

# MT: Who needs it?

Robert François vs. Muriel Vasconcellos

## on machine translation

I recently received the following letter which says a lot about the dreams skillfully nurtured by unscrupulous manufacturers and the supporters of so-called translation machines.

*Dear Sir: It has been brought to my attention that your company uses freelance technical translators at home. I would appreciate very much any information you can supply me on how I might become a translator for your company. I do not have previous experience but I have a Epson Equity II+ computer with software that is capable of translating from English to Spanish to English and could possibly acquire software that could translate other languages as well. Enclosed please find my resume with references. Thank you for your prompt response. Respectfully yours,*

Her résumé indicated that she was married, had a 3-month old baby and that she was born in 1957. She majored in Computer Information Systems (a catch word for almost everything nowadays). She was responsible for the complete computer operations in accounts payable and receivable and payroll for a supply company, and responsible for the computer files for an affiliate of TRW Credit Systems. She had also sold residential properties and had set up and maintained all bookkeeping and probation records on computer for an adult probation office in Texas. My overall impression of that person was that she was educated, not stupid but simply naive with good intentions. However, she does not make any mention in her letter or her résumé of a knowledge of either French or Spanish.

When you have a 3 month old and you don't know a foreign language, it seems absolutely logical to buy a MT computer and to try to make a buck while taking care of the little one.

We should excuse her candor, because she is simply one more victim of farfetched claims of technological breakthroughs and dramatic improvements made everyday by unscrupulous manufacturers of MT, and a cohort of supporters mainly paid by taxpayers through large international and governmental organizations and universities. In fact, up to now, hard realities snag the quest for artificial intelligence and for very efficient MT in particular. Newspaper articles, technical reviews and professional publications, including the *ATA Chronicle* are unwillingly involved in this vast conspiracy of publishing unproven results and

*Continued on page 25 . . .*

Translators will never be dead in the water. Machine translation, though it may change the way some of us do our work, is never going to replace the human contribution. On the contrary, as a screening mechanism, it is going to greatly increase the demand for translations in general. The fact that MT is available has already spurred information-mongers to look at foreign sources that were not being translated before. At the same time, MT can reduce the drudgery, and it can retrieve technical terminology in place in an era of increasing specialization. The important thing is for translators to keep an open mind, explore the possibilities, and come to an understanding about its uses and limitations. If we all do this, we will undoubtedly find many areas of common agreement.

For instance, whether or not we are in favor of MT, the thought of a monolingual housewife hoping to earn big bucks with "machine translation" software on her PC at home is enough to raise the hackles of us all. Though this is an extreme example, alas it is indicative of a trend. The megaphenomenon of personal computers running sophisticated applications which a short time ago were undreamed-of has led a jaded public to expect all manner of miracles as a matter of course. Hardly a day goes by that the Committee on Translation and Computers doesn't receive an inquiry about MT on PCs, some of them quite naive. We talk to these people, or write to them, and we send them background material in the hope of educating them about the realities of machine translation.

Indeed, the housewife is not to be blamed for her innocent idea of cashing in on yet another kind of software. But I submit that she is not so much a victim of farfetched claims by unscrupulous manufacturers as she is an echo of our collective American consciousness. As a people, we are entrepreneurial, ingenuous, and given to wishful thinking. And the average American knows little if anything about foreign languages, much less about translation. It is not surprising, therefore, that MT has caught the public imagination. People want this miracle to happen. It is they, not the manufacturers, who have created the expectations.

While admittedly MT sales personnel have been known to make exaggerated claims, the vendors of the earlier mainframe-based systems got into trouble with this approach and abandoned it some time ago. What we are seeing today is a new generation of vendors with products

*Continued on page 28 . . .*

... continued from *Joly*, page 1

delegation to the World Congress of FIT in Vienna and Maastricht in 1984. A member of the FIT Council since 1987, he was chairman of the FIT Committee for the Press and Public Relations. He is also a founding member of the FIT Regional Center for North America and was recently appointed chairman of the Executive Committee of the Center. □

... continued from *FIT*, page 4

continue our plans for Montreal 1992 when we meet in New Orleans.

