations may be considered an important difference between both corpora. Our initial hypothesis expecting a clear overuse of the pre-modifying position is not confirmed here, mainly because the pre-modifying position is also very frequent in original Spanish for this particular adjective *nuevo*. There is indeed a small overuse of *nuevo* as a single pre-modifier of 3%.

4.2.2. *The case of importante.*

As for the Spanish adjective *importante*, the analysis has revealed that this adjective is typically of the post-modifying type, with only rare occurrences in the pre-modifying position in both original and translated texts. One of the reasons for this clear trend may be its length, with 4 syllables, which makes pre-modification by *importante* sound cumbersome in Spanish. Table 5 shows the number of cases found in each corpus in each syntactic position. The chi-square test applied to this set of data revealed a p-value of 0.0001, so the differences found are statistically significant.

Table 5: Syntactic functions of *importante* in original and translated Spanish.

Function in context	Number of cases in	Number of cases in
	CREA	PACTRES
Single post-modification	112 – 29.7%	59 – 30.7%
Predicative position	95 – 25.2%	78 – 40.6%
Multiple postmodification	94 – 25%	27 – 14.06%
Pre- + post-modification	50 – 13.2%	24 – 12.5%
Single pre-modification	24 – 6.3%	3 – 1.5%
Adjective used as noun	1 – 0.2%	-
Multiple pre-modification	-	1 – 0.5%
Total	376	192

There are important divergences in the use of this adjective in original and translated texts. The single post-modifying position is the most frequent one in original texts, with nearly 30% of the cases, as in (11), and a very similar percentage was found in translated texts, as in (12):

(11) *una decisión importante*; ‘an important decision’

(12) *un riesgo importante*; ‘an important risk’

However, the predicative position is much more frequent in translations (40% of the cases, as in (13)), than in original texts (25% of the cases, as in (14)):

(13) *es importante recordar aquí ...*; ‘it is important to remember here ...’

(14) *le afectan numerosos factores y es importante analizarlos*; ‘many factors affect him and it is important to analyze them’.

This seems to point towards the influence of the English source texts. The English adjective *important* is also very frequent and may appear more often in the predicative position than its Spanish cognate, leading to this clear difference in use. The predicative position is the biggest contributor to the statistically significant difference between original and translated Spanish, with 42.7% of chi-square.

The third most common option found shows that original Spanish texts present another 25% of the cases of *importante* in a multiple post-modifying structure, where our descriptive adjective occurs immediately after a classifying adjective:

(15) *información estadística importante*; ‘important statistical information’.

Translated texts present this combination in only 14% of the cases:

(16) *un valor simbólico importante*; ‘an important symbolic value’

What we find here is a common language pattern which is under-represented in translations and contributes 30.3% of the chi-square.

The combination of pre- and post-modifying positions occurs in a very similar number of cases in both original (13%) and translated texts (12%), as shown in (17):

(17) *una importante presencia military*; ‘an important military presence’.

As in the case of other adjectives analyzed here, the descriptive adjective occurs in pre-modifying position mainly because there is a classifying adjective occupying the position immediately after the noun.

The remaining positions identified are rather infrequent in both corpora, although the single pre-modifying position occurs more often in original texts (6%) than in translations (1%), in another case of under-representation of a possibility which is more common in original texts. This result actually contradicts our initial hypothesis which expected an overuse of the pre-modifying position in Spanish translations. The reason in this case is obvious: the length of the adjective *importante*, with 4 syllables, plays against its idiomaticity in the pre-modifying position, up to the point of making it even less common in translations than in original texts. The single pre-modifying option is rather marginal in both original and translated texts, with less than 10% in both cases, but because the difference is so large, it contributes 26.5% of the chi-square.

Figure 2 illustrates the differences found in the four most common syntactic positions of the adjective *importante* in original and translated texts.

