

APPROACHING EPISTEMIC MODALITY IN ENGLISH AND SPANISH: A CORPUS-BASED CONTRASTIVE STUDY OF *PROBABLY* AND *PROBABLEMENTE*

Noelia Ramón García
University of León (Spain)

Abstract

Modality is a broad and fuzzy area and modal meanings are mapped in quite different ways cross-linguistically. This paper is an attempt to describe the expression of epistemic modality in English and Spanish by focusing on the most commonly employed epistemic adverbs in these languages: probably and its cognate probablemente. This paper reports on a corpus-based contrastive study of these adverbs, which analyses issues related to their syntactic function, placement, and the co-occurrence of these adverbs with other modal markers, mainly modal verbs in English and tense, mood and lexical verbs in Spanish. Empirical data are extracted from two large monolingual corpora of contemporary English and Spanish, Cobuild/Bank of English and CREA (Corpus de Referencia del Español Actual), in order to carry out the analysis on authentic instances of language in use. The results not only provide a description of the use of the two adverbs, but also a general picture of the different grammatical resources associated to the expression of epistemic modality in English and Spanish. The future applications of the results lie in areas related to interlinguistic communication involving the two working languages, in particular in foreign language teaching and translation.

Key words: English-Spanish contrastive analysis, epistemic modality, modal adverbs, corpus.

Resumen

La modalidad es un área semántica amplia y compleja que se representa gramaticalmente de formas muy diversas en las distintas lenguas. Este trabajo describe la expresión de la modalidad en inglés y español partiendo de los adverbios modales más frecuentes en las dos lenguas: probably y su cognado probablemente. El objetivo de este estudio es llevar a cabo un contraste de estos adverbios centrándonos especialmente en sus funciones sintácticas, su posición en la oración y su combinación con otros marcadores modales, sobre todo verbos modales en inglés y categorías como el tiempo o el modo verbales y unidades léxicas en español. El trabajo está basado en datos empíricos obtenidos de dos grandes corpus monolingües representativos de las lenguas inglesa y española en la actualidad, Cobuild/Bank of English and CREA (Corpus de Referencia del Español Actual). El estudio contrastivo proporcionará una descripción del uso de los dos adverbios en cuestión, y además también ilustrará la distribución de los medios gramaticales disponibles en inglés y español para expresar modalidad epistémico. Las aplicaciones de los resultados se orientan principalmente hacia actividades interlingüísticas como la enseñanza de lenguas y la traducción.

Palabras clave: análisis contrastivo inglés-español, modalidad epistémico, adverbios modales, corpus.

1. INTRODUCTION

This paper is a contrastive study of epistemic modality as it occurs in English and Spanish clauses in which the modal adverbs *probably* and *probablemente* act as reinforcement. Modality is a broad and fuzzy area and modal meanings are actualised in different ways cross-linguistically. Modal verbs are the most outstanding modality markers in English, whereas tense, mood and lexical verbs are used in various combinations by Spanish speakers to express modal meanings. *Probably* and its cognate *probablemente* are pragmatically inequivalent, as the English adverb is more common, particularly in spoken registers, whereas the Spanish counterpart is less frequent, and stylistically marked for written language mainly. Nevertheless, these adverbs share syntactic and semantic features, and this will be used to approach the expression of epistemic modality in their corresponding languages.

This contrastive study will follow a functional approach to the description of *probably* and *probablemente*, focusing on issues related to their syntactic function, placement, and the co-occurrence of these adverbs with other modal markers. The description will be based on

500 authentic language samples in each language extracted from two large monolingual corpora, Cobuild/Bank of English and CREA (Corpus de Referencia del Español Actual). The contrastive description will be especially useful in interlinguistic activities such as foreign language teaching and translation.

2. ADVERBS AND EPISTEMIC MODALITY

The grammatical category ‘adverbial’ is a heterogeneous group, not only in terms of form, but also in terms of meaning. Adverbials are generally optional, and are thus considered structurally peripheral. This peripherality implies a relative positional freedom of adverbials in the sentence. These factors are all interrelated. Different classifications have been put forward to describe adverbials. The adverbs analysed here, *probably* and *probablemente*, fall into the category of epistemic stance adverbials (or content disjuncts), as they indicate a level of doubt.

