

## FAMT IS ALIVE AND WELL

**Marjorie León**

Chief, Translation Services

525 23rd Street, N.W.

Pan American Health Organization (PAHO)

Washington, DC, 20037 USA

leonmarj@paho.org

### Abstract

This invited talk describes the use of fully automatic machine translation (FAMT) at the Pan American Health Organization. Statistics covering 1998 are presented and analyzed in terms of productivity and cost savings. Feedback from several outside users of PAHO's translation software is also reported. Problems encountered in implementing machine translation in an international organization are discussed from the points of view of managers, translators, and end users. The talk concludes with a quick glimpse at what PAHO's MT development staff has been working on this year.

### Introduction

In April of this year, I attended the meeting of the Joint Inter-Agency Committee on Computer-Assisted Translation and Terminology (JIAMCATT) in Geneva, Switzerland. The group fosters discussions and exchanges between members of language services units. United Nations agencies are facing serious budget restrictions these days, and managers have zeroed in on translation as one area where costs might be cut. This year's JIAMCATT meeting focused on computer tools such as translation memory, terminology data banks, and speech recognition systems. Those tools are designed to be operated in real-time by a live translator. PAHO's fully automatic machine translation system was the exception. Many members of the group had long since lost interest in FAMT, and they were surprised to learn that PAHO was using it successfully.

### 1 Why Does MT Work at PAHO?

Why does MT work at the Pan American Health Organization when most other international organizations haven't been able to make it work? I would like to try to shed some light on this question. Let me mention some of the key ingredients of our success. They are not listed in order of priority, because they are all essential parts of the picture.

#### Our Own Software

PAHO uses its own proprietary MT software: Spanish to English (SPANAM®) and English to Spanish (ENGSPAN®). The software is maintained and enhanced by our staff computational linguists. The program has been tuned to the types of text we usually encounter, and we can introduce improvements and fix problems when the need arises.

#### Management Commitment

After about 10 years of in-house software development, a controlled experiment showed that the quality of postedited MT was acceptable and that MT had the potential to produce cost savings through increased productivity. In order to realize this potential, PAHO's Director approved a general directive in 1989 which established MT as the primary mode of translation to be used by Translation Services. This commitment has been reconfirmed this year; the revised translation policy now provides for use of MT by the technical and administrative units at Headquarters and in our field offices.

#### Electronic Input

Our policy requires that the units requesting translation services provide an electronic copy of the text, whenever possible. Since much of our work is generated by PAHO staff, and the same word-processing software is used throughout the Organization, obtaining electronic input is not difficult. For documents from outside sources, we urge the units to obtain the electronic version, especially if the work was done under contract to PAHO. The time it takes to locate the file is more than offset by the speed with which the final translation can be produced.

#### Translator Skills

Our translation unit has two Spanish translators and one English translator. They are skilled professionals who also know how to use word-processing software. They are committed to delivering the highest quality translation that time permits, and they do it by postediting the raw machine output on the screen. Dictation is not an option, because we have no transcribers. Since we have such a small in-house staff, we outsource a large amount of postediting to freelance translators. Our staff translators review the contractor's

work and revise it on the screen. We have no typing pool.

**Support Staff**

This does not mean, however, that our translators are responsible for all aspects of MT processing. Our three office assistants play an important role. They are the ones who run the translation and provide the translator with the output in a standard word-processing document. The translator does not even have to learn how to operate the MT software. Clerical support is also available for coping with complicated formats.

**Customized Dictionaries**

The PAHO MT dictionaries are the result of 25 years of input from translators and linguists. The lexicon has been gleaned from the more than 40 million words that have been translated using the software and postedited by staff and contract translators. Each set of dictionaries contains over 80,000 stem entries, expressions, and context-sensitive rules. The dictionaries are updated daily, based on the feedback received from the translators.

**Broad Domain**

Although the capabilities of our software and the contents of its dictionaries reflect the types of text that they have been exposed to, the system is robust. It is not restricted to a particular text type or sublanguage. PAHO uses MT for many different types of text. In 1998, the unit translated a total of 2,977,763 words between English and Spanish, and 90% of them were done using MT.

Some people have the impression that the PAHO system is only for medical texts. While it is true that our dictionaries contain a lot of medical terms, the system performs well on texts in many other domains. We routinely texts dealing with many aspects of public health, including political, economic, and social issues. They are written for varied audiences: our Governing Bodies, other international forums, funding agencies, