# TRANSLATION IS ABOUT PEOPLE

*A look at integrating automatic translation into the localization process*

**SERGE GLADKOFF**

Processor power continues to skyrocket in obedient observance of Moore's law. The Internet, not so long ago a notoriously slow WorldWide Wait, is rapidly turning into worldwide cable TV. Automatic translation, translation databases and controlled authoring language tools increase in power.

It is often stated that "using a sophisticated translation system to produce a first draft, with consistent terminology throughout all documents, can increase a translator's productivity by several factors." I believe that this is a purely theoretical claim.

## QUALITY AND PRODUCTIVITY

Quality standards across the localization industry are quite demanding. It is not acceptable to have a single translation error in a 100-page document (about 30,000 words or 3,000 sentences). This is a less-than-0.03% allowance for translation error, or accuracy above 99.97%. This is a very high level of quality indeed; this error level is not feasible even for a single professional translator. These quality standards require a sophisticated localization document workflow with every translation mandatorily passing through at least three people: translator, editor and proofreader.

Our editors and proofreaders know our translators and rank them as "perfect," "very good" or "good." No translators worse than "good" are in the business, and nothing below "good" is accepted from translators. Here "good" means that a translator has 1) understood every sentence of the source document and 2) translated it correctly. The key word here is *understood*. It is the understanding of the source text that holds the key to the entire process.

The very first part of the process is to achieve a full rapport with the author. Only after that can you start putting it into the target language, watching for appropriate context, following terminology, observing client style, then passing through the stage of literary editing, proofreading and so on. Human comprehension is the keystone of the translation process; source and target language grammars are only technical tools to facilitate comprehension.

These are only the language-specific stages. Technical, project-specific or software-specific stages may appear at any time.

This critical path of source document translation has certain throughput limitations per capita. Productivity at every stage has been thoroughly benchmarked, and the averages are well known. Average productivity in the translation stage is about 2,500 words per person per day; the speed of the QA stage is only twice as high, about 5,000 words per person per day. These are average numbers valid for a final text of average difficulty. If the source is very simple and familiar, translation productivity

may reach as much as 10,000 words per day. If much research is to be done in the new field or in the software, productivity may drop to any number close to zero.

## INTEGRATING AUTOMATIC TRANSLATION

I am not challenging anybody to undertake a full replacement of this workflow. Let's start with a modest goal and try to build automatic translation into this process. Where? Is it possible to replace the translation stage with automatic translation and pass its output to the QA stage?

One could try to modify the traditional workflow as follows: automatic translation, then post-editing and then traditional QA (not being willing to sacrifice quality). Unfortunately, *quality* means observing all the requirements, and therefore all the original stages remain; they only move to the next stage.

The point is that with software translation, the task of making the process yield something meaningful actually has become more complex. At the post-editing stage you have to read two sources instead of one: the one in the source file and the other in the translation output. The whole idea of post-editing is approval and correction output by an expert. To approve someone else's work, you have to read and understand the source text again and, in addition to that, compare and check if the translation is good.

With or without automatic translation, we just derived that post-editing is a pure definition of rework. Not only is rework something that professionals hate, it is unacceptable *by definition*. The ultimate goal of fine-tuning the traditional production translation process is to avoid rework as much as possible because it is time consuming and excessive. And comprehension speed is limited to only as much as 3,000 words per person per day.

When I have a file for editing from an experienced translator, I know that it is extremely unlikely that there's a translation error. I have no need to compare his/her writing to the source text, and I carry out only a finite number of compliance and QA checks. When I have a file that is automatically translated, I know in advance that I have to check every sentence. This does not promise any productivity gain whatsoever. All this promises is pain, low quality,

complaints from my people and an overall drop in productivity.

Translation is all about understanding and comprehension; there is no way to get rid of this crucial stage. You can wrap this into different packaging, but translation is all about people.

## ONE UNFORTUNATE CASE

Two years ago we had a client who asked us to review an automatic translation process and possibly to take part in the project at the post-editing stage. The job was to translate a 100,000-line description of some communication operating system and manuals. The client believed that the volume of the job and especially the price level of localization industry professionals were prohibitive for taking solely a

### Translation Processes

**Without Automation**

Translation Stage
  Comprehension
  Writing

→

Quality Assurance Stage
  Editing
  Proofreading

**With Automation**

Translation Stage
  Automated translation
  Comprehension
  Writing

→

Quality Assurance Stage
  Editing
  Proofreading

manual translation approach. A decision had been made to "train" an automated software translation product, run large volumes of source text through it and do human post-editing thereafter.

We were asked to evaluate the feasibility of such post-editing and come up with an acceptable productivity and pricing offer. To do the evaluation, we were given the source English text, text after translation and the same text after post-editing, which the client considered acceptable.

At first glance, the translated text was close to acceptable. The software "knew" some terms of the field, and the sample post-editing was done on a very basic level, obviously to minimize the time and manual effort in the final stage. To evaluate the scope of the work, we did a very simple thing: we carried out a TRADOS analysis and compared post-edited "final," "acceptable" provided text against the pre-translated text. There were no matches. Every sentence of the final text had been edited and corrected.

Even with this obviously deliberately sketchy post-editing provided by the client, there were changes in every sentence.