Epistemic modality “is concerned with the speaker’s assumptions or assessment of possibilities and, in most cases, it indicates the speaker’s confidence (or lack of confidence) in the truth of the proposition expressed” (Coates 1983: 18). *Probably* is the most common epistemic adverb in English, and *probablemente* one of the most common ones in Spanish (Kovacci 1999: 755). From a semantic perspective, these adverbs express doubt, but imply “the speaker’s reasonable certainty that the propositional content of what he is saying is in fact true. [...] *Probably* is the most common item and the strength of likelihood it conveys is best understood if it is contrasted with the much weaker adverbs *conceivably* and *possibly*.” (Hoye 1997: 194).

Many of the most common adverbs in English and Spanish are lexical cognates, and also share similar syntactic and semantic properties. But from a pragmatic point of view *probably* and *probablemente* are clearly inequivalent: “-ly forms have a much wider currency and acceptance in English than their Spanish counterparts ending in -mente.” (Hoye 1997: 258). In Spanish there are strong stylistic constraints in the use of all adverbs in -mente. Alternative constructions or verbal mood are more commonly exploited in Spanish for expressing meanings denoted in English by -ly adverbs (Butt & Benjamin 1988). With regard to register, *probably* occurs most often in spoken conversation. In contrast, *probablemente*, like all -mente adverbs, is located higher on a formality scale than its English counterpart, and thus occurs less often in the language, and more frequently in written texts. This pragmatic inequivalence, however, does not affect the semantics of the adverbs nor their modal surroundings.

Syntactically, *probably* and *probablemente*, like many single-word adverbs, may occur either as modifiers of other categories or as full adverbials. Like most adverbials, the two items selected for this study can be modified themselves by intensifiers such as *most* and *mu*y, respectively.

With regard to position in sentence structure, in the vast majority of cases *probably* occurred in medial position, usually immediately before or after the operator. There were only five cases of clause-initial position of *probably*, and in none of these cases was the adverb separated from the rest of the sentence by a comma: *Probably it was his imagination*. There was no single case of end position of *probably* in the 500 samples. Only in three cases did the adverb occur interpolated in the middle of the sentence, between commas (*You won't be able to forgive her, probably, for what she has done, but ...*), which means that there is a high degree of integration of this adverb in sentence structure. The Spanish adverb *probablemente* mostly occurred in medial position also: there was just one case of initial position and none of final position. However, in 79 cases the adverb occurred interpolated

between commas in medial position: *se darán a conocer, probablemente, la próxima semana en San Sebastián*. The pragmatic characteristics of the Spanish adverb make it less prone to be easily integrated in clause structure.

In English the co-occurrence of modal verbs and adverbs is “largely pleonastically” (Palmer 1979: 57), that is, with adverbs that express modal meanings themselves. *Probably* is a good example indicating a level of likelihood as perceived subjectively by the speaker. When the meaning of the modal adverb fits the meaning of the modal verb it goes with, the resulting collocation is considered modally harmonious (Lyons 1977: 807).

Spanish, on the other hand, lacks a category of modal verbs. Modal adverbs may co-occur with different moods and tenses, those indicating lower degrees of likelihood being more closely associated to the subjunctive and conditional moods and the future tenses, and those closer to certainty more prone to co-occur with the indicative mood, although most adverbs, including *probablemente*, may occur with any of these. In addition, Spanish speakers exploit a small set of fully-inflected lexical verbs, in various tenses and moods, to convey modal meanings. *Poder* and *deber* are especially relevant in epistemic modality.

3. DATA AND METHOD

This descriptive study based on authentic language in use follows a functional contrastive approach, and starting from a perceived similarity in the two languages – the expression of epistemic modality –, 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 fact that epistemic adverbs tend to be used to reinforce other modal structures in the language is used to attempt a description of the expression of epistemic modality in English and Spanish as it occurs when the adverbs *probably* and *probablemente* appear in discourse.

