

Analysis of the Translation Quality of the WordMagic English ↔ Spanish Interpreter Professional

OLGA BEZHANOVA

McGill University, Montreal, Canada

This paper presents an effort to analyze the capacities of the new and improved English-Spanish Interpreter Professional by WordMagic¹, which is now capable of performing translations from English to Spanish and from Spanish to English.

I have conducted a series of tests aimed at establishing the accuracy of translations in both directions offered by this system. To this purpose I have used texts provided by different news sources on the Internet. In particular, I used the system to translate recent articles posted on the website for the Spanish newspaper *El mundo*, as well as news items from www.cnn.com.

My selection of texts to be translated by English-Spanish Interpreter Professional included articles on various topics, which allowed me to see how well the system manages to translate texts with different types of vocabulary. At the same time, the texts that I chose contain sentences in different grammar tenses, as well as some more complicated expressions that it is usually quite difficult to translate for a MT system. I am going to present the results of my analysis according to the accuracy of the translation of

- a) vocabulary;
- b) grammar.

This way of organizing the results of the analysis will make it easier to appreciate what kinds of texts the system will be able to translate most successfully. At the same time, the prospective users of WordMagic products should bear in mind that the news items I have chosen for the test represent a kind of material that is quite complicated, due to the particular style commonly adopted by newspaper articles and other news communications.

I. SPANISH-ENGLISH TRANSLATIONS

Three different texts have been translated from Spanish into English by the system: an article about a political demonstration demanding peace for the Basque country, an account of an interview with the writer Umberto Eco, and a book review. These texts offer a wide variety of vocabulary and grammar constructions and, thus, allow to test the system more effectively.

A) Vocabulary

As for the vocabulary, it is obvious that the system in question is very well-equipped to translate Spanish texts dealing with politics. The first of the articles I have translated using the system contains a significant number of words used to describe the Spanish system of government. Such words as *ayuntamiento*, *alcalde*, *policía municipal*, *casa comistorial de la villa* were translated correctly allowing for a good understanding of a text made more complicated by the use of such terms.

The quality of the translation suffers in the instances where the system encounters verbs, which are always used with certain prepositions forming one lexical unit. The following expressions can be given as examples of this phenomenon:

- a) *respaldado por* is translated word by word, instead of as one fixed expression. As a result, the translation offered is *backed for*,
- b) *unido a* (in the phrase “cultura unida a internet”) fails to be translated as *united with*.

The original and the translation look as follows:

El único antidoto contra la esquizofrenia de una cultura cada vez más unida a internet son los libros.

The one and only antidote against the schizophrenia of more and more culture united they are Internet books.

It is obvious that the quality of the translations provided by the system could be improved by adding to the dictionary verbs that are always used with certain prepositions.

In the texts that deal with common interest stories, almost no words are left untranslated, which allows to understand the general meaning of the text. At the same time, the finer shade of meanings get lost and the sentences sometimes become confusing due to the system’s tendency to choose the first meaning for a word to be translated, which is not necessarily the most common one. To give an example, the verb *descubrir* is translated as *unveil*, although translating this word as *discover* would suit a greater

number of texts, due to the fact that this word is encountered more often in general interest texts. As a result of using a rarer synonym, the meaning of the phrase gets much less precise:

Los jóvenes que viajan por la red descubren nuevas curiosidades.
Young people that travel for the net unveil new curiosities.

While this particular mistake seems to be rather minor, there are more of the same order that create a much greater confusion as to the meaning of the text. Thus, the system tends to fail to recognize certain set expressions, which could also be entered directly into the system' dictionaries in order to make the quality of translation better. For example, in a set expression *ante este argumento* the first word is perceived by the system as a noun and translated accordingly:

Ante este argumento. - *Elk this argument.*

In the same manner, in the phrase *el tipo de información* the word *el tipo* is translated in its very colloquial meaning, which would only very rarely appear in printed text that is to be translated by an MT system.

El tipo de información. - *The fellow of information.*

It would seem more useful to relegate this translation of the word to a secondary place in the dictionary, while assigning the word *type* as the first-choice translation for the Spanish word *tipo*.

A similar problem arises with the verb *abarcar*, which is not recognized by the system as a verb, but rather as a noun with a very rare meaning:

Pero la crítica del escritor no solo se queda en el mundo de las nuevas tecnologías, sino que también abarca el tipo de información . . .
Bul the writer' critic not only remains in the world of new technologies but also wooden shoe, the fellow of information...

While this mistake might be due to a grammatical error made by the system that confuses a noun and a verb, coupled with the incorrect rendering of the phrase *el tipo de información* it makes this whole sentence appear senseless. Keeping in mind that *abarca* as a noun is encountered on extremely rare occasions, while the verb *abarcar* is quite common, it might be advisable to remove the translations of the word as a noun altogether if it will result in more precise translations.

