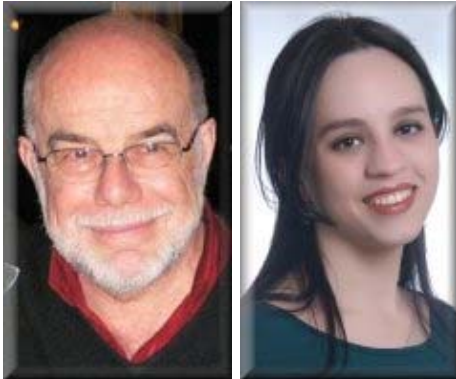Danilo Nogueira and **Kelli Semolini** are about the unlikeliest of pairs. Born 40 years apart and living a good 200 kilometers away one from the other, they have met less than half a dozen times. They have practically nothing in common but a sweet tooth—and the love for languages and translation.

Yet, they work well together and are full of plans, three of which they are sharing with Translation Journal readers first hand. First in the pipeline is an online English <> Brazilian Portuguese business dictionary that should reach the testing stage in early January, 2008; next comes a translation course in distance-learning format. There is a book, also, obviously on translation, but they won't mention the details under torture.

If you read Portuguese, you will probably like their blog, which is currently hosted at http://tradutor-profissional.blogspot.com.

Danilo e Kelli can be contacted at danilo.tradutor@gmail.com and kelsem@gmail.com, respectively.

Front Page

Select one of the previous 52 issues.
Select an issue:

The Profession

Will We Be Here Tomorrow?

by Danilo Nogueira and Kelli Semolini

*M*achine translation is here to stay, and there is nothing we can do to make it go away. Are we, professional translators, here to stay too? Good question!

Any product or service, including translations, can be defined in terms of three parameters: cost, delivery time, and quality. All machine translation systems offer unbeatable speed and very low prices, principally if you are going to deal with high volumes. Web-based systems are free, at least for the time being.

All this means that we humans must compete on quality alone. Unfortunately (for us, of course), MT quality is visibly improving. The results are often better than many professionals can provide—although not yet as good as the results delivered by a good professional. Will they ever be?

However, the point is—or may be—that quality is an elusive and complex parameter. What is quality, after all? What is a good translation? How good a translation must be before it is considered "acceptable?"

> Today's translator will probably be tomorrow's pre-editor or post-editor, but those who are competent and can adapt will be here tomorrow.

Some translations (human and otherwise) will be unanimously considered unacceptable. Above that level, however, quality may be a matter for disagreement. Quality means different things to different people and that is one of the reasons why there are so many translation theories. In fact, you can only evaluate the quality of a translation against the backdrop of a specific theory. That is why Dr. Eugene Nida's work is admired by so many and execrated by so many more, for example. Let's ask a simple question: what theory does Google Translate follow? Dynamic equivalence? Formal equivalence? Or is it more modern and follows Dr. Vermeer's skopos theory? Can it "deconstruct" a text in Derrida's manner? Of course not!

Focus, Please!

This is a point all MT providers evade: the absolute lack of theoretical focus in their systems. It should be possible to build a rule-based system according to some current theories—although in the case of skopos we would need different systems for different aims, a requirement that would probably make the job impossible in practice. However, Google and Bing Translator, two emblematic and highly praised systems, are based on an uncritical alignment of whatever could be found on the web and thus are an irremediable hodgepodge of styles and theories.

That is one of the reasons why so many translations provided by those systems seem to require some polishing —because although not unacceptably bad, they fail the test of compliance with whatever theory whoever asked for the translation had in mind.

Don't even try to tell us you are a practical professional who simply translates the way you think is right and do not pay much attention to this theory thing. All of us translate according to some theory or the other. Most "practical translators" have developed their own theories and, if have failed to formally develop them, it was for lack of appropriate training or plain prejudice against "theories." It is a theory, even if not fully developed or expounded, that allows us to choose between two acceptable alternatives. And different translators make different choices because they support different theories.

One Size Fits All...

The one-size-fits-all approach extends to language varieties. Most MT packages cannot tell Brazilian from European Portuguese, for instance, and will even merrily mix both varieties in the same sentence, possibly in a well-intentioned attempt to reunify the two flavors of Portuguese, an endeavor wherein the so-called orthographical agreement has utterly failed. The same should apply to German, French and Spanish—not to mention English—and perhaps to several other languages in different measures.

If you deal with any of those languages you know how dangerous it is to use the variety that is wrong for your public. It may look funny or annoying and in some cases it will make no sense at all. In the worst cases, it will be misleading. But Google Translate and most other systems ignore the problem, like the client who wants "a translation into correct and elegant Portuguese that will be accepted on both sides of the Atlantic." It is a fact that such clients will readily find someone willing to translate into "mid-Atlantic" Portuguese, producing texts that will look equally funny and unacceptable on both sides of the pond, but that is a different problem. Remind us to write on the ethics of translating some of these days, please.

... and Sometimes None!

But we know some of the results provided by MT systems

are simply wrong, under any theory and for all linguistic flavors. Even MT suppliers will admit that and that MT output for publication must be "post-edited." Depending on output quality, post-editing may be a difficult and thus expensive task. Editing time and money can be saved if the text is written in some kind of controlled language —meaning short sentences, simple constructions, a limited vocabulary, and no ambiguity. This may range from very easy to utterly impossible.