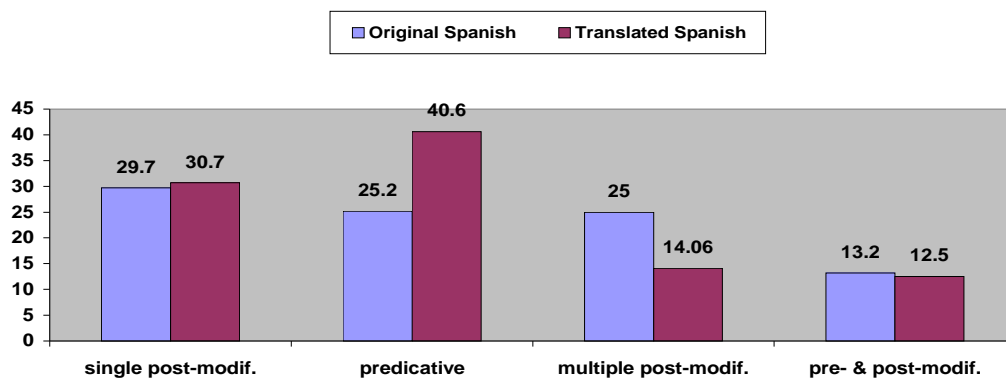


Fig. 2: Percentage of occurrence of the various functions of *importante* in original and translated Spanish.

The post-modifying position is dominant in original Spanish for the adjective *importante*, be it as a single post-modifier (nearly 30% of the cases) or in multiple post-modifying combinations (25%). All in all, 55% of occurrences of *importante* in original Spanish are post-modifying positions. In principle, this would make it a good candidate for a deviation in translations showing an overuse in the pre-modifying position. However, the data reveal that this is not so. The pre-modifying position is very infrequent in original Spanish (only 6%), mainly because of the length of the adjective *importante*, and this tendency is even more marked in translations (1.5%), as discussed above.

So where is the difference then? The difference between original and translated texts here lies in the predicative position. In original texts the predicative position accounts for 25% of the cases, but this position is dominant in translations with over 40% of the cases. This deviation may only be explained by the influence of the source language.

To sum up, in the case of the adjective *importante* we can see a trend towards fewer post-modifying and also fewer pre-modifying positions in translations than in original texts. Here we find many more predicative uses of this adjective in translations

than in originals. A different distribution pattern arises as the consequence of the translation process, resulting in untypical collocations and syntactic combinations.

4.2.3. *The case of largo.*

Finally, the Spanish adjective *largo* shows a clear preference for the pre-modifying position in both originals and translations. Table 6 shows the number of cases in each corpus with the corresponding percentage of occurrence. The chi-square test applied to this set of data revealed a p-value of 6.55e-27, so, as in the previous two cases, the differences found are statistically significant.

Table 6: Syntactic functions of *largo* in original and translated Spanish.

Function in context	Number of cases	Number of cases
	in CREA	In PACTRES
Fixed expression ‘a lo largo de ...’	195 – 52.5%	8 – 4.4%
Single pre-modification	85 – 22.9%	74 – 41.5%
Pre- & post-modification	35 – 9.4%	29 – 16.2%
Noun	26 – 7%	24 – 13.4%
Single post-modification	8 – 2.1%	21 – 11.7%
Predicative	8 – 2.1%	6 – 3.3%
Multiple post-modification	7 – 1.8%	14 – 7.8%
Multiple pre-modification	5 – 1.3%	2 – 1.1%
Numeral	2 – 0.5%	-
TOTAL	371	178

The analysis has revealed that in original texts the adjective *largo* appears mostly (52% of the cases) as part of the adverbial expression *a lo largo de ...*, meaning ‘along’, as in example (18):

(18) *a lo largo de la historia*; ‘along history’.

This is interesting because only 4% of the cases in translations show this option. This difference contributes 56.3% to the chi-square value. Here we come across a similar case to the one discussed above for the fixed expression *de nuevo*: a clear under-use of a particular expression typical of the target language. Of course translated texts are conditioned by their source texts, but translators seem to use other expressions with similar temporal meanings instead of *a lo largo de ...*, perhaps because no exact cognate with the word *long* is available or equally frequent in English. As in the case of *de nuevo*, I claim that this is the result of source language interference. In the case of *largo* the deviation is so important that the remaining syntactic positions will always be far more frequent in translations.

