Glossaries

A linguistic feature that is sometimes neglected, perhaps because it is not a feat in terms of computer sciences, is the capacity to build glossaries.

Glossaries are VERY appreciated in the world of translators, whenever you have to deal with a subject that requires some expertise. Please make the test: find a commercial website in your own language that proposes very precise devices to catch the perch. Then try to find the corresponding technical word in another language that you know... and count your time...

If you had a good glossary at hand, your speed would be multiplied by ten. And not only you would translate fast, but you would rest assured to provide a reliable translation, not ridicule yourself in a subject that you don't know well...

By "glossaries" you may understand something like an average two columns bilingual table: "red = rouge"; "red= rot".... That is more like a terminology, or a short thematic bilingual dictionnary. In such a file, an entry exists because it is difficult to find the corresponding word in another language, but the knowledge of the object is close to universal (water = eau = wasser...). You only have to set the convention, so that one word will be used and not other correct words, and the terminology will be consistent throughout the translation unit. For instance this will be "mutton" and not "sheep in English", although both would be logical.

In order to build a superior quality glossary, not only you have to know technical words that are not in the common language, but you need to provide some expertise with the facts those words describe, which leads to much better accuracy in the use of those words. A glossary is a set of words that you may not be able to use properly even in your own language, and therefore a glossary can be monolingual, and useful as such.

Therefore a glossary cannot be written by linguists, language professors, etc. It has to be written by people of the trade. It is interesting that a cms be built in order to help the building of glossaries by people who don't work in language, but are specialists of their own trade. Their expertise is highly useful in the translation feedback phase othe translation process, and in the capitalization of knowledge in the wiki way.

The existence of websites on very precise subjects is a unique opportunity to create glossaries. The building of a glossary should begin by selecting a word in an already existing text and click on a button "Make a glossary out of this".

This should lead to a special page designed for Glossary words. In a way you could say in a way it creates just another wiki page: you click on a word in order to know what it means. But experts should be able to fill comments or definitions fields.

Such a special template should contain a linguistic form, with fields such as gender, plural, etc. But this linguistic form should be collapsed, because the method is to have the glossary page set up by people of the trade first, and not linguists. Linguists will know how to expand the linguistic form. Priority must be given to people with topical expertise.

In the translation interface, with a column for source text, words that are glossary entries should be visually distinct, with a glossary pop-up if clicked. There could also be a list of glossary entries available somewhere.

Possible automatic help in the creation of a glossary page:

- automatic link to the glossary pages for this proposed word into another target language
- automatic quotation of all sentences in the wiki where this source word occurrs, with links to concerned pages.