Translating the ‘Predictive’ and ‘Hypothetical’ Meanings
English-Spanish

Rosa Rabadán

Résumé de l’article

Cet article vise à explorer les différents choix de traduction des modalités de la prédiction et de l’hypothétique dans le sens anglais-espagnol. La recherche proposée ici se fonde sur un corpus et combine l’analyse contrastive des techniques propres aux études descriptives. Deux corpus monolingues « comparables » permettront de déterminer si et dans quelle mesure le sens est porté par des formes perçues comme des équivalents interlinguistiques. Un corpus de traductions permettra de confirmer les « solutions » employées pour combler les différences. En résultera un inventaire de choix de traduction « corrects d’un point de vue descriptif », choix applicables dans diverses situations de traduction.
Translating the ‘Predictive’ and ‘Hypothetical’ Meanings English-Spanish

ROSA RABADÁN
University of León, León, Spain
dfmrра@unileon.es

RÉSUMÉ
Cet article vise à explorer les différents choix de traduction des modalités de la prédiction et de l’hypothétique dans le sens anglais-espagno. La recherche proposée ici se fonde sur un corpus et combine l’analyse contrastive des techniques propres aux études descriptives. Deux corpus monolingues « comparables » permettront de déterminer si et dans quelle mesure le sens est porté par des formes perçues comme des équivalents interlinguistiques. Un corpus de traductions permettra de confirmer les « solutions » employées pour combler les différences. En résultera un inventaire de choix de traduction « corrects d’un point de vue descriptif », choix applicables dans diverses situations de traduction.

ABSTRACT
The aim of this article is to explore the different translation options into Spanish for the ‘predictive’ and ‘hypothetical’ English modals. Contrastive analysis and DTS techniques are combined in this corpus-based research. Two monolingual ‘comparable’ corpora are used in order to establish whether and how these meanings are conveyed by forms perceived as being ‘cross-linguistic equivalents.’ A translation corpus contributes evidence to the ‘solutions’ provided to fill in the gaps. The result is an inventory of ‘descriptively correct’ translational options that can be used for different translation applications.

MOTS-CLES/KEYWORDS
predictive and hypothetical meanings; contrastive analysis, DTS, corpus-based research

1. Background and aim

The ultimate aim of the ACTRES’ project is to produce a translation application/aid to help translators quickly identify the inventory of correct and acceptable translational solutions available for a particular translation problem. The ACTRES team’s work concentrates on those areas of meaning or on those formal resources that are a source of recurrent problems when translating from English into Spanish (Rabadán, Labrador and Ramón 2004). The underlying language conception is Bondarko’s functional-semantic fields (1991) where one core meaning is represented primarily by one formal resource, which is the typical central one, and by a varying number of other expressive means which are more or less peripheral according to a typicality scale, a conception which can be seen as parallel to that of translational regularities/norms in DTS. This latter notion2 (Toury 1995) has been extremely influential in research directed towards unveiling particular translation practices by a given translator, in a given period, concerning a text form/type, etc. But it has done little to help good reliable translating, mainly because of the disregard for the concept of ‘linguistically correct translation’ coupled with the widespread habit of considering the language of translating as necessarily different from non-translated language. It is

Meta LII, 3, 2007
obvious, however, that the expressive resources the translator works with are those of the target language and if a certain amount of ‘creativity’ is often involved, translation solutions, particularly at the grammatical level, are to be found within the limits of the formal possibilities of the TL.

It is the aim of this paper to find the (correct) possibilities Spanish offers for the translation of the semantic functions encoded by WILL, SHALL and WOULD in English. These forms are approached mainly as modal markers in the English language grammatical tradition, whereas Spanish considers this area primarily as tense and mood marking. In order to map out the actual cross-linguistic correspondences between form(s) and meaning(s), a (double) corpus-based perspective is taken: first, two separate monolingual corpora, one English, the other Spanish, will be used as an input to our ‘comparable corpus,’ which will be the source of ‘correct’ options; second, a ‘parallel corpus English-Spanish’ will provide qualitative evidence on formally divergent translation solutions, which in turn will be verified by means of new ‘comparable corpus’ searches and analyses. The experimentation will follow ACTRES’s standard procedure: 1. identification of a problem area and reasons for it being so (selection phase); 2. description of meaning function and typical expressive resources both in English and in Spanish; 3. juxtaposition of expressive means and typicality scales in both languages; 4. diagnostic translation solutions as shown by the ACTRES parallel corpus, and 5. inventory of ‘descriptively correct’ translation solutions available.

2. The translation problem: mood, tense, and modal uses of mood/tenses English-Spanish

The interaction between modal meanings, modal forms and tense and/or mood marking and the way form(s) and meaning(s) are associated tend to be extremely language-specific (Larreya 2000), and it is particularly obvious in the case of English and Spanish. Using the criterion of ‘perceived cross-linguistic similarity’ WILL, SHALL and WOULD and the Spanish futuro and condicional tenses will be reviewed so as to establish the commonalities and the divergences between the resources of English and Spanish to convey ‘prediction’ and ‘hypotheticality’.

WILL and SHALL are taken to constitute the ‘volition and prediction cluster’ (Coates 1983: 169), as most of the meanings in the gradient fall within this semantic area. From the ‘prediction’ senses, which may amount to little more than future tense marking or include ‘intention,’ to those cases of strong volition in which an ‘obligation’ interpretation is called for, it is evident that these forms are used as a futurity marker and also in a range of non-factive modal meanings (Lyons 1977: 816; Haegeman 1983, Stage 2003; Sundell 2003).