In translations, the most frequent use of the adjective *largo* is as a single pre-modifier (41%):

(19) *su largo cuerpo*; ‘his/her long body’

And this is also the second most common use in originals, though with only half of the cases (22%):

(20) *el largo verano*; ‘the long summer’.

There is, consequently, an overuse in the pre-modifying position in translations, an expected result if we consider the source language English. We may consider this

again a case of untypical collocation in the target language, contributing 11.1% to the chi-square value between originals and translations.

The combination of pre- and post-modification occurs in nearly 10% of the cases in original Spanish (21) and in 16% of the cases in translated texts (22):

(21) *un largo proceso evolutivo*; ‘a long evolutive process’

(22) *un largo camino que desciende hacia el río*; ‘a long road which leads down to the river’.

In this case again we can see a trend towards a more frequent use of the pre-modifying position in translated texts, contributing 3.7% of the total chi-square value.

The use of the word *largo* as a noun, mainly in expressions of measure indicating length occurs twice more often in translations than in original texts, with 13% and 7% of the cases, respectively, contributing 4.2% of the chi-square value:

(23) *tiene más de 4.500 kilómetros de largo*; ‘it is over 4,500 km long’

This phenomenon may be attributed to the influence of the source language English, where the use of *long* is required in expressions where the adjective *largo* may be idiomatically omitted in Spanish.

Finally, the use of *largo* as a single post-modifier is also much more frequent in translations (24) than in originals (25), (11% versus 2%), contributing in this case 16.1% of the chi-square value:

(24) *memorizar un poema largo*; ‘learn a long poem by heart’

(25) *tenían el morro largo*; ‘they had long snouts’

Similarly, *largo* in multiple post-modifying structures is also more frequent in translations (26) than in original texts (27), with 7.8% versus 1.8%, contributing 8.5% to the chi-square value:

(26) *el pelo rubio y largo*; 'long and blond hair'

(27) *un trazo largo y recto*; 'a long and straight line'

Both cases are examples of untypical collocations in the target language, since the data have revealed that native speakers of Spanish clearly prefer the adjective *largo* in pre-modifying positions or in the fixed expression *a lo largo de ...*

The remaining cases are very infrequent in both corpora. Figure 3 shows in percentages the most frequent positions in original and translated texts.

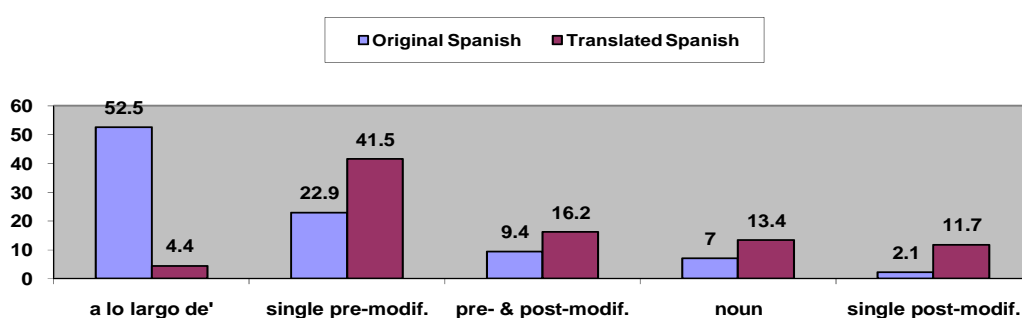


Fig. 3: Percentage of occurrence of the various positions of *largo* in original and translated Spanish.

Figure 3 illustrates very clearly that the findings are extremely biased by the different frequency of use of the fixed expression *a lo largo de* in original and translated texts. From a lexical perspective, we may consider this difference as an under-use of a unique item in the target language (Tirkkonen-Condit 2004). The target texts do not become ungrammatical or difficult to read by this different distribution, but it certainly detracts from idiomaticity. The explanation for this fact seems to be the non-existence of an expression with a similar meaning including the functional equivalent *long* in the source language English.