But all work and no play makes Jack a dull translator. Our Yugoslav colleagues invited the delegates to the Statutory Congress to a garden party at their headquarters. We attended a reception at city hall. The Ambassador of Bulgaria invited us to a reception at the Bulgarian Embassy. A tour of Belgrade was included in the registration fee for the Open Congress. We visited the Belgrade Cathedral, Tito's former residence and tomb, and an ancient fortress overlooking the not-so-Blue Danube. The final banquet of the Open Congress ended with Yugoslav songs and dancers.

One of the more enjoyable and worthwhile aspects of a FIT Congress is the opportunity to browse through the exhibits and talk with and learn from colleagues from other countries. We were astonished to learn, for example, that the failure rate for the official Norwegian accreditation exam is between 90 and 94%. The main reason: a candidate must take the exam both into and out of Norwegian!

In short, the opportunity to meet with colleagues from all around the globe is an invaluable one. For that reason, we are glad that there is an International Federation of Translators. We are, however, particularly enthused because the new Council promises to be very dynamic and responsive to the member societies. We will make efforts to provide regular updates as the committees and Council go to work.

The next FIT Congress will be held in Brighton, England, in the summer of 1993. □

... continued from *François on MT*, page 7

results and ridiculous cost-cutting claims. For the most part, these claims completely ignore the realities of the market place by conveniently omitting such simple criteria as the heavy cost of learning curves in MT, the amortization of expensive equipment and the huge amount of work involved in setting up machines, which perform poorly in very narrow and closely-structured fields to which should be added time-consuming pre- and post-editing work.

Even Muriel Vasconcellos, the ATA guru of MT states candidly in her "View of Japan and a Japanese View of Translation Technology" (*ATA Chronicle* No. 6 June 1989) "With regard to the current situation in Japan, MT is generally looked upon as 'the basic technology for future high level information processing.'"

In addition to the government's official Mu2 project, MT systems are being developed by Oki, Canon, Sanyo, CSK, Sharp, Toshiba, IBM Japan, Systran, NEC, Hitachi, etc. But the report does not specify which of these are in practical use. To date, 4,000 MT systems have been sold although "... many are said to have been returned to the seller and some are not used and are idle." In clear language, it means that in spite of large subsidies and a generous host, the Japan Electronic Industry Development Association and its members have not succeeded in finding enough suckers willing to keep their white elephants in service. Because, despite the acceptance of expensive expert systems as advisers on specialized tasks, no system makes important decisions autonomously or possesses even a modicum of common sense. The technology is so feeble that this question is frequently asked: Can expert MT systems be beefed up to duplicate human problem-solving abilities, or will they remain simply business tools for very narrow translation applications?

Another disturbing fact is that most manufacturers of MT products have the bad habit of closing their doors, merging, or simply stop servicing their equipment; in these cases users are left holding the bag.

A close encounter of the first kind with an allegedly large direct user of MT equipment can reveal a lot of the pitfalls of these miraculous systems. During the EXPOLANGUES show held last year in Paris I contacted some translators working for the European Community (EC) which is presently the world's largest user of translated material covering thousands of fields in several languages. I was told that they are using one of the existing MT systems. I learned from the horse's mouth, that at a tremendous expense, the machines churn out fifty pages a month and that the texts must be pre- and post-edited. I tried to obtain some additional information, such as the cost per page including the maintenance cost of the equipment and the number of people involved in the project. They did not know, and I don't think they really care about the European taxpayers who are paying for the experiment. One advanced the price of \$200+ a page ...