Although, as it has been pointed out, some set expressions and commonly used words were not translated correctly, there has also been a significant number of frequently encountered expressions, which were translated perfectly. To give but a few examples:

- sin calcular los riesgos - *without counting the cost*
- sin previo aviso - *without previous warning*
- aprovecha sus amplios conocimientos históricos - *makes good use of his ample historic knowledge*

In general, it can be said that vocabulary-wise the Spanish-English Interpreter affords high-quality translations for texts relating to politics and current news, while more general interest texts, which abound in set expressions, are not translated quite as well. The confusion produced by the fact that the system sometimes fails to recognize such expressions and translate them adequately results in an outcoming text parts of which it is difficult to understand. At the same time, if the general dictionary that the system uses is amplified, the quality of the translation will become much better.

B) *Grammar*

The texts that were used to test the system allowed me to see how well it recognizes and translates different grammar tenses and grammar constructions. The book review that I translated with the system was written using the present and the future tenses. As a result, it became evident that the Spanish-English Interpreter is capable of translating such constructions perfectly. As to the past tenses, they are also translated well enough, especially in the case of simpler constructions. At the same time, it is obvious that the system still makes a considerable number of mistakes when it has to deal with more complicated grammar constructions and longer sentences.

One of the more significant mistakes, which could be eliminated in order to improve the overall quality of translation, usually occurs in longer sentences that contain the conjunction *que* (that). In such sentences, the system fails to recognize that the part of the sentence introduced by the word *que* usually refers to the noun from the preceding part of the sentence.

Los asistentes a la marcha, que partió de la Plaza del Sagrado Corazón, desfilaron por la arteria principal de Bilbao.

The spectators to the running, that he departed from the Sacred Heart's Plaza, paraded for Bilbao's principal artery.

This mistake has a tendency to occur every time the system encounters a sentence like this. As the Interpreter fails to recognize the part of the sentence that follows after the conjunction *que* as referring to the preceding noun, it tends to provide a pronoun to go with the verb that is actually used to describe the preceding noun. As a result, a confusion is created as to the general meaning of the sentence:

A la marcha, que estaba respaldada por todos los partidos políticos, excepto la suspendida Batasuna, asistieron unas 2.500 personas, según informaron fuentes de la Policía Municipal de Bilbao.

To the running, that she was backed for all of the political parties, except the suspended Batasuna, a 2,500 people assisted, according to the sources gave information about the Municipal Bilbao's police.

This error produces even more confusion in sentences that contain more than one conjunction *que*:

Pero la crítica del escritor no solo se queda en el mundo de las nuevas tecnologías, sino que también abarca el tipo de información y programas que ofrecen las televisiones, que “provocan un embotellamiento mental”.

But the writer's critic not only remains in the world of new technologies but also wooden shoe, the fellow of information and programs that they offer televisions, that they provoke a mental traffic jam.

Coupled with the insufficiencies of the vocabulary that were discussed earlier, this mistake renders the sentence quite incomprehensible. It is evident that the simple elimination of the personal pronoun that the system tends to add after *que* will help to clarify the meaning of the phrase that, overall, is translated very well and is quite comprehensible.

Taking into consideration that the texts I used to test the Interpreter contain a considerable number of quotations marked in the text by quotation marks, the quality of translation of such texts is not as high as could be hoped for, due to the fact that the system fails to recognize quotation marks as bearing any significance within the text. As a result, such sentences are not recognized as consisting of two separate units with different meaning within one sentence:

Las declaraciones las hizo el premio Príncipe de Asturias de Comunicación y Humanidades de 2000 en Venecia durante un seminario en el que ha expresado la necesidad de “**leer para crearse un propia identidad cultural** que las redes informáticas no pueden aportar”.

*The declarations did them the reward Comunicación's Prince of Asturias that he has expressed the need to **read to create for oneself in and Humanidades of 2000 in Venice during a seminary one of one's own cultural identity** that the information-technology nets can not contribute.*

If the system were capable of translating the part of the sentence between the quotation marks separately from the rest of the sentence, the errors that can be seen in the preceding example could be avoided. This version of the system, though, pays no attention to the quotation marks and translates all the parts of the sentence together, mixing the elements of the sentence that are located in and out of the quotation marks.

To conclude this part of the analysis, it can be said that the system allows for a good comprehension of a text that does not contain many complex sentences. As for the texts that do contain a significant number of such sentences, the end result of a translation often allows for the understanding of the general meaning of such a sentence, while finer shades of meaning are lost.

II. ENGLISH-SPANISH TRANSLATIONS

In order to test the quality of translations from English into Spanish provided by the system, I used a news communication by CNN and a longer editorial, as these texts contain vocabulary on a wide variety of topics and some quite complicated grammar constructions. Overall, it can be said that the English-Spanish Interpreter Professional provides a high-quality translations from English into Spanish. Even sentences that could be considered quite complicated were translated well. In general, English-Spanish Interpreter Professional allows to understand the meaning of the original text very well.