As for the remaining syntactic functions, we may say that the adjective *largo* shows a strong trend towards the pre-modifying position, with or without additional post-modifiers in the same NP, both in original and in translated texts. This means that it is

not a good candidate for being overused in the pre-modifying position in translations, which was our initial hypothesis. The data show that, apart from the great difference with respect to the fixed expression *a lo largo de ...*, there are no major divergences in the positioning of this adjective in original and translated texts. In fact, the second largest contributor to the chi-square significance value (16.1%) is the single-post-modifying position, which occurs nearly 10 times more often in translations than in originals, reflecting a case of untypical collocations in the target language Spanish which may not be attributed to the influence of the source language English.

5. CONCLUSIONS

From a typological perspective English and Spanish are very different as far as adjective position is concerned, English having the pre-modifying position as the unmarked one, and Spanish having the post-modifying position as the unmarked one, but admitting the pre-modifying position in descriptive adjectives with certain semantic connotations. Previous studies have shown a generalized overuse of the pre-modifying position in translated Spanish (Rabadán *et alii* 2009), something which may be attributed to the influence of the source language English. This paper has tested this initial hypothesis by looking at the patterns of the most frequent Spanish adjectives, in originals and translations.

In this paper I have analyzed raw frequency data and representative numbers of concordance lines of the most common Spanish adjectives in original texts and in texts translated into Spanish from English originals. The adjectives analyzed are *grande*, *bueno*, *nuevo*, *importante*, and *largo*, with all their morphological variants in Spanish.

The empirical data for the analysis have been extracted from two large corpora: CREA for original Spanish and P-ACTRES for translated Spanish. A corpus-based methodology is the most reliable way to confirm or reject our initial hypothesis of a systematic overuse of the pre-modifying position of adjectives in Spanish translations. The study has revealed the various usage patterns in original and translated texts for the most frequent Spanish adjectives, and the differences found have been interpreted with the help of statistical tests. For each case, I have tried to provide explanations for the differences in terms of translation universals and/or source-language interference.

The analysis contains two separate parts: (1) a quantitative study for determining overuse or under-use of the raw figures of the most frequent adjectives in original and translations, and (2) a qualitative analysis of the actual use of the most common adjectives in context, in particular with respect to their syntactic position with respect to the noun they were modifying.

The quantitative analysis revealed interesting differences between original and translated language, with a statistically significant overuse in translations of the adjectival forms *grande*, *bueno*, and *largas*, but also a significant under-use of *nuevo* and *importante* in translations. The cases of overuse may all be explained by the simplification hypothesis, as all these adjectives are highly frequent in Spanish. The simplification hypothesis states that translations tend to present already common items in the target language in even higher frequencies of occurrence in translations. In contrast, the cases of under-use point in the opposite direction, namely towards an attempt to avoid the homogeneity conveyed by high-frequency adjectives in the target language.

The qualitative analysis has also yielded very interesting data, although not in the sense intended initially with this study. The findings of the empirical analysis have

actually rejected our hypothesis that the pre-modifying position would be statistically more frequent in translations than in original Spanish texts. The adjectives *nuevo* and *largo* show a clear preference for the pre-modifying position in original Spanish texts already. There is, indeed, a slight overuse of this position in translations, but the difference is either very small, or the number of cases is too low to make the difference relevant. In fact, the significant differences lie elsewhere. In the case of the adjective *importante*, the pre-modifying position is even significantly *less* common in translations than in original texts!