Two large monolingual reference corpora of contemporary English and Spanish were used to extract the necessary concordance lines to carry out this contrast: the Cobuild/Bank of English and the CREA (Corpus de Referencia del Español Actual). These two corpora differ somewhat in scope and size, but they were made comparable by exclusively selecting texts from between 1990 and 2004, and written (no oral texts were included) in British English and European Spanish. This limited our search to two sample corpora of roughly 30 million words each. Once the sample corpora were delimited, 500 instances of *probably* and 500 of *probablemente* were randomly extracted, analysed, classified and contrasted. The use of authentic language material in the two languages affords great reliability in the results obtained.

4. RESULTS AND DISCUSSION

Figure 1 shows the percentages of occurrence *probably* and *probablemente* as adverbials and as modifiers.

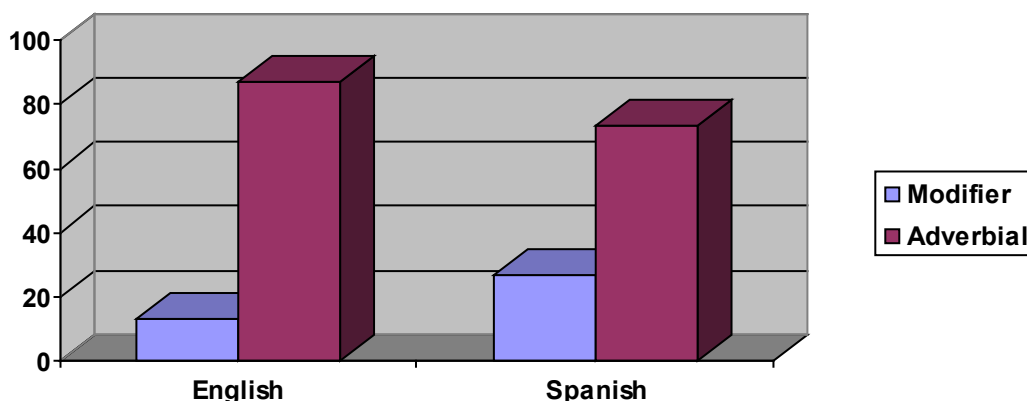


Fig. 1. Syntactic functions of *probably* and *probablemente*.

Probably occurs as a modifier in 13.2% of cases, and as an adverbial in 86.8% of cases. *Probablemente* is a modifier in 26.8% of cases and an adverbial in 73.2% of cases. There is a similar trend in the two languages with both adverbs mainly related to the expression of modality as adverbials. However, in Spanish there is a higher frequency of *probablemente* as a modifier. This may be related to the different types of grammatical means used in the two languages to actualise epistemic meanings.

The co-occurrence of *probably* and *probablemente* with other modal categories has been analysed and quantified, taking into account modal verbs in English and mood, tense and lexical verbs in Spanish. In our Spanish sample corpus there were four instances of a common non-standard use of a past perfect in the subjunctive mood instead of the perfect conditional. These cases have been left out of the study.

The analysis was applied onto 434 instances of *probably* as an adverbial. Figure 2 shows the co-occurrences of *probably* with modal verbs in our sample corpus:

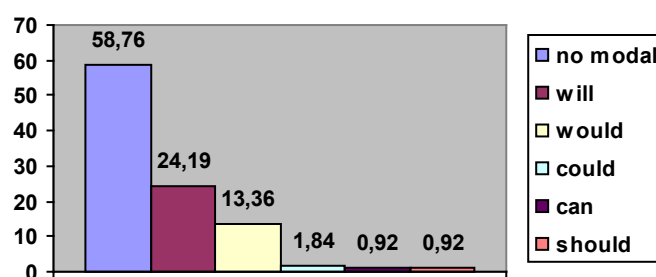


Fig. 2. Distribution of modality co-occurrence patterns for *probably*.