What does this mean? Our productivity benchmarking shows that with a match less than 80%, the time it takes to correct the sentence is equal to the time needed to do a brand-new high-quality translation. Period. There's no final engine-generated text without post-editing, and to do post-editing, a person must read and understand the entire translation unit and verify that the outcome adequately and correctly conveys the original thought. The key word here, again, is *understand*. I claim that comprehension is the critical task that takes most of the translation time. There's no shortcut in sight.
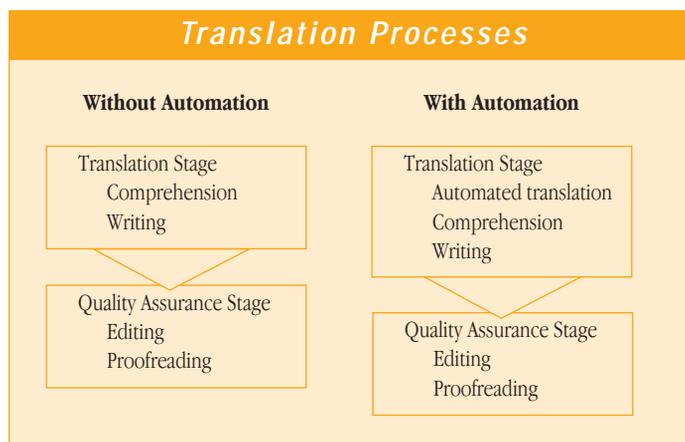
It is well known that if translation quality is below a certain level, it is easier and faster to retranslate from scratch than to improve the incorrect translation. And this threshold happens to be very high, very far from the reach of current products. As of this writing, I believe that this threshold lies somewhere close to what we call Artificial Intelligence.

## ACQUISITION AND OWNERSHIP COSTS

Translation is all about process as well as people. It takes a great deal of effort to get things up and running, and every sensible manager is aware of the eternal principle "when it runs fine, don't touch it." A change takes courage, and the more radical the change is going to be, the more courage it takes, especially if the outcome is uncertain. Luigi Riboldi, Logoscript general manager, points out that another major decelerating factor of future automatic translation deployment in the localization industry would be high acquisition and ownership costs. Ownership cost is not so much the cost of licensing the technology itself (the software product), but rather the cost of its deployment in the company, changing the process to incorporate automatic translation, training personnel, implementing changes in workflow and so on.

The more complex a technology is, the higher are the costs of ownership. With technology as complex as this, the cost of ownership would be enormous. One would have to keep dedicated people to maintain all the glossaries, always continue to "train" the system — the fight for quality through glossary extension is a battle without an end in sight, if not a losing battle.

## ON THE CLIENT SIDE, TOO

Considerations about high deployment (acquisition) costs of translation technology are fully applicable to the client side as well, no less than to the vendor. Actually, they are worse because while the vendor may see translation as a core technology, for the client it may be of the same perceived importance as drinking-water delivery. All attempts from outside solution providers or consultancies to embed translation software into the client's process will encounter substantial difficulty simply because, as Detroit Translation Bureau president and CEO Rick Woyde has said, using translation software "requires the client's process to be reengineered, and I don't know anybody who wants to be reengineered."

## A PROCESS WITHOUT THOUGHT

Is the picture really so hopeless? What about the controlled language applications? I do not deny at all what such translation can do in the field of information gathering, "gisting" translation or surfing Web sites in unfamiliar languages.

But now consider this: Automatic translation is a process without human thought, or at least a process that applies only very limited thought. If something can be successfully converted to another language with this tool, then the outcome most probably is of very limited use by itself. The idea is very simple. If there is no extensive use of all the facilities of the language, then it is probably largely only technical data, and we are speaking not so much about translation as we understand it but more about generation or representation of information. It falls out of the human dimension and, therefore, is not much of a translation *per se* any longer.

On top of that, if there's a large volume of such material to be "translated," the very first cost-cutting step I would take, if I were in the customer's place, would be to reduce the volume of that material. Chances are that there is not much need for such huge volumes of this kind of information in the first place. If this is only dry technical data, then there must be a cheaper and easier way of making it available and useful than making an ocean of words of it to be translated.

## TRUST, QUALITY AND BRANDS

Translation is, I believe, also about reputation. This is a service business. Everybody knows that there are lawyers and lawyers, physicians and physicians. Similarly, there are translators and translators. If you outsource something to a company with a reputation, you are guaranteed to receive a certain level of quality simply because there is a certain corporate standard that ensures expertise, quality level, communication, reliability and so on. You don't care very much what is inside, automatic translation or not.

It is the business of the localization company to decide what works best. And you entrust them your most beloved child — your product. The statement "we cannot trust a machine to do the work" is actually all about brand building, and the key word here is *trust*. Contracting with someone for outsourced work, you enter a business relationship, a relationship between people. The localization service business is, again, all about people.

## THE WAY IT MAY WORK, AFTER ALL

In 1782 a well-known French scientist, mathematician and astronomer wrote that "it is ultimately and finally proven that people are unable to ascend or even maintain their elevation in the air with the aid of apparatus heavier than the air." In the next jubilant year, 1783, the brothers Joseph-Michel and Jacques-Etienne Montgolfier launched their balloon into the open sky. I might be as unlucky as this French scientist to underestimate the potential of translation technology and the progress it made for the past fifty years, but again I am not trying to prove that it is not possible or useless or fake or unnecessary.

Jaap van der Meer of ALPNET, Inc., says, "All in all, I don't think automatic translation is a threat to our business, but rather a business opportunity we must exploit. I do not rule out an option that such technology will be of service for indirect uses, for example to help people to achieve better productivity or as an internal tool at a translation service enterprise."

It is the professional dimension, indeed, that is important. Instead of contrapositioning technology and professional translators, I would recommend looking at both from a human viewpoint. Translators will use computers, the Internet and software tools to increase their productivity in many ingenious ways. I believe that the tools, word processors, CD-ROM glossaries, Internet, translation databases and automatic translation programs are to be used as productivity tools to assist professional translators rather than replace them. Simply because translation is all about people — this is just the way it works. 🌐

..........................

*Serge Gladkoff is president of Logrus International Corporation. He can be reached at sgladkoff @logrus.ru*