Sometimes good results are achieved by having someone straightjacket the source text into controlled language, something the industry refers to as "pre-editing". Pre-editing can cost as much as human translation and is economically feasible when the same text is to be translated into several languages, because the same pre-edited source text will then be used for all target languages, thus diluting the cost.

Both pre- and post-editing must be in the hands of humans, of course.

For Screening Purposes Only

Translations for screening purposes, also known as "gisting," that is, to determine whether the text should be translated, however, do not need either pre- or post-editing, let alone theoretical consistency.

In the old times they were done "diagonally": a translator would have a quick look at a document, summarize it aloud in a couple phrases and the client would decide whether to have the document translated or not. The process offered at least three advantages: it saved time, translation costs and also a very scarce resource: translator capacity. Because good translators are very, very scarce and squandering the time of a good one in doing a text that will prove useless is an unjustifiable waste.

Diagonal translation has been on the wane for a long time and in many places has already been entirely replaced by raw MT output, since even texts produced by very primitive MT systems are good enough for screening purposes.

A Matter of Discretion

It may be reassuring to learn that a few colleagues have been asked to sign an undertaking not to use MT translation at all. There may be several reasons for this.

Some clients are afraid translators will supply them with raw, unedited MT output—and charge for it. We can understand that. Once, we were sent one of those jobs to edit and the PM, who did not know Portuguese, was very surprised to learn he had been fooled into paying for that kind of job. Of course, there is no problem in taking advantage of MT if it is properly edited into a quality job. Or is there?

If an offline system like Systran is used, we can see no peril, but online systems such as Google Translator, for instance, may be a different matter. Google has repeatedly

stated they do not store or use whatever they are asked to translate, but those statements should be taken with a grain of salt. Many colleagues who have used Google Translator together with CAT software have reported some strange observations, which may indicate that accepted translations are in fact added to Google's database. Since this process is not widely known, allow us to explain how it works. If you already know, just skip the next paragraph.

CAT software searches the translation memory for materials that can be "recycled" into a new translation and offers the translator a suggestion that will be edited into a correct translation for the particular segment at hand. Sometimes nothing is found and the software leaves the target field blank. If the MT option is used, however, the software will ask Google for a translation, enter it in the target field and give the translator a chance to edit and accept it.

According to software developers, it is a one-way system: The program will ask Google for a suggestion but does not inform Google what the translator did with it. That, plus Google's statements should calm everybody down , including us and our clients. However, many colleagues have reported strange facts. For instance, you are translating a contract and the suggestions offered by Google for the first few segments contain a certain type of error, which you correct before "accepting" the translation. All of a sudden, Google seems to notice it was wrong and stops making the same mistake again. How come it could do it without "reading" your corrections?

So, what does really happen? Some colleagues will say that it is just our impression or a coincidence—but others claim that although the CAT software does not send finished translations back to Google, Google places a zillion cookies in our computer during translation and those cookies capture the translation as soon as we accept it. We do not pretend to know what happens, but we find those things a bit odd.

On the one hand, there is the point of "take a little, give a little": you are using Google Translate for nothing, so why not contribute a little too? Yes, agreed. On the other hand, so much of our stuff is confidential, that we cannot risk.

It should be interesting to notice that web-based alignment services as well as those websites that convert Adobe pdf files into MSWord format are equally dangerous. Who knows what they will do with the stuff they process?

Will we Have a Job Tomorrow?

Everything we wrote above is of no importance to us, human translators, except for the bits about confidentiality. The question, the real question is *will we be chased out of the market by MT?* The decision rests with our clients, not with us and our arguments.

Will our clients drop us for Google or Bing? Some indeed will, others will not, and the decision is not likely to be based on the points we made above—among other things

because not all our clients read the *Translation Journal*.

The type of client who wants the lowest possible price and does not care a hoot about quality can be counted among the first casualties. Other clients simply cannot tell good from bad. This will mostly be the case where the client wants a translation into a foreign language. A German client who wants a translation into German is likely to read the translated text and be able to tell it lacks quality and requires post-editing. But the same client may be quite happy with a bad quality translation into Portuguese, for the sole reason he cannot tell Portuguese from Chinook.

Clients who care about quality will have MT jobs edited by a human translator, at least just to make sure it is good enough, because it is undeniable that MT has its shortcomings, some of which are recognized in the description of controlled language: MT does not deal well with long sentences and complex structures and, of course, cannot translate words or senses not included in its vocabulary. Neologisms will baffle MT, as will puns, metaphors, rhyme, alliteration... all those things that enrich text are wasted on MT. That should spare literary translation, of course, but do not make the mistake of believing that only *belles lettres* make use of those resources. Even financial translation, which is the daily fare of the authors, is full of word games.

In addition, MT takes texts at their face value and translates what the author wrote. A badly written phrase will look even worse after MT, and MT cannot correct errors. It takes a human to do that. However, it is very likely that editing MT output will play an increasingly important role in our professional lives, for it also takes a human to edit MT.

In short, surviving is just a case of adapting to change. That should be no news to anybody: we have been adapting to change since the times when we lived in caves. Many people said CAT would be the end of us, that agencies would accumulate huge memories and do without translators, didn't they? Nothing of that kind: the quantity of work going round has increased incessantly, and it is likely to continually increase in the foreseeable future.

Today's translator will probably be tomorrow's pre-editor or post-editor, but those who are competent and can adapt will be here tomorrow, and remember the "good" old days when you actually had to type every single letter of a translation.