In almost 60% of cases, *probably* does not co-occur with any modal adverb at all, but is the only modal item in the sentence: *He was probably doing her a favour*. Figure 2 shows a clear trend towards the co-occurrence with the modal verbs *will* and *would*, thus indicating futurity or hypothetical meanings: *we will probably lose our house*; *Britain would probably have far fewer undergraduates*. The meaning nuance, however, is not a pure future, but involves also prediction, and the adverb confers the attitude of the speaker towards this prediction. In very few cases *probably* co-occurred with other modal verbs: *could* (*you could probably use them on rivers*), *can* (*it can probably occur elsewhere*), and *should* (*He should probably have been immediately replaced*). The meaning nuance provided by the adverb remains the same, even when the semantics the modal verb is non-epistemic, but deontic, as

in the case of *should*. In any case, *probably* shows a higher degree of collocability with *will* and *would* than with any other modal verb.

There were 362 instances of *probablemente* as an adverbial in Spanish. Figure 3 shows the distribution of *probablemente* as it co-occurred with modal markers involving tense, mood or lexical verbs in Spanish:

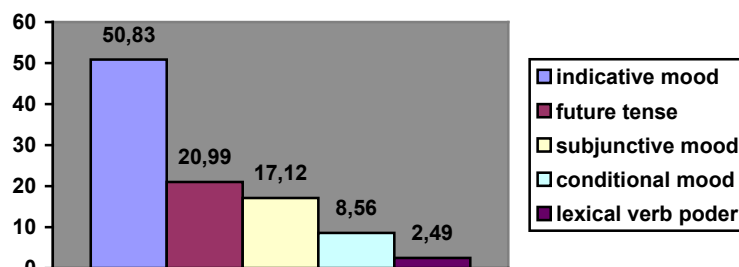


Fig. 3. Distribution of modality co-occurrence patterns for *probablemente*.

In 50.83% of cases, the adverb was the only modal marker: *es probablemente su última creación*. Additional modal resources combined with the adverb were mainly the future tense of the main verb in the clause (20.99%) (*sus expectativas se verán probablemente defraudadas*), and the subjunctive mood of the main verb (17.12%), which indicates non-factuality (*probablemente se encuentre en un error*). There were 8.56% of cases of the conditional mood as in *probablemente ya no estaría entre nosotros*, which also includes reference to futurity and hypotheticality. All present and past tenses are included under the headings for subjunctive and indicative. Finally, in our samples there were 2.49% of cases where *probablemente* co-occurred with the lexical verb *poder* followed by the indicative mood: *las firmas discográficas probablemente no pueden restringir a las divas que ...* A few additional cases of *poder* and *deber* occurred in the future or the conditional and are included under those headings: *un escritor que probablemente debería ser estudiado y leído*.

The data show that epistemic modality is mapped in very different ways on the grammars of English and Spanish. Figure 4 shows the distribution of the different modal markers found in the syntactic environment of *probably* and *probablemente*.

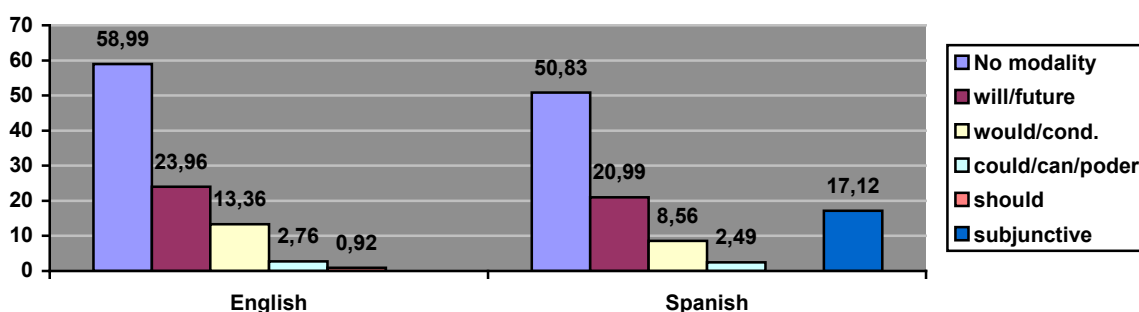


Fig. 4. Contrasting modality co-occurrence patterns for *probably* and *probablemente*.