Recently I received a copy of *Electric Word*, which is published in the Netherlands. In it was an interesting article entitled, "Eight Case Histories from the Tower of Babel." Eight large international companies expressed their opinion and outlined their own experiences for dealing with several foreign languages. Except Philips, none were willing or

## NORTRANS

*Glossaries for the Year 2000!*

Perfect-Bound Hard-Copy Printouts  
Newly Developed Electronic Format  
for MTX Version 2

[Mercury/Termex™ (MTX) through LIS, Inc.  
Software Product of LinguaTech International]

Practical \* Proven \* Economical

Quality-Controlled Standardized Terminology!  
(Finance/Commerce, Mechanical  
Engineering, Metallurgy)

**FRENCH GERMAN ITALIAN  
PORTUGUESE SPANISH**

*True Translating Tools*

**R. M. Northrop**

To English Only 1118 West Lake Avenue  
No Price Change Peoria, IL 61614-5936  
(List on Request)

**(309) 685-8041**

even had the desire to use MT systems. All of them relied more on human translators working with computer assisted translation systems. These "biggies" are not interested in fancy experiments from academia, and they rely on their own experience and on the modification of existing systems which are tailored exclusively for their specific needs and suited for their in-house requirements. They emphasize training in company terminology and technical background, and writers and translators collaborate constantly. ICL, the British flagship computer maker, goes a step further and states that no terminology management software or computer-aided translation are used. Organizational skills rather than technical wizardry are the key cards in their game. So equipped, the team is confident that they can tackle anything.

I don't want to discourage the MT fanatics and experimenters, and it would be absurd and premature on my part to declare emphatically that machine translations will never replace translators and the PC. I just deplore the fact that for the last 12 years we have been told every day that we are dead or dying because MT is just around the corner. The plain fact is that there is no easy-to-use and cost effective system; freelance translators and even agencies should stay away from MT until more experiments have been conducted and major, honest and verifiable technical breakthroughs have been made: Believe me, that will not happen next year. □

... continued from *Vasconcellos on MT*, page 7

for the PC covering a wide range of quality, all the way down to simple-minded dictionary lookups (not to mention the hand-held toys being sold to tourists). In the meantime, the mainframe systems have continued to mature (only one of the companies has closed its doors), and there are also several new products running on intermediate "workstation" computers that are quite sophisticated. Although hardware complexity doesn't necessarily equate with product quality, the point is that today's PC is not yet host to a general-purpose system with the power and flexibility that a number of the others have.

While the MT picture has become increasingly muddled by the plethora of systems and products, there are a few basic principles that hold true across the board, and they are worth reviewing briefly.

The first principle is that output quality improves to the extent that the input text is constrained. High quality can now definitely be achieved with a linguistically robust system when the grammatical structures are fairly simple and when the vocabulary is limited, consistently used, and restricted to a single subject area. Proof of this fact is *Météo 2*, which has translated more than 100 million words of weather forecasts in Canada. In *Météo's* output, 96.3% is good enough to be used as is (Chandioux 1989). On the other hand, as texts become more general, quality starts to break down. However, the more viable systems--those with powerful grammars and parsers, large lexicons, and various approaches for picking context-sensitive glosses--can still be useful tools.

The \$64,000 question is useful for what and for whom. MT is not for everyone. Which brings us to the next key principle: the real value of an MT system depends on the requirements of the specific application. It is the entire process that is of interest, not just the product in isolation. Only after using a system can one know for certain how effective it is. A number of factors enter into the equation:

- o **Purpose of the translation.** By producing something automatically where nothing existed before, MT creates in fact the possibility of providing a less-than-ideal translation which may be of use to someone who needs to assimilate information (rather than disseminate it). When the name of the game is information, raw output or nearly raw output is now already sufficient to give monolingual scientists a competitive edge in the scanning of foreign technical literature. Bostad (1990) reports that at the Air Force Foreign Technology Division, Systran's raw output from Russian,

German, and French is being directly accessed by scientists as often as 625 times a month. In addition, the Division delivers 50,000 to 60,000 pages a year of material that has been partially postedited. Thanks to an automatic editing facility, only 20% of the output needs to be seen by posteditors.