When looked at from a cross-linguistic perspective, WOULD emerges as the general marker of ‘unreality’ in conditional utterances. Yet this is the typical context in which WOULD is unlikely to create any translation problem as the tense correspondences are highly regulated. Transfer problems mirror roughly those of the ‘future forms’ – from marking ‘unreal condition’ or ‘hypotheticality’ to encoding other more specialised meanings such as ‘future-in-the-past,’ ‘volition,’ ‘habit’ or ‘irrealis,’ there is a wealth of meanings which are generally attributed to formally different grammatical resources in languages other than English.
English does not feature specific mood-conveyor forms; rather it makes use of modal forms which include these notions among their various meanings (Palmer 2001: 201-202). The ongoing discussion about whether English has got a future tense marked by WILL and/or SHALL (Wekker 1976, Huddleston and Pullum 2002) is not relevant for our purposes, as independently of the answer to that question, this is not a problem trigger when translating from English into Spanish.

In Spanish the grammaticalization procedure has been the opposite: modal meanings are not clearly associated with a ‘central expressive resource’; rather verbal forms show a basic mood adscription to either indicative or subjunctive, and they can be affected by ‘temporal dislocation or displacement.’ This phenomenon accounts for the double functionality of certain verbal forms as conveyors of both temporal content and modal meanings when the primary time reference is challenged (Rojo 1974; Rojo and Veiga Rodríguez 1999: 2895). Thus, starting from a ‘present time’ form and context, Veiga Rodríguez (1991: § II. 3.1.) distinguishes five possible actualizations resulting from the interaction mood-modal meanings, including ‘uncertainty’ (the future indicative and the conditional), and ‘unreality’ (the conditional and the imperfect indicative and subjunctive). The ‘uncertainty’ of the simple future form tends to display a ‘conjecture/distinctiveness/predictability (confident probability)’ function, whereas the conditional form in this use can be seen as indicating the same meaning in the past. The Spanish imperfect tense is – in terms of its translation potential – a valuable multifunctional form, capable of conveying different types of meaning: past time (tense), duration (aspect) and irrealis (modality) (Gutiérrez Araús 1996). If combined with the subjunctive, the resulting imperfect tense conveys the meanings of both ‘hypothetical’ and ‘irrealis,’ competing with the conditional tense in this semantic area. This picture is still further complicated by the treatment given to these ‘displaced’ uses in the grammar.

Although the Spanish subjunctive is normally associated with syntactic subordination (Borrego, Asencio and Prieto 1985: 8-9), subjunctive forms may also appear in independent clauses expressing a range of modal meanings (Sastre Ruano 1997: 43-58). The precarious historical relationship between futurity and subjunctive in Spanish has been resolved by favouring the use of the indicative forms and making the subjunctive futures redundant and obsolete to the point of their virtual disappearance. The reason for this change lies in the overlapping of values between the future indicative and the subjunctive: being virtual and indicating future perspective at the same time makes some of the forms redundant (Sastre Ruano 1997: 38-42).

3. The tools

3.1. The comparable monolingual corpora: the Bank of English and CREA

In order to establish potential translational options we need real life observational data from which to derive our proposals. Our primary tools in ACTRES are two corpora, one comparable, i.e., made up of original texts in English and in Spanish, and a parallel corpus, i.e., a corpus containing original English texts and their translations into Spanish. The monolingual corpora are ‘source corpora,’ from which topic-specific smaller corpora are derived. They are two large, independent monolingual corpora, one English, the other Spanish, featuring an equivalent internal archi-
tecture concerning subcorpora, intralinguistic varieties and statistical dimensions. Corpus selection started from the very basic condition of availability in both languages. For Spanish, there is certainly a range of possibilities, although the one best suited to our purposes is the CREA (Corpus de Referencia del Español Actual http://www.rae.es), sponsored by the Real Academia de la Lengua Española; and this choice determined the selection of the English corpus, which for reasons of equivalence is Cobuild’s the Bank of English (http://titania.cobuild.collins.co.uk).

The Bank of English includes vast quantities of text (524 million words in October 2004) but only 56 million are available online through CobuildDirect. It comprises 12 different full-text subcorpora, selected according to geographical provenance, textual mode, and physical format. Although materials cover the last 20 years the majority of the texts originate after 1990. The corpus is constantly updated and tagged for morphological category. Search tools include wildcards, part-of-speech tags, word combinations, etc. The ‘actual corpus’ used here comprises over 30 million words.

CREA consists of approximately 148 million words (July 2004) and is also a full-text corpus available online. The criteria for text selection are comparable to those of the Bank of English: geographical provenance; textual mode; physical format; and an extra bonus: field area, which can prove immensely helpful in certain types of research (e.g., phraseology). CREA is also subject to continuous updating, keeping only those texts produced in the last 25 years (more representativeness is granted to the most contemporary materials), which means that regularly some of the older materials are moved to the diachronic CORDE database. Chronological restriction is a choice. The ‘actual corpus’ used in our sampling consists of approximately 37 million words. CREA’s query syntax also features wildcards but is much more restricted than that of the Bank of English, as part-of-speech tagging is not yet complete. This difference in software tools, together with typological and distributional differences have forced us to come up with a variety of searching strategies, all geared towards one unique aim – to extract ways of naturally expressing one meaning in English and in Spanish.

For purposes of corpus comparability, the subcorpora chosen are, for both ‘actual’ corpora, those comprising written texts: newspapers, magazines, books and ephemera. In both the Bank of English and CREA the chronological span has been selected by default and the language variety choice is in both cases the European one. The only criterion which is available and has not been used as a filter is what we have called “field/topic area,” as it is only applicable in CREA.