The no-modality category groups instances in which *probably* did not co-occur with any modal verb and instances where *probablemente* occurred with a non-modal lexical verb in the indicative mood. Figure 4 shows that in English there is a slightly higher percentage of

cases where *probably* is the only modal marker. Co-occurrence with *will* and *would* in English has been compared to co-occurrence with non-modal lexical verbs in the future tense and the conditional mood in Spanish. In general, there is a similar trend in the two languages, with a slightly higher frequency in English. The two modals *could* and *can* may be compared with the occurrences of the modal lexical verb *poder* inflected in present and past tenses in the indicative. The type of modal meaning represented by the modal verb *should* is actualised in Spanish by the combination of a lexical verb, *deber*, and the conditional mood: *debería*. There is just one such case in our corpus (0.27%), but this overlap cannot be represented graphically. This case is included in the category ‘conditional’. The most outstanding difference between the two languages is the high number of cases in which the Spanish adverb *probablemente* co-occurs with a verb in the subjunctive mood. The subjunctive indicates non-factuality, often involving futurity and uncertainty too, and thus expressing meanings covered by the whole range of modal verbs found in the English sample corpus. The subjunctive mood as such is almost non-existent in Present-Day English grammar, and there was no case in our 500 sample concordance lines. The results show that when *probably* co-occurs with a modal verb expressing futurity, prediction or hypotheticality, similar nuances may be expressed in Spanish by adding the adverb *probablemente* to the main verb in the subjunctive mood.

5. CONCLUSIONS

Authentic language data were used in this paper to study the expression of epistemic modality in English and Spanish when the adverbs *probably* and *probablemente* were part of the modal force indicating degree of likelihood. Several conclusions may be drawn from this analysis. First, *probably* is more commonly used as an adverbial than *probablemente*, which is, in its turn, more commonly employed as a modifier than its English counterpart. For stylistic reasons, the Spanish adverb has been found to be less integrated in clause structure. Regarding co-occurrence patterns, the results indicate that English relies heavily on modal verbs for expressing epistemic meanings co-occurring with *probably*, in particular futurity and hypotheticality. In contrast, in Spanish, tense, mood and lexis are all involved in the process of expressing this type of modality in co-occurrence with *probablemente*. The category of modal verbs in English has grammaticalized networks of meanings that are distributed across several different grammatical categories in Spanish. The subjunctive mood in Spanish is the second most common modal category co-occurring with *probablemente*, revealing that in cross-linguistic relations the semantics of English modal verbs will often find functional-semantic equivalents in the subjunctive mood in Spanish.

The expression of modality is a broad and complex issue in grammar and further studies are needed to allow for the inclusion of additional modal markers in Spanish that may not co-occur with modal adverbs, such as adjectival constructions, for example, *es probable que ...*, *es posible que*, *seguro que ...*. I claim that small studies like this one are very useful in approaching modality and its mapping onto different languages.

REFERENCES

- Butt, J. and C. Benjamin. 1988. *A New Reference Grammar of Modern Spanish*. London: Arnold.
- Chesterman, A. 1998. *Contrastive Functional Analysis*. Amsterdam/Philadelphia: John Benjamins.

- Coates, J. 1983. *The Semantics of the Modal Auxiliaries*. London: Croom Helm.
- Hoye, L. 1997. *Adverbs and Modality in English*. London: Longman.
- Kovacci, O. 1999. "El adverbio". *Gramática descriptiva de la lengua española*. Eds., I. Bosque and V. Demonte. Madrid: Espasa. Vol. I: 705-786.
- Lyons, J. 1977. *Semantics*. Cambridge: CUP.
- Palmer, F.R. 1979. *Modality and the English Modals*. London: Longman.