

[From: *Ezine Articles*, submitted 18 March 2007]

Machine translation – how to get a good machine translation

By Stefan Dowd

The quality of the translation depends on two aspects:

- A text written in formal language.
- A professional machine translator.

In this first article, we will discuss the aspect related with the source text. For this a definition of language will help us.

Definition: A language is a collection of words, of conventions and semantic rules.

We will now analyze the elements of the definition.

Vocabulary: Collection of valid words in a language.

Grammar: System of rules which govern how to form the words, and how they combine to form sentences.

Spelling: Branch of the normative grammar that points out the rules of correct representation of a language.

Semantics: The study of the meaning of the expressions of a language.

Syntax: The order and the relationship between words and other structural elements in the sentences.

All language that can be developed on the base of this definition is denominated *Formal Language*. Any other expression way is denominated *Informal Language*.

Formal language: The language that is used in schools, in text books, in businesses and in the government's transactions. (academic)

Informal language: The language that is used in daily conversations. (social)

The idioms of the informal language vary according to the region, the time, people's habits, the tradition. The informality implies the nonexistence of rules.

On the other hand, the formal language is based on rules. Having rules, it is not difficult to implement an automatic system that recognizes them. The objective of a translator is to translate a text written in formal language.

To get a good machine translation, the text should be:

- Grammatically correct : without spelling errors, adjusted to the rules.
- Semantically correct : meaning in the text.
- Syntactically correct : words and punctuation marks used correctly.

To proof it, this article was written in Spanish and it was translated with a professional machine translator.

In the next article we will discuss the aspect related with the translator, their two fundamental elements and how to recognize a professional translator.

A machine translator produces the translation, using rules that transfer the grammatical structure from the origin language to the destination language. A good translation depends on the quality of the analysis of the translation and of the size of the used dictionary.

The most important factors that define a professional translator are:

- Wide vocabulary : possess the biggest quantity possible of words of a language.
- Powerful analysis : possess the biggest quantity possible of semantic and grammatical rules.

These are the two fundamental elements of a machine translator. Better they are implemented, better the translation.

Vocabulary:

If a machine translator doesn't recognize the words, the grammatical structure of the sentence won't be understood. Let us see an example of Spanish to English:

Original sentence in Spanish:

Aunque esta definición es muy simplista, es conveniente al propósito que se persigue.

Correctly translated sentence:

Although this definition is very simplistic, it is convenient to the purpose that is pursued.

Translated sentence using the on-line most popular translator:

Although this definition is very simplista, is advisable to the intention that is persecuted.

Here we find two vocabulary problems:

- it doesn't recognize the word: **simplista**.
- the meanings of the words: **conveniente** and **persigue**, are erroneous in the context.

Analysis:

If a machine translator doesn't understand the grammatical structure of the sentence, the words won't be located correctly, altering the meaning of the original sentence.

Original sentence in Spanish:

Sintaxis: El orden y la relación entre palabras y otros elementos estructurales en las oraciones.

Correctly translated sentence:

Syntax: The order and the relationship between words and other structural elements in the sentences.

Translated sentence using the on-line most popular translator:

Syntax: The order and the relation between structural words and other elements in the orations.

Locating a word incorrectly, the structure of the sentence is modified.

In the two analyzed cases, the sentence lacks sense.

How to recognize a professional translator.

The test of the inverse translation is the best way to know if a translator is good, that is to say to check that the translation of the translation maintains the main idea. The vocabulary errors and analysis errors, will appear clearly in the inverse translation. Let us see a good example:

Original sentence in Spanish:

Un traductor automático produce la traducción, usando reglas que transfieren la estructura gramatical del lenguaje de origen al lenguaje de destino.

Correctly translated sentence:

A machine translator produces the translation, using rules that transfer the grammatical structure from the origin language to the destination language.

Translation of the translation:

Un traductor de máquina produce la traducción, usando reglas que transfieren la estructura gramatical de la lengua de origen a la lengua de destino.

Although the produced sentence is not exactly similar to the original one, it maintains the concept.

To get a good machine translation:

- Avoid the informal language.
- Avoid long and complex sentences.
- Use the active voice.
- Don't use capital letter unnecessarily.
- Make sure that the original document is grammatically correct.
- Use a spelling checker.

The most common errors:

- Incorrect use of the accents. (Spanish to English).
- Incorrect use of the commas.
- Incorrect use of the punctuation that separates sentences
- Words incorrectly spelled.

Conclusion:

Following the suggestions and avoiding the errors, the best satisfaction that gives a good translation is obtained: the test of the inverse translation.

Don't translate the complete document, translate paragraphs carrying out the inverse validation. If is necessary, change some word or invert the order of the sentence. The final result will surprise you.

To proof it, this article was written in Spanish and it was translated with a professional machine translator.