3.2. The parallel corpus ACTRES

The ACTRES parallel corpus has been modelled on the selection parameters of both the Bank of English and CREA so that the materials can be taken to be representative of the same areas of written language as those in our chosen comparable subcorpora. This means that the materials in the corpus belong to one of the following areas: newspapers, magazines, books and ephemera/miscellaneous. The chronological span dates from 1999 onwards (open) and is determined both by the date of the translation and that of the original English text.
TABLE 1
Corpora comparability and distribution

<table>
<thead>
<tr>
<th>Bank of English subcorpora</th>
<th>CREA subcorpora</th>
<th>Parallel ACTRES</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK books</td>
<td>Libros-España</td>
<td>Books</td>
</tr>
<tr>
<td>UK Today</td>
<td>Periódicos-España</td>
<td>Newspapers</td>
</tr>
<tr>
<td>UK Times</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK Sun</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK mags</td>
<td>Revistas-España</td>
<td>Magazines</td>
</tr>
<tr>
<td>UK ephemera</td>
<td>Miscelánea-España</td>
<td>Miscellaneous</td>
</tr>
</tbody>
</table>

Materials range from fiction to essays of all types to popular science for ‘books’; the ‘magazines’ category contains materials from a variety of sources, including *National Geographic* or *Cosmopolitan*; and ‘ephemera/miscellaneous’ includes beauty products leaflets, texts from airline internal publications, gadget ‘instructions for use’ texts, education institutions prospectus, tourist leaflets, etc. For copyright reasons parallel ACTRES is not a complete-text corpus. On the contrary, an average of 10% of the total textual material for each title finds its way into the corpus. Although the intended full size is 2 million words, in July 2004 the ACTRES corpus comprised 550,487 words evenly distributed between English and Spanish.5

Bank of English and CREA involve different search tools – more general or more delicate – which are built into the corpus interface and/or are available from the corpus site. For parallel ACTRES the chosen program is MULTICONCORD, a multilingual parallel concordancer developed mainly for language teaching purposes (http://artsweb.bham.ac.uk/pking/multiconc/lingua.htm).6

3.3. The informants

A further ‘tool’ in the course of this inquiry was a group of 10 informants. Their task was to provide native speaker information whenever necessary. Their sociological profile is ‘university educated speaker,’ ‘middle class,’ and their age range ‘28-45 years.’ Five of them have some variety of European English as their first language; the other five are native speakers of Castilian Spanish. All the informants have had some training in linguistic analysis, although only two in each group are professional linguists. In each subgroup there is one person that has access only to his/her language whereas the other eight can communicate both in English and in Spanish.

4. Procedure

The procedure followed consists of the following: i) starting from qualitative and quantitative evidence from the Bank of English, streamline translationally operative labels for each meaning function and analyze English data; ii) search CREA for qualitative and quantitative evidence and analyze Spanish data in terms of the labels already proposed; iii) verify degree of cross-linguistic overlapping and/or divergence in meaning functions – expressive means and iv) search parallel ACTRES for answers to Ø solutions.

This strategy has been adopted in an attempt to i) minimize the possible negative effects of the as yet non-representative dimensions of the parallel corpus, and ii) make
the most of the findings of descriptive and (where possible) corpus-based studies of non-translated language for our language pair. In practice this means that i) acceptable non-translated uses and their typicality of occurrence in both languages will be obtained from the ‘comparable’ corpora, ii) qualitative evidence concerning translational solutions for Ø slots will be collected by analysing 100 pairs from parallel ACTRES, and iii) correctness of these uses (other than lexical paraphrasing) will eventually be verified by analysing CREA materials in the corresponding grammatical area, though this step exceeds the limits of this paper and will be undertaken in a future phase of the project along with basic research into tense/aspect translation related problems.

4.1. Selection

4.1.1. Selection: English

Our choices for searching the Bank of English are marginal items within the modal system of English, in the sense that they are either 'low-degree' modals – namely items which are apparently modal verbs but which convey ‘little discernible modal meaning’ (Huddleston and Pullum 2002: 187-88) of their own – or are not basically modals, but have modal uses. WILL is considered as the ‘central’ form of the ‘prediction’ cluster displaying a number of futurity-related modal senses. Although used very restrictively, SHALL contributes at least one specialized modal sense to the ‘prediction’ area, as well as being an alternative to WILL in other modal uses. WOULD is the key item of the ‘hypotheccality’ area and is also associated with modal meanings (see section 2).

The distribution parameters of occurrences for both English and Spanish are mirrored in the statistics for the samples to be analysed. The size of the sampling was arrived at by applying the following formula: \( n = N / ((N-1)E^2 + 1) \) where ‘n’ is the sample to be analysed, ‘N’ the population, i.e., the total number of occurrences yielded by our searches, and ‘E’ the estimative error (5%).

<table>
<thead>
<tr>
<th>FORM</th>
<th>N</th>
<th>SAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>WILL</td>
<td>95,331</td>
<td>398</td>
</tr>
<tr>
<td>SHALL</td>
<td>1,886</td>
<td>330</td>
</tr>
<tr>
<td>WOULD</td>
<td>58,362</td>
<td>397</td>
</tr>
<tr>
<td>TOTAL</td>
<td>155,579</td>
<td>1,125</td>
</tr>
</tbody>
</table>

4.1.2. Selection: Spanish

Two main considerations have been taken into account when deciding what and how to select from the CREA corpus. The first concerns mood and simple/compound form choice, as this, in principle, could affect the type of meanings conveyed by the forms being analysed; the second has to do with querying strategies and types of information obtainable thereby.

Research into the modal and temporal values of verbal forms in Spanish has demonstrated that these values belong in the same meaning functions irrespective of
WHETHER WE ARE DEALING WITH A SIMPLY OR A COMPOUND FORM  

whether we are dealing with a simple or a compound form (Rojo and Veiga Rodríguez 1999: 2871). This has a direct implication for our analysis, i.e., analysing just the simple forms will yield a complete insight into the ‘displaced values’ of future and conditional forms. Furthermore, although mood selection has evident implications for our purposes, it is a well documented fact that the subjunctive future(s) have virtually disappeared’ from present day Spanish (Borrego, Asencio and Prieto 1998: 14-20) and that their functions have been taken up by subjunctive forms such as the ‘presente simple’ (Sastre Ruano 1997: 38-42) and the imperfect (Veiga Rodríguez, 1989; Palmer 2001: 209-216). Based on our proposed ‘principle of similarity perception’ this means that only the cross-linguistic resources perceived as correspondents will be mapped out initially. Possible ‘gaps’ in the ‘future-to-future’ grid will be addressed by means of parallel ACTRES data.

