It's raining cats and dogs…

Wiki in translation of idioms and proverbs: the importance of collaborative work.

Dora Montagna

Università degli Studi di Torino – Università Sapienza di Roma

Presentation.

The reason why I would like to propose my experience to this workshop is that I'd like a confrontation about some possible uses of Wiki. I don't have any computational education, nor am I an expert in programming: I'm a linguistic and a translator. I discovered the potential of Wiki when, together with my colleagues, we had to complete a multilingual glossary for the web. The problem was that each of us lived in a different city in Europe. When we discovered the Wiki, we experienced the potential of collaborative work (so useful in translation!) and we realized that the Wiki was going to be the solution to make our work complete, usable and sharable. That's how our first "twenty-hand" glossary was born!

Now, I am starting to use the same support to cope with a very specific aspect of the language and translation: idioms and proverbs. Actually, idioms represent a window on a community's linguistic culture, they are not necessarily the same for all speakers and they often change from a generation to another. Proverbs, from which idioms often draw on, are an expression of popular wisdom, which is why they exist in a big variety of forms and nuances depending on the geographical and social context, the dialect, the historical moment. It is evident that such expressions represent a great difficulty for a translator. But the cultural debate that emerges from their study involves translators, linguists and the whole community of speakers! Such oral knowledge, for its changeable nature, is hardly held and registered in collections and conventional dictionaries. For this reason the Wiki as a collaborative and open tool, seemed to me the ideal instrument to open the confrontation.

The research.

To start my research, I limited the corpus to a selection of idioms and proverbs containing the term "*mosca"/ "mucha"* (fly) in Italian, Spanish and Polish. I consulted recent idioms and proverbs collections, both monolingual and multilingual, and I registered all entries I found in ancient editions as well. Afterward, I compared expressions in different idioms, establishing semantic and/or formal equivalence relations. However, that meant to restrict the research to already registered and published expressions, and to exclude many geographical, social, and chronological nuances. Moreover, idioms and proverbs are speakers' domain, and their interpretation is up to a speaker. As I wanted to establish relations and create spaces for comments to each entry, it was essential to me to deeply understand the sense, or the senses, of the expressions. Thus, it seemed important to me to open a confrontation between the speakers about the meaning of certain expressions, and the more people got involved in the confrontation, the more new expressions and idioms came out, handed down

from a generation to another and never had been written anywhere. So, I decided to put my corpus on-line, and to open the debate.

The application.

The application is a synthesis between a multilingual on line glossary of idioms and proverbs and a translator's forum. It presents a multiple index of idioms and proverbs (one index for each language). Search is possible either by index or by key word. Selecting each entry, a page with all variants and equivalents in other languages will be shown. From this page, it is possible to proceed to a commentary page, with useful translating, linguistic and cultural observations. Users can contribute to each page in order to add new entries or their interpretation of the idiom. A discussion page can also be used to discuss crucial issues.

Why Wiki.

One of the main problems of idioms and proverbs, and in general of multi-word expressions collection is their classification, and, in consequence, their finding, especially in multilingual corpora. An index of such corpus can be done by key words (e. g. the *fly*), by motivation (e. g. the *poverty*), or in alphabetical order, for instance. Wiki allows a very extensive textual corpus as a data base, and allows its navigation in different levels, by double or multiple index, as well as by searching by key words. Despite, the issue of classification remains an open question and subject to a continuous debate. In fact, classification is essential also in a Wiki support, as it stays at the base of the codification of the corpus and of every following page: to each entry of a corpus, as well as to each Wiki page, must necessarily correspond a logic code.

Another problem is represented by the idioms' and proverbs' innumerable variants, and their interpretation. Moreover, two given expressions presenting formal equivalence, may not present a semantic correspondence in the other language. Wiki allows creating additional pages with notes and in-depth examination of the expression, to which all users can contribute if they notice, for instance, "false cognates", different interpretations or variations, or if they don't agree with the given meaning. In that way, each comment, annotation, cultural and linguistic observation useful to the translator, would be easily available and would not interfere with the classification of idioms and proverbs in the corpus.

Moreover, collaborative approach is truly essential in this field. Idioms and proverbs are expressions from the oral linguistic and cultural heritage, often not catalogued in libraries, changing from one region to another and domain of different generations. For this reason, it is important to maintain an open access to modification of idioms and proverbs corpora so each user could contribute to their interpretation and add new expressions and new variants. A collaborative and dynamic environment, such as the Wiki, reflects well, in my opinion, the mutable characteristic of the language concerning idiomatic and proverbial expressions.

As a translator, I find this application a useful tool for its open access character. Moreover, proverbs databases are not easily accessible on line. I'm not an expert of computational issues, thus I'm sure this application could be improved a lot through discussion with other linguistics and Wiki designers, and that it could be enriched by their experience!