

## A CORPUS -BASED STUDY ON THE MULTAL AND PAUCAL SERIES IN ENGLISH AND SPANISH

M<sup>a</sup> BELÉN LABRADOR DE LA CRUZ

Universidad de León

Before going into the matter, I'd like to devote a word of gratitude to the people who have provided me with the necessary materials for this paper. Without the help of Professor Guillermo Rojo, in the case of ARTHUS, the Spanish corpus, and without the help of the Cobuild team, I wouldn't have been able to carry out the present research.

As a sort of introduction, I shall describe, very briefly, the two corpora involved. As I couldn't find any bilingual corpora specifically designed for contrastive linguistics, that is, corpora made up of source texts in two different languages, not texts in a language and their translations in another (in which case, the target texts might be influenced by the source language), I chose to use two monolingual corpora. ARTHUS, "archivo de textos hispánicos de la Universidad de Santiago", is a one million word corpus (1,141,431), selecting only the variety of Spanish used in Spain, and it is mainly composed of books (novels, essays and plays) and oral transcriptions. As for the corpus of Cobuild, it has 323 million words, 50 million of which are on line. I decided to take only the subcorpora of British books and of oral British English, with 14,626,841 words. In any case I have calculated all figures per million so that the results are comparable.

The next step was the choice of the formal items that were to constitute the input, the search queries, the "prototypical equivalents": "one can start from L2 establishing a prototype and family resembling constructions in L2 and pass on to the analysis of matching constructions from L1... Finally it may be shown which properties are shared by the two prototypes." (Kalisz, 1988:42). I decided on the following items: *little, few, much, many* and *lot* and *poco, poca, pocos, pocas, mucho, mucha, muchos* and *muchas*. *Lot* was the most controversial unit; although it looks like a partitive, I included it in the set because I considered it had a long tradition of being studied along with *much* and *many* due to one functional fact: it is said to bear a relationship of complementary distribution with them, although this fact

will be put in doubt in the analysis. As for the others, they appear to be the main representatives of the multal and paucal expression of quantity. The two main guidelines for this purpose are as follows: "Underlying such specific instances of language use, are more general functions which are common to all cultures", in this case, quantification and "we shall consider language in terms of its use", which can be better seen in a corpus (Halliday, 1980: 141-142). Hymes also talks about use, about observing whether and to what degree the linguistic elements can be possible, feasible, appropriate and actually performed. (Hymes, 1971: 281). The time has arrived to go a step further and add a fifth variable: to see whether and to what degree they are typical, which joins the notions of performance and frequency, that is, how often they are done, how common and idiomatic they are. The first part of the analysis is then to do with the rates of frequency, shown in this chart:

	SPANISH QUANTIFIERS AND FREQUENCY FIGURES		ENGLISH QUANTIFIERS AND FREQUENCY FIGURES	
MULTAL QUANTIFIERS	MUCHO	1,216	MUCH	2,256.9 / m.
	MUCHA	197	MANY	1,496.9 / m.
	MUCHOS	316	A LOT	1,315 / m.
	MUCHAS	362		
PAUCAL QUANTIFIERS	POCO	1,317	LITTLE	1,449.3 / m.
	POCA	67	FEW	912 / m.
	POCOS	118		
	POCAS	66		
TOTAL		3,659		7,430.1 / m.

FIGURA 1

The fact that the total frequency rate of the English quantifiers is double the number of the Spanish group can have different explanations: the selected Spanish terms are in fact grammatical variants, in gender and number of two lexemes *poco*, *mucho* whereas in English, we have selected five different lexemes. Secondly, due to the fact that Spanish is a more analytical language and it is morphologically richer than English, the lexemes are often modified by suffixes. But searching for all the word forms starting with their roots would have allowed for the inclusion of words such as *muchedumbre*, *muchacho* and *poquedad* apart from *muchísimo*, *poquito*, etc. The inclusion of these morphological variants would obviously increase the present figures. Some other noteworthy details about these data are the following:

a) Most grammars say that *much* and *many* usually occur in non-assertive contexts and *a lot* in assertive ones. It is surprising then that the most frequent of them is *much* and not *a lot* (assuming that assertive utterances are more common), which is, in addition, not limited to being followed by only countable or uncountable nouns.



certain reality. And finally, regarding *bastante* and *demasiado*, they are still more subject to relativity from the speaker's perspective; they add certain connotations and enter different nets of relations: it makes sense to say "poco pero bastante" but "poco pero mucho" is either paradoxical or nonsense.

And to round off the section concerning paradigmatic relations, let me just mention the case of *much* as opposed to *many* and *little* as opposed to *few*. It has long been said that their choice depends on the nature of the noun they modify; whether it is mass or count. However, this feature of nouns has been called into question: "Countability is not a fact characteristic of nouns per se, but of NP's; thus it is associated with nouns in syntagmata, not with nouns as lexical entries"; he talks about countability preferences and levels of countability (Allan, 1980:546). Now it looks as if the relationship between two members of a paradigm can be better observed through their syntagmatic links. This takes us to the text, the words in context. And that's why the following step was to get lists of the most frequent collocates (the words on either sides of the quantifier), the top 25, to be more precise. Some of the conclusions that were drawn from the lists of collocations and supported by observing the concordances, (the stretch of text where the items studied occur) are as follows:

**1. Partition:** One of the highest collocates in the lists was the preposition *of / de*. It indicates a quantity that is restricted to the limits of a larger group, eg. "he displayed many of the weaknesses outlined above", "...con un poco de aquel entusiasmado escepticismo suyo". The figures show that the partitive use of this type of quantifiers is slightly superior in English as compared with Spanish, being a common function in any case. Not all of them are examples of partition, eg. "me cuido mucho de contradecirlo".

**2. Comparison.** The use of comparative elements, especially *más / more* and also *como, menos, as, like, less, better, easier* and *longer* seems to be very common in the surroundings of these words except for *poca* and *pocas*. One of the resources used to express imprecise quantity is by means of comparison, by reference to another object or entity that acts as a sort of scale (Lamíquiz, 1991:93-114). There are some cases of idiomatic combinations, eg. "their populations turned out to be much the same", "Designing a garden is much like decorating a room", "pasado mañana o al otro como mucho" and "cualquier crítica es poco menos que una herejía".

**3. Intensification.** Another case could be thought of as being the union of two quantifiers: "a little bit" but it could also be considered a quantifier premodified by an adjective of size, just the same as "a great many" and "tiny little bit". Adjectives of quality are common as premodifiers as well, eg. "an awful lot", "pretty much" "bien poco", "a good few" and "a fair few". There is a difference, however, as refers paucal quantifiers: whereas "bien poco" suggests that it is less than *poco*, e.g. "Pero bien poco iba a durarme la euforia, porque apenas hubé dado el primer paso en dirección a la calle...", "a good few" and "a fair few" suggest more than *a few*, e.g. "when my children were smaller they used to go fishing there... But that's a good few years ago now". Similarly "a whole lot" and "a hell of a lot" are frequent