Obtaining comparable information when searching the Bank of English and CREA for ‘cross-linguistic equivalents’ requires different querying strategies. As Spanish futures and conditionals have a full conjugation paradigm, it is not possible to use a single lexical base to search the corpus, as was done with SHALL, WILL and WOULD in the English corpus. A second possibility is to use the wild cards available, but this strategy would render the total number of verbal forms – irrespective of tense, mood, etc. for that entry. An alternative is to adopt a ‘token-strategy’ – the three token verbs used as models for each of the conjugation paradigms (cantar, comer and escribir) can be searched in the corpus in all their forms for the futuro simple and the condicional. A further option contemplates using frequency lists and using, for example, the 10 most frequent lexical verbs as querying nodes, but the statistical information of CREA does not yet release global data; only the distribution by document and/or author of each particular form is accessible. However, frequency lists elaborated on the basis of the CORPUS DEL ESPAÑOL ACTUAL, have been made available for this inquiry.³ This corpus mirrors CREA’s parameter distribution and materials on a minor yet representative scale. So, a frequency list strategy was adopted and ten verbs were chosen at random among the top 100. Of these, the top 10 were discarded because of their basic auxiliary nature (ser, estar, haber…). Selected entries are creer, dejar, encontrar, tomar, perder, acercar, aparecer, escribir, comer, acordar. The smaller number of cases to be analyzed in Spanish does not preclude representativeness as it is due to the non–auxiliary nature of the Spanish future and conditional tenses. A search of the Bank of English, reversing the strategy and using frequency lists does not work either, as the distribution of meaning functions can be significantly constrained by the non-inflectional nature of the verbal markers. While it may not prove to be definitively conclusive in quantitative terms (no corpus-based study is), the information to be derived from the decision to use frequency lists in Spanish can be taken as ‘representative’ of the ‘state-of-the-art’ use of the future and the conditional in Spanish, and it will yield the basic meaning functions these verbal forms may convey.

It is also interesting to note the rather limited use of the conjugated future tense in Spanish as compared with other tenses (246 examples of futuro simple vs 3796 examples of presente simple, verb escribir). In the light of these data, and for the time being, it makes sense to hypothesize that there must be other forms of expression that have taken up part of the meaning functions normally associated with the formal future. Evidence from parallel corpus ACTRES will hopefully provide some verifica-
tion to support this hypothesis. The size of the sampling to be analyzed was determined by means of the formula given above. In order to ensure the representativeness of the sample, the distribution of cases to be analyzed among the different conjugated forms was arrived at by applying the regular simple proportional rule. For the futuro simple the total number of cases retrieved from the corpus is 7138, and the sample to be analyzed 379 cases. For the condicional simple (also Pos-pretérito) the total number of examples obtained from the corpus is 2806 cases, and the sample 350 instances of condicional.

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection statistics for futuro and condicional tenses</td>
</tr>
<tr>
<td>FORM</td>
</tr>
<tr>
<td>FUTURO SIMPLE</td>
</tr>
<tr>
<td>CONDICIONAL SIMPLE</td>
</tr>
<tr>
<td>TOTAL</td>
</tr>
</tbody>
</table>

4.2. Description

4.2.1. Cross-linguistic labelling

Although our semantic function labels are by no means original, as we draw on many different sources, the designations used here have been streamlined according to two major criteria. The first concerns meaning relevance from an English-Spanish cross-linguistic point of view; i.e., the semantic value is distinctive and has crucial consequences for translational decision-making. The second concerns the applicability of the labels from the point of view of the practitioner.

**Prediction [PDN]:** This basic value is taken here as the default meaning (Wekker 1976; Sundell 2003), also comprising modal notions such as ‘non-factuality, intention and willingness.’

- a.1. The great champions will be running in France and America,” said the little Scot. [PDN-FUT] (26)
- d.1. Roca dejará de ser portavoz, no el escaño. [PDN] (51)

**Conjecture [CJ]:** stands for those uses where the meaning is ‘hypothesis,’ ‘conjecture,’ ‘inference,’ also identified as futur épistémique (Stage 2003). In both languages the ‘inference’ conveyed by a future ‘is not based on a process of logical inference. Instead it is based on common sense, or on repeated experience’ (Coates 1983: 177).

- a.2. But people with outside experience will be reluctant to stand, and to suffer a big cut in income, penalising their families at a time of maximum expense. [CJ-WILL] (71).
- d.2. No sé si te acordarás de mí, pero el año pasado hiciste un comentario contando una anécdota y me gustaría saludarte. [CJ] (376).

**Contingency [CT]:** This label stands for those uses of ‘will/shall’ or ‘would’ where futurity or hypotheticality are not overtly explicit, but rather indicate that the happening is dependent on some external factor which is implied (Westney 1995: 198-200). This narrow definition seems to fit those English uses which roughly correspond to subjunctive forms in Spanish.
a.3. FAVOURITE Deep Sensation has been well backed after his 15-length defeat of Clay County at Nottingham, but I doubt that he will be able to cope with YOUNG POKEY. [CT-WILL] (9)
c.1. I doubt if Grandad would have kept in touch with his sister at all if it hadn’t been for Nan [CT] (356)

**Distinctiveness [DST]:** The crucial feature of this function is that it expresses different types of ’pure all-time reference’, such as habit, general and physical laws, characteristics of a person, a place, lifeless things, proverbial statements, etc. (Haegeman 1983: 22-23). A difference between will/future ’conjecture’ and will/future-’distinctiveness’ is that the former refers to a single situation or event, while the ’distinctiveness’ cases indicate a) a general law or truth or b) a series of reiterated events.

a.4. H. Plate Stacker – Protect your dinner service against chipping with this 4-tier rack in heavy plastic. Will hold up to 6 different pieces on each tier. [DST-WILL] (316)

**Obligation [O]:** This function has received different labels depending on the author: futur injonctif (Sundell 2003), futur déontique (Stage 2003). The cross-linguistic perspective used here calls for a unified account of all those uses ranging from a command or a prohibition to prescriptions and directions of some type. For the sake of applicability and terminological simplicity ‘obligation’ (Coates 1983: 190 and ff.) has been favoured over more technical options (Davidsen-Nielsen 1990: 161).