So the data clearly contradict our initial hypothesis, but other very significant differences have been unveiled contributing to our understanding of translation universals and source language interference in translated texts. The main finding refers to the fact that two of the adjectives analyzed (*nuevo* and *largo*) are used less frequently in fixed expressions in translations than in originals (*de nuevo*, *a lo largo de*), and this under-use is statistically significant and, therefore, not due to chance. This clear trend illustrates a lower degree of typicality and idiomaticity in translated language, and can be explained in terms of the unique item under-representation in translated texts. Because these two adjectives appear in common everyday expressions, it was possible to determine that translators under-use them, possibly because no expression with a similar meaning exists or is equally frequent in the source language English including the functional equivalent adjectives in that language, *new* and *long*. Similar studies carried out with texts translated from a different source language where these expressions do exist (for example French), may yield different results, and would then confirm the source language interference factor here.

Another relevant finding refers to the behavior of the adjective *importante* in our two corpora. This adjective occurs much more often in the predicative position in

translations than in original texts and this difference is statistically significant. Because the cognate English adjective *important* is equally frequent in that language, we may argue that this difference in the syntactic position is due to source language interference. It may well be the case that the adjective *importante* occurs more often in the predicative than in the attributive position in English, and this trend is then replicated in the Spanish translations, as this option also exists in the target language, though with a much lower frequency of occurrence. At any rate, we may talk here of a case of untypical collocations in the target language, most probably due to source language interference.

An additional difference found in the data shows that the adjective *importante* occurs less often in multiple post-modifying structures in translations than in original texts. Translators show a clear trend to under-use this particular resource available in the target language, and this fact may be related to the length of the adjective in Spanish, with 4 syllables, which makes its use rather cumbersome in multiple modifying constructions. Here we are confronted again with the under-representation of a unique item in the target language, but also with what seems to be a trend towards simplification in translations, with fewer complex structures as modifiers.

In conclusion, each adjective shows different usage patterns in original texts and different divergences with respect to their translations. The initial hypothesis of an overuse of the pre-modifying position in translations has not been confirmed by the empirical data of the study, and what's more, no general trend can be applied to all the adjectives studied. Traces of translation universals have been identified, as in the case of simplification or under-representation of unique target language items, but other pieces of evidence argue against simplification. Different patterns of use with respect to the occurrence of the predicative position have been identified, and attributed to the influence of the source language. Most commonly, however, untypical collocations

have been detected in all cases as causing statistically significant differences between original and translated Spanish.

This study has focused on a very small number of adjectives, and no general conclusions can be drawn for the whole class of adjectives in Spanish as such. But the data very clearly illustrate different types of divergences between original and translated language, from simplification, to untypical collocations or unique item under-representation. The findings highlight here the importance of the lexical factor, as each item behaves in its own particular way. It is the untypical combinations resulting from locating one particular adjective next to one particular noun that often gives translations their peculiar foreign flavor. This fact makes it extremely difficult to pinpoint any syntactic category or lexical item as the one responsible for the foreign touch of a translation, as this effect is not due to a single unit, but to a syntactic position of many different lexical items, which vary greatly from one text to another. In Mauranen's words, "detailed analyses tend to show that the behavior of different linguistic items is not identical; therefore we need to consider factors other than overall tendencies of a very general kind." (Mauranen 2008, 40).

The results of this study shed light on the various possible translation universals existing between Spanish originals and Spanish translations. Untypical collocations were found to be the most common differences identified in our analysis, together with simplification or unique item under-representation. These findings may be useful in fields such as translator training or translation quality assessment, where the linguistic quality of the target language is essential.

ACKNOWLEDGEMENTS

This paper has been partly funded by the following grants: Ref. LE025A09 (Junta de Castilla y León) and Ref. FFI2009-08548 (Spanish Ministry of Science and Technology).