On the other hand, there is translation that will be disseminated and is expected to undergo some degree of scrutiny. A translator is needed to capture the connotations so that the reader will understand precisely what the author intended to say: here the human interpretive role cannot be dispensed with. But MT can be enlisted to leverage the task. There have been many positive experiences with translators working as posteditors. Magnusson-Murray (1985), for example, has presented data showing that the productivity of translators more than doubled with the Weidner system used in a bureau setting. Gains achieved at four different sites with Logos are the subject of an article by Lawson (1984), and Logos is now being used in the Canadian Government (Lévy 1989). For my part, I kept both short- and long-term statistics over a period of eight and a half months at the Pan American Health Organization and found that very high levels of productivity could be achieved with Spanam and Engspan (Vasconcellos 1989).

- o **Nature of the source text.** We already know that text which is simply structured is easier for the machine. Of the more general texts, those that are technical documents are particularly good fodder for MT because research on terminology is greatly reduced and consistency and uniformity are enhanced. With text that is dull and repetitive--some of today's equipment manuals run 30,000 pages and more--MT speeds up the task and frees up the translator's creative energies for where they are needed most. The market for translations of this kind is growing rapidly.
- o **Languages and pre-editing.** When the source text is to be translated into several languages simultaneously ("one-to-many"), the added expense of pre-editing is often justified. In the one-to-one situation, the particular languages in question may also affect the need for pre-editing. In the case of Japanese-to-English, for example, there is considerable linguistic distance to be bridged, coupled with a priority need for information and a shortage of traditional translators. In these circumstances, the added investment of pre-editing pays more dividends than it would for language pairs that are less at a premium and closer in structure--in which

case pre-editing is the exception rather than the rule.

- o **System potential.** What separates the men from the boys in the MT world is the system's potential for being tailored to a given user's needs (Vasconcellos 1990). The assessment of this potential is a complex issue and goes beyond the scope of the present debate.
- o **The human element.** After all is said and done, MT is people-driven. A system's effectiveness will depend in large measure on the attitude of those who deal directly with the output--in many cases, translators-turned-posteditors. Postediting is a special skill; not every translator takes to it readily. And there's nothing wrong with that. In my experience training more than 30 posteditors (self-selected volunteers) over the last 10 years, I have found that about half of them are able to adapt to MT and half are not. Interestingly, the free lances who do postedit are willing to accept 40% less pay because they find that they still come out ahead.

Our task, as translation professionals and members of ATA, is not to avoid MT but rather to get involved in it. Progress in MT is incremental; there are no major breakthroughs to wait for. What MT needs is the long-term commitment of translators to develop replicable strategies for postediting, exercise the systems, build dictionaries, and provide feedback to developers. The more they participate, the sooner they will have a tool they can really use.

#### REFERENCES

- Bostad, Dale A. 1990. "Aspects of Machine Translation in the United States Air Force." In: *Benefit of Computer Assisted Translation to Information Managers and End-Users*. North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development, 1990. (AGARD Lecture Series No. 171)
- Chandioux, John. 1989. "Météo: 100 Million Words Later." In: *Proc American Translators Association Conference 1989* (Washington, D.C., 1989), ed. Deanna L. Hammond. Medford, NJ: Learned Information, 1989. pp. 449-453.
- Lawson, Veronica. 1984. "Users of Machine Translation System Report Increased Output." *Language Monthly* 11:6-10.
- Lévy, Moise. 1989. "Consolidating a Machine Translation Project at the Post-Implementation Stage." In: *Proc American Translators Association Conference 1989* (Washington, D.C., 1989), ed. Deanna L. Hammond. Medford (NJ): Learned Information. pp. 479-487.

---

Magnusson-Murray, Ulla. 1985. "Operational Experience of the Machine Translation Service." In: *Tools for the Trade: Translating and the Computer 5*, ed. Veronica Lawson. London: Aslib. pp. 171-180.

Vasconcellos, Muriel. 1989. "Long-Term Data for an MT Policy." *Literary and Linguistic Computing* 4(3):203-213.

Vasconcellos, Muriel. 1990. "Perspectives on the Assessment of Machine-Translated Output." In: *Miscellany on Translation Criticism*, ed. Milan Hrala. Fédération internationale des traducteurs. In press.