A) *Vocabulary*

Although it is evident that the system's English-Spanish dictionary is quite rich, there are still some instances in which its lack of some important and often-used set expressions influences the quality of the translation. For instance, the system failed to recognize the expression *as well*, translating it not as a set expression, but rather word by word. As a result, the word *well* in this expression is translated as a noun:

We will use these actively in dealing with schools, and we'll use them in cases we're litigating as well.

usaremos estos activamente en tratar con escuelas, y los usaremos en los casos que litigamos tan pozo

Taking into consideration that this expression is encountered very often in all kinds of texts, it seems to be advisable to add it to the general dictionary. The same happens with several other set expressions, which are translated word by word, instead of as one unit:

Home front - la parte delantera de la casa;
 Spelling out - *mejor trabajo de ortografía fuera*;
 A joint call - *una llamada de la juntura*;
 Pointed out - *apuntado apagado*;
 Put in harm's way - *metido en la forma de daño*, etc.

It is always necessary to keep in mind the high content of colloquial and set expressions that abound in the English language and are often encountered in such texts. That is why it is necessary to amplify the general dictionary of the system in order to improve the quality of translations from English into Spanish. At the same time, it might be advisable to add to the dictionary some geographic locations, and maybe even names of most important political leaders of today, in order to avoid the system translating such names, as happened in the case when George W. Bush was mentioned in one of the texts:

Not once — not to the students at his alma mater, Yale, not to any gathering of the College Republicans, nor the young business leaders — has *President George W. Bush* publicly ever urged, asked or even suggested to anyone that any one of them serve his or her nation by enlisting in the military.

No una vez que - para no los estudiantes en su alma máter, Yale, para no cualquier reunión de los College Republicans, ni los líderes comerciales jóvenes - tiene W President George. Crezca espesamente públicamente alguna vez urgido, preguntado o aun sugerido para cualquiera que cualesquier de ellos sirve la de él o su nación enlistándose a en las fuerzas armadas.

There is no doubt that the name of the current president of the United States will appear often in a considerable number of texts, and the quality of translation will improve if such words are added to the general dictionary.

B) *Grammar*

The poor quality of translation as seen in the preceding example is generally not customary in English-Spanish translations provided by the system. This sort of confusion only occurs in longer sentences, which abound in complicated grammar constructions. Many parts of the texts that I used to test the system were translated perfectly, with only some minor mistakes that do not prevent good understanding of the meaning of the text that is being translated. To give but a few examples of such high-quality translations:

i) The sons of the nation's most affluent and most influential families are overwhelmingly missing from the enlisted ranks of the military. The most privileged are woefully underrepresented, particularly in the enlisted ranks.

Los hijos de las familias más adineradas y más influyentes de la nación faltan abrumadoramente de los rangos alistados de las fuerzas armadas. Lo más privilegiado está tristemente poco representado, particularmente en los rangos alistados.

ii) But the department also offers more details on such long contentious matters as moments of silence and prayer in student assemblies.

Pero el departamento también ofrece más detalles en tales materias bastante contenciosas como los momentos de silencio y oración en las asambleas estudiantiles.

It can be seen that these sentences were translated very well, and anybody who reads them will not have any problems understanding the meaning of the text in general and these phrases in particular.

As for the system's weaknesses, it seems worth mentioning that the same kind of complex sentences containing the word *that*, which the Spanish-English Interpreter also had problems translating, are sometimes not translated too well, although it does not happen quite as often as it does in Spanish-English translations. The following sentence could be given as an example of this phenomenon:

I'm very excited about the clarity, and very optimistic that these guidelines will go a long way in solving issues related to students' religious speech.

Estoy muy excitado acerca de la claridad, y muy optimista ese estas líneas directivas llegarán muy lejos en asuntos que soluciona guardaron relación con el discurso religioso de estudiantes.

In this sentence, the word *that* is not translated correctly when it appears in the English text, while it is added by the system to the Spanish version of the text in the place where it is not supposed to appear, confusing the meaning of the whole sentence. At the same time, it is important to remember that there are many instances where such sentences (the ones containing the word *that*) are translated reasonably well:

To avoid controversy, schools may issue disclaimers clarifying that such speech does not represent the school.

Para evitar controversia, las escuelas pueden emitir descargos de responsabilidad aclarando que tal discurso no representa la escuela.

By way of conclusion, it should be said that the WordMagic English-Spanish Interpreter Professional is capable of providing its users with translations that, although varying in quality according to the level of the difficulty of the text, can give a fairly good idea of the general meaning of the text. In the case of Spanish to English translations, finer shades of meaning and the details of the texts that are translated are sometimes lost, while translations from English into Spanish provide the users of the system with reasonably accurate understanding of the text as a whole. The improvements suggested in the present article will allow to further improve the quality of translations provided by the system in question.

NOTES

- 1 The version of the Interpreter that is analyzed here is the one offered by WordMagic as of February 2003. Currently the company is working on a new version of the system containing significant improvements and changes.

OLGA BEZHANOVA
McGill University, Montreal, Canada
E-mail: olga_bezhanova@hotmail.com