b.1. A journalist shall not accept bribes nor shall he/she allow other inducements to influence the performance of his/her professional duties. [O] (321)
d.3. Bajo el título No tomarás el nombre de Dios en vano, el pasaje más controvertido del artículo hace la siguiente invitación a los católicos catalanes: … [O] (204)

**Volition [V]:** Although examples indicating ‘request,’ ‘offer,’ ‘invitation’ (Haegeman 1983: 22) or ‘addressee’s volition’ (Coates 1983: 186) are not always regarded as having a meaning distinct from ’obligation’ (see above), it makes sense to postulate a further meaning when they are considered from an English-Spanish cross-linguistic point of view. Basically, this function appears in the contexts, either future or hypothetical, of attenuated requests (Leech 1987: 126-7), wishes, polite imperatives, etc. Both in future and hypothetical contexts the presence of an explicit conditional structure (of whichever type) does not seem to be decisive in the distribution of semantic functions either in English or in Spanish.

a.5. You will keep the secret for a bit, won’t you? [V] (263)
c.2. And we have kept our word by reuniting you with your daughter. We would now expect you to do the same. [V] (341)
e.1. Pero don Fermín parecía inclinado a satisfacer previamente un más perentorio apetito. Y así, lo hizo saber diciendo: -Yo comería algo. [V] (341)

**Hypothetical [HY]:** It is the default value in this area of meaning, and as such covers a gradient going from being a marker of uncertainty or unreal condition (hence condicional tense in Spanish) to hypothetical prediction, intention, etc. (Coates 1983: 211, 213). Whether and how these potential distinctions are differently grammaticalized in our two languages is vital information for mapping out a pool of correct translational solutions to choose from.

c.3. It is all about confidence and a victory over the Springboks would provide a massive injection of that [HY-P] (16)
e.2. Entonces te encontrarías con tu hijo sumido en el ostracismo, sin compañeros de juegos. Y esto puede acarrear problemas más importantes. [HY-P] (167)

**Future-in-the-past [FUP]:** This value indicates an event or happening foreseen in the past that did take place. This use exists both in English and in Spanish associated with would/ condicional forms and seems to be typical of narrative. According to Leech (1987: 53-54) this meaning is equivalent to a Past Tense and it can be paraphrased by ‘was/were destined to.’

c.4. Made available only in 1989, they cast new light on Paul’s commitment to what would become Gaudium et Spes. [FUP] (335)

e.3. Pero poco después fue detenida por los nazis y deportada a Bergen-Belsen, donde encontraria la muerte. [FUP] (125)

**Habit-in-the-past [HBP]:** Cross-linguistic descriptive data suggest considering ‘habit in the past’ separately (see note 11) so as to reflect the different expressive choices in Spanish: the condicional can convey ‘hypothetical action typical or characteristic of someone/something,’ but not ‘habit –in-the-past.’

c.5. From that day on, whenever Coward came to London, Greenwell would go round to the Savoy before the performance. [HBP] (145)

**Irrealis [IRR]:** According to Palmer (2001: 18ff.) there is no binary contrast realis/ irrealis, but it is a useful concept, as “it links modal systems to mood in the overall category of modality.” However, it is also true that neither modal ‘would’ nor any mood choice in Spanish can be said to have as their primary use to mark off the meaning ‘irrealis.’ In our taxonomy ‘irrealis’ means ‘past unreality’ (Fleischman 1995: 524), and will only apply to unreal events that did not happen and will not happen in any factual or hypothetical universe.

c.6. Frank would have been proud at the turnout. He would have loved all the attention [IRR] (226)

e.4. Me explico, es el encargo de un editor a un viejo escritor antifranquista y tengo que contar la vida de Franco como la escribiría el propio Franco. Es una obra para los jóvenes del año 2000 porque el editor piensa que ninguno sabrá nada sobre Franco. [IRR] (320)

4.2.2. Description: English

In our corpus WILL is the most frequent of the future encoding resources in English (87.68%), and it is associated with the basic notion of ‘prediction.’ ‘Volition’ cases add up to 1.75% of our corpus and the context is in all cases one of polite request, attenuated command, etc. ‘Distinctiveness’ is also represented in our corpus (4.02%) and it expresses habit, general and physical laws, characteristics of a person, a place, lifeless things, proverbial statements, etc. ‘Conjecture’ (0.75%) is a rather infrequent value and constitutes a separate group as its meaning is ‘deduction’ based on common sense, or on repeated experience (Coates 1983: 177). ‘Contingency’ instances of WILL account for 5.27% of the cases and will typically suffer a shift in both verbal mood and tense when translated into Spanish. The ‘obligation’ value appears just twice (0.50%), and in both cases it means ‘prescription.’ In this function the future is roughly equivalent to an imperative or to the central resources of the ‘necessity’ modal cluster.
Selection data show that SHALL is a low frequency form in our English language corpus (1886 examples) and it makes sense to hypothesize a small number of highly specialized semantic functions. Semantic labelling for SHALL values along these lines has produced four basic functions. ‘Prediction’ is the default meaning (71.81%), and it applies whenever the basic meaning of prediction is embedded in the indication of future time. ‘Obligation’ emerges as a typical function of ‘shall’ (17.87%), particularly in formal statutory or regulatory textual environments. A further function of ‘shall’ is ‘volition’ [V] (10.30%), which appears in contexts where the addressee’s wishes and/or opinion are consulted. In this use, the temporal meaning of projection into the future is neutralized, as the action targeted is marked for present time. Table b shows the distribution of the uses of ‘shall’ according to the semantic functions which seem to be relevant for translation from English into Spanish.