REFERENCES

- Baker, Mona. 1993. "Corpus Linguistics and Translation Studies." In *Text and Technology. In Honour of John Sinclair*, ed. by Mona Baker, Gill Francis, and Elena Tognini-Bonelli, 233-250. Amsterdam/Philadelphia: John Benjamins.
- Frawley, William. 1984. "Prolegomenon to a Theory of Translation." In *Translation: Literary, Linguistic and Philosophical Perspectives*, ed. by William Frawley, 159-75. Newark: University of Delaware Press.
- Gellerstam, Martin. 1986. "Translationese in Swedish Novels Translated from English." In *Translation Studies in Scandinavia*, ed. by Lars Wollin, and Hans Lindquist, 88-95. Lund: CWK Gleerup.
- Gellerstam, Martin. 1996. "Translation as a Source of Cross-linguistic Studies." In *Languages in Contrast*, ed. by Karin Aijmer, Bengt Altenberg, and Mats Johansson, 53-62. Lund: Lund University Press.
- Izquierdo, Marlen, Knut Hofland, and Øystein Reigem. 2008. "The ACTRES Parallel Corpus: An English-Spanish Translation Corpus." *Corpora* 3 (1): 31-41.
- Laviosa, Sara. 1996. "Comparable Corpora: Towards a Corpus Linguistic Methodology for the Empirical Study of Translation." In *Translation and Meaning: Part 3*, ed. by Marcel Thelen, and Barbara Lewandowska-Tomaszczyk, 153-163. Maastricht: Hogeschool Maastricht.

- Laviosa, Sara. 1998. "Core Patterns of Lexical Use in a Comparable Corpus of English Narrative Prose." *Meta* 43 (4): 557-570.
- Laviosa, Sara. 2002. *Corpus-based Translation Studies. Theory, Findings, Applications*. Amsterdam: Rodopi. 138 pp.
- Laviosa, Sara. 2011. "Corpus-based Translation Studies: Where does it Come from? Where is it Going?" In *Corpus-based Translation Studies. Research and Applications*, ed. by Alet Kruger, Kim Wallmach, and Jeremy Munday, 13-32. London: Continuum.
- Mauranen, Anna. 2000. "Strange Strings in Translated Language. A Study on Corpora." In *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*, ed. by Maeve Olohan, 119-141. Manchester: St. Jerome.
- Mauranen, Anna. 2004. "Corpora, Universals and Interference." In *Translation Universals. Do They Exist?*, ed. by Anna Mauranen, and Pekka Kujamäki, 65-82. Amsterdam/Philadelphia: John Benjamins.
- Mauranen, Anna. 2008. "Universal Tendencies in Translation." In *Incorporating Corpora. The Linguist and the Translator*, ed. by Gunilla Anderman, and Margaret Rogers, 32-48. Clevedon: Multilingual Matters, Clevedon.
- Olohan, Maeve, and Mona Baker. 2000. "Reporting *that* in Translated English. Evidence for Subconscious Processes of Explicitation?" *Across Languages and Cultures* 1 (2): 141-158.
- Rabadán, Rosa, Belén Labrador, and Noelia Ramón. 2009. "Corpus-based Contrastive Analysis and Translation Universals: A Tool for Translation Quality Assessment English-Spanish?" *Babel* 55 (4): 303-328.
- Ramón, Noelia. 2009. "Translating Epistemic Adverbs from English into Spanish: Evidence from a Parallel Corpus." *Meta* 54 (1): 73-96.

- Ramón, Noelia, and Belén Labrador. 2009. "Translations of -ly Adverbs of Degree in an English-Spanish Parallel Corpus." *Target* 20 (2): 275-296.
- Schäffner, Christina, and Beverly Adab. 2001. "The Idea of the Hybrid Text in Translation: Contact as Conflict." *Across Languages and Cultures* 1 (2): 167-180.
- Tirkonnen-Condit, Sonja. 2004. "Unique Items – Over or Under-represented in Translated Language?" In *Translation Universals. Do They Exist?*, ed. by Anna Mauranen, and Pekka Kujamäki, 177-184. Amsterdam/Philadelphia: John Benjamins.
- Toury, Gideon. 1995. *Descriptive Translation Studies and Beyond*. Amsterdam/Philadelphia: John Benjamins. viii + 311 pp.
- Trosborg, Anna. 1997. "Translating Hybrid Political Texts." In *Analysing Professional Genres*, ed. by Anna Trosborg, 145-158. Amsterdam/Philadelphia: John Benjamins.