<table>
<thead>
<tr>
<th>FUNCTIONS</th>
<th>CASES</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREDICTION</td>
<td>349</td>
<td>87.68</td>
</tr>
<tr>
<td>CONTINGENCY</td>
<td>21</td>
<td>5.27</td>
</tr>
<tr>
<td>DISTINCTIVENESS</td>
<td>16</td>
<td>4.02</td>
</tr>
<tr>
<td>VOLITION</td>
<td>7</td>
<td>1.75</td>
</tr>
<tr>
<td>CONJECTURE</td>
<td>3</td>
<td>0.75</td>
</tr>
<tr>
<td>OBLIGATION</td>
<td>2</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>398</td>
<td></td>
</tr>
</tbody>
</table>

The ‘hypothesis senses’ are the meaning(s) displayed by 73.8% of the examples of WOULD in the corpus and they include hypothetical prediction, intention, willingness, etc. Quantitative data for the ‘future-in-the-past’ meaning yield a 7.3% share of WOULD cases. This function signals something that did actually happen in the past, and seems to be largely restricted to historical narrative. ‘Volition’ represents 6.54% of the examples and is associated with contexts such as attenuated requests, polite imperatives, wishes, etc. ‘Habit-in-the-past’ is the function of 5.54% of the cases in our sample. For the sake of congruence, and because of their ‘translation behaviour,’ cases which convey a ‘hypothetical action typical or characteristic of someone/thing’ have been considered as ‘hypothetical’ in the statistics. ‘Irrealis’ amounts to 4.28% of the examples in the corpus sample. ‘Contingency’ is a low frequency function in the hypothecality area (0.50%), and it tends to appear in utterances where the semantics and pragmatics of the main clause call for a subjunctive form in the target language.
4.2.3. Description: Spanish

The analysis of the Spanish futuro simple has revealed that the inventory of meanings conveyed by these forms is consistently much more restricted than the possibilities of their ‘perceived’ English counterparts.

The ‘prediction senses’ account for 97.62% of the examples in the sample, thus corroborating the fact that in Spanish these forms mark primarily ‘future time,’ and are generally associated with the relative (un)certainty of ‘plans and predictions.’ ‘Obligation’ (1.58%) appears both in prescriptive and regulatory texts and in common language. The meaning labelled ‘conjecture’ is not a central value of these forms, as it only yields 0.79% of cases in the CREA sample.

The Spanish condicional simple emerges from our translation-oriented analysis as a multifunctional tense, able to cover a wide range of meaning nuances. On the basis of quantitative results, two of these functions appear to be more central (in Bondarko’s terminology) and therefore closer to the core meaning associated with these forms. The rest seem more peripheral, which may indicate that they alternate with other formal resources to express these meanings.

The ‘hypothetical senses,’ which constitute the core meaning of the conjugated conditional form(s), represent 91.71% of the sample. The data corroborate the hypothesis that the Spanish condicional is basically a tense marker. The ‘future-in-the-past’ is the second most frequent use (7.14%). Our sample has yielded one single case meaning ‘characteristic’ and, as in English, it will be considered as ‘hypothetical.’ ‘Irrealis’ is a low-degree function of the condicional simple (0.85%) in contexts where it can alternate with subjunctive forms and ‘volition’ (0.28%) is mostly restricted to contexts indicating ‘wish.’
4.3. Juxtaposition English-Spanish

Juxtaposing the results from our comparative corpus analysis has shown the following degree(s) of overlap between English and Spanish resources for the semantic functions identified in the description of both areas, 'prediction' and 'hypothealcy':

The quantitative analysis of our 'prediction' materials yields the different typicality of occurrence that the meaning functions have in both languages. The 'prediction' meaning area is by far the most frequent both in English (87.68% in WILL cases and 71.81% in SHALL examples) and in Spanish (97.62%). 'Obligation' yields a 17.87% frequency rate in SHALL cases and a modest 0.50% in WILL uses, which seems to indicate a specialization of SHALL in this area. Spanish 'obligation' examples amount to 1.58% of uses in the corpus. 'Conjecture' is present in both the English (0.75%) and the Spanish (0.79%) corpora and they show practically equivalent typicality rates. 'Volition' is found in 1.75% of SHALL cases and 10.30% of the WILL-corpus; 'distinctiveness' is the meaning of 4.02% of the cases in the WILL-corpus, and 'contingency' is found in 5.27% of WILL-cases. However, there is no evidence of these functions in the Spanish futuro simple corpus.

The juxtaposition of the data for the 'hypothealcy' area makes it clear that there is a considerable degree of cross-linguistic overlapping in the meaning functions expressed by WOULD forms in English and the condicional simple in Spanish. Still, two 'gaps' have been identified in the Spanish corpus: no evidence has been obtained that the function 'habit-in-the past' is ever conveyed by means of conditional forms (5.54% rate in the English corpus); no case of 'pure contingency' (0.50% in the English corpus) has been identified either. Both languages can express 'hypothesis' by means of WOULD/condicional and the typicality rate, 75.8% in English and 91.71% in Spanish, indicates that in both languages this is by far the most frequent value.
‘Future-in-the-past’ shows virtually the same typicality rate in both languages, with English yielding 7.30% and Spanish data revealing a frequency of use of 7.14% of cases. ‘Volition’ accounts for 6.54% of the English sample, as against only 0.28% in Spanish, and ‘irrealis’ corresponds to 4.28% of the examples in English but only 0.85% in Spanish.\textsuperscript{12}

<table>
<thead>
<tr>
<th>WOULD</th>
<th>FUNCTIONS</th>
<th>CONDICIONAL SIMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>75.8%</td>
<td>HYPOTHESIS</td>
<td>91.71%</td>
</tr>
<tr>
<td>7.30%</td>
<td>FUTURE-IN-THE-PAST</td>
<td>7.14%</td>
</tr>
<tr>
<td>6.54%</td>
<td>VOLITION</td>
<td>0.28%</td>
</tr>
<tr>
<td>5.54%</td>
<td>HABIT-IN-THE-PAST</td>
<td>Ø</td>
</tr>
<tr>
<td>4.28%</td>
<td>IRREALIS</td>
<td>0.85%</td>
</tr>
<tr>
<td>0.50%</td>
<td>CONTINGENCY</td>
<td>Ø</td>
</tr>
</tbody>
</table>


The ‘gaps’ detected in the analysis lead us to hypothesize the existence of other non-future marked formal resources in Spanish that have specialized in expressing these meanings. In order to gain an insight into the translation options for these ‘gaps,’ evidence has been obtained from the parallel ACTRES corpus. A sample of 100 pairs for each semantic area has been analyzed for meaning and translation option(s) into the target language with the following qualitative results:

4.4.1. Prediction meanings

Examples featuring the core ‘prediction’ function are normally translated by a straight \textit{futuro simple}, which in our written corpus materials is the most frequently favoured solution followed by the \textless ir a + inf\textgreater option. ‘Obligation’ is always translated in our sample by a straight \textit{futuro simple}. ‘Volition’ and ‘distinctiveness’ cases are translated by a Spanish \textit{presente} (i.e. Simple Present tense), with the latter also being translated by the aspectual periphrasis \textless soler + inf\textgreater. ‘Contingency’ is encoded by means of a subjunctive present, as is ‘conjecture,’ although this is contrastive, not ‘diagnostic,’ evidence.

4.4.2. Hypothetical meanings

In our ‘diagnostic’ sample the ‘hypothetical senses’ examples are translated by a straight \textit{condicional simple}, as are ‘future-in-the-past’ cases, which alternatively use a \textit{pretérito} (simple past tense). ‘Volition’ is consistently translated by means of a Spanish \textit{presente} and ‘Habit-in-the-past’ by an \textit{imperfecto}.\textsuperscript{13} ‘Irrealis’ examples are all translated by either an \textit{imperfecto}, or a ‘subjunctive past perfect,’ and ‘Contingency’ cases are expressed in the Spanish translations by means of a subjunctive \textit{imperfecto}.

The ‘diagnostic’ information obtained from the ACTRES corpus helps fill in the ‘gaps’ obtained from the first (contrastive) analysis, thus contributing real transfer data that greatly improve our understanding of the problematic areas. In the ACTRES program these ‘diagnostic’ data will also be used as querying input for subsequent verification of translational (grammatical) options in CREA.
5. Translation options into Spanish: Results and conclusion

The analyses carried out here show that English and Spanish grammars handle the problem of mood/tense/modality overlapping differently. An ‘instruction-like’ formulation will favour the direct applicability of the results. For the sake of clarity, the results are also visually presented in tables h and i below. Whenever possible the Spanish solutions have been identified using the most widely accepted labels in the Spanish-speaking world; mood is only stated when other than the indicative.

5.1. Translating WILL and SHALL + infinitive.

i) The ‘prediction-futurity’ meanings can be translated by a straight futuro simple or by the periphrastic future <ir a + inf> option.

ii) ‘Obligation’ is always conveyed by a straight futuro simple.

iii) When the meaning is ‘conjecture’ the translation solution in Spanish is a futuro simple.

iv) ‘Distinctiveness’ is commonly expressed in Spanish by means of a presente, or the aspectual periphrasis <soler + inf > in the present tense as well.

v) ‘Volition’ is conveyed by a Spanish presente.

vi) ‘Contingency’ is encoded by means of a subjunctive present.

| TABLE H |
| Translation options into Spanish for the ‘prediction’ English forms |

<table>
<thead>
<tr>
<th>WILL</th>
<th>SHALL</th>
<th>FUNCTIONS</th>
<th>TRANSLATION OPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>√</td>
<td></td>
<td>PREDICTION SENSES</td>
<td>FUTURO SIMPLE</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>OBLIGATION</td>
<td>FUTURO SIMPLE</td>
</tr>
<tr>
<td>√</td>
<td></td>
<td>CONJECTURE</td>
<td>FUTURO SIMPLE</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>DISTINCTIVENESS</td>
<td>PRESENTE</td>
</tr>
<tr>
<td>√</td>
<td></td>
<td>VOLITION</td>
<td>PRESENTE</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>CONTINGENCY</td>
<td>PRESENTE SUBJUNTIVO</td>
</tr>
</tbody>
</table>

5.2. Translating WOULD + infinitive.

vii) The ‘hypothetical senses’ are regularly conveyed by a straight condicional simple in Spanish.

viii) There are two regular translation options for ‘future-in-the-past’: a straight condicional simple or a pretérito (simple past tense).

ix) ‘Volition’ is conveyed by a Spanish presente or a condicional simple.

x) ’Habit-in-the-past’ is always encoded by means of an imperfecto.

xi) ‘Irrealis’ can be translated by an (indicative or subjunctive) imperfecto, a ‘subjunctive past perfect’ or a condicional simple.

xii) The only translation option for ‘contingency’ into Spanish is a subjunctive imperfecto.
Translation options into Spanish for the ‘hypothecality’ English form

<table>
<thead>
<tr>
<th>ENGLISH ST</th>
<th>FUNCTIONS</th>
<th>TRANSLATION OPTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYPOTHESIS</td>
<td>CONDICIONAL SIMPLE</td>
<td></td>
</tr>
<tr>
<td>FUTURE-IN-THE-PAST</td>
<td>CONDICIONAL SIMPLE</td>
<td></td>
</tr>
<tr>
<td>VOLITION</td>
<td>PRESENTE</td>
<td></td>
</tr>
<tr>
<td>HABIT IN THE PAST</td>
<td>IMPERFECTO</td>
<td></td>
</tr>
<tr>
<td>IRREALIS</td>
<td>IMPERFECTO</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IMPERFECTO SUBJUNTIVO</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PLUSCUAMPERFECTO SUBJUNTIVO</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CONDICIONAL SIMPLE</td>
<td></td>
</tr>
<tr>
<td>CONTINGENCY</td>
<td>IMPERFECTO SUBJUNTIVO</td>
<td></td>
</tr>
</tbody>
</table>

This ‘de-centralization’ of meanings in a rather extensive inventory of forms is of extreme interest for translational applications, as is the possibility of using the results in practice directly as a pool of correct solutions for one given translation problem. Previous application-oriented analyses, including the seminal work by Vinay and Darbelnet (1977) for English-French, Vázquez-Ayora (1977), López Guix and Wilkinson (1997) and Whitley (2002) for English-Spanish, offer useful contrastive information done on the basis of ‘acquired knowledge.’ Other recent proposals, such as Álvarez Lugris (2001) and Santos (2004) take a corpus-based perspective using exclusively parallel corpora English-Galician and English Portuguese respectively. The originality of our proposal resides in the double-tiered, corpus-based experimentation. Translation materials from the parallel corpus are used as ‘diagnostic’ data and serve two functions: identify ‘grammatical transfer’ resulting in ‘translationese’ and provide formally dissimilar translation equivalents to the area being targeted. Empirical data obtained from the original languages corpora supply vital information concerning the differences between the two languages, and using non-translated language corpora neutralizes the problem of validation of the contrastive analysis results. A further innovative contribution is the cross-linguistic semantic labelling. These labels avoid ‘universal’ categorizations that tend to reflect the structure of English, but not that of other languages, and they can also be used directly by practitioners as an accessible analytical tool together with the inventory of ‘solutions available.’

It is obvious however that more work is needed into the transition from the implications to the actual application(s) of results. While users’ attitude can be described primarily as normative, in that they expect to be supplied with ‘ready-to-use’ tools, the researcher’s duty is basically that of a facilitator, as it is his/her role to produce readily available answers to the problems encountered. Work to turn results into a computerized tool compatible with state-of-the-art translation aids (e.g. TMs) is already underway.
NOTES
1. The Spanish acronym stands for ‘Contrastive analysis and translation English-Spanish.’ Research funded by project BFF-2001 (MCyT and FEDER). I wish to thank Roda P. Roberts and an anonymous reviewer for their helpful comments.
2. Only the ‘classical’ view of norm as ‘regularities of translational behavior’ will be considered here. Other approaches and/or redefinitions such as Chesterman’s (1997) or Herman’s (1999) tend to be embedded into their respective models and are therefore not very productive for our aims. Baker’s reinterpretation of norm as typicality (1993: 239-40) brings the concept conveniently closer to corpus-based studies, a conception which is not to be mistaken for her ‘universals of translation’ (Baker 2001). On ‘universal laws of translation behaviour’ see Toury (1995) and Mauranen and Kujamäki (2004).
3. The RAE postulated a new mood, potencial, in its 1931 edition. The Esbozo (1973) allocates these tenses to the indicative mood. Alarcos Llorach (1994: §§ 234 and ff.) includes the future tenses and the conditional tenses in what he labels modo condicionado. This distinction has not made its way into the verbal paradigm.
4. At the time of going to the press CREA had 160 million words (latest update April 2005). All our quantitative data refer to the situation when researching for this paper.
5. As of September 2005 the corpus contains over 1,800,000 words between both languages. Statistics in this paper refer to the situation at the time of the analysis reported here.
6. For the pairs used in this article. At the time of going to the press, all corpus materials are being aligned (and re-aligned) using the TCA (Translation Corpus Aligner). Searches are done by means of WebTCE (Translation Corpus Explorer for the Web). See http://khnt.hs.uib.no/files/align3.pdf and http://www.hf.uff.br/iba/prosjet/ENPC/manual.html#_Toc445194185 [10/11/2005]. Thanks are due to Knut Hofland (University of Bergen, Norway) for his invaluable help.
7. CREA yields a total number of 5 concordances for the subjunctive future dejare when searched within the same restriction parameters as its indicative counterparts. Compare this figure with those obtained for the indicative future (1753) in this same section.
8. The 50 most frequent verbs would be split into 3 groups as follows: the first 20 would be discarded as they are bound to include auxiliaries and ‘empty’ verbs, and the last 20 would not be considered either, as a way of marking some type of symmetrical boundary. These would leave us with 10 high frequency verbal entries. This strategy was suggested to me by Raphael M. Salkie (University of Brighton, UK).
9. All relevant information concerning this corpus can be found at http://www.sintx.usc.es/EspArthurs.html. I am grateful to Guillermo Rojo (University of Santiago de Compostela and Real Academia Española de la Lengua) for offering and providing the frequency list on the evidence of the CORPUS DEL ESPAÑOL ACTUAL.
10. It is advisable to consider a rather large chunk of text to accurately identify this function. Otherwise it may be difficult to distinguish it from a straight hypothetical prediction. Substituting the WOULD form by a simple past tense and/or paraphrasing it by ‘was/were to become, was destined to’ (Leech 1987: 53-54) can help.
11. There is agreement among the informants that it seems to be a (practically) fossilized idiomatic use <e.5. No pasó de ahí ni el cambio de simple y premeditada buena intención. Genio y figura, no le falta ninguno. Qué bollo es vivir, que diría / escribiría Tono. Qué bollo. [HY-CH] (332)>
12. It has to be taken into account that the search in the English corpus also includes WOULD HAVE forms, whereas in Spanish only the condicional simple tense has been considered.
13. Our ACTRES sample has yielded one isolated case of WOULD meaning ‘inference’ which has been rendered into Spanish by means of <deber de + inf>. It has not been considered separately because it is not relevant from the cross-linguistic viewpoint, as the translation choice is still an imperfecto. Besides, there is no evidence of this function in the English language sampling and it seems logical to assume that the incidence of ‘would-inference’ is extremely low <p.13. Father Martin had been on the staff when he last visited but must have retired long ago; he would be eighty by now.//Aunque el padre Martin seguía allí cuando visitó el lugar por última vez, sin duda se habría retirado ya; debía de tener ochenta años. (74)>.

REFERENCES


Real Academia Española de la Lengua (1973): Esbozo de una nueva gramática de la lengua española, Madrid, Espasa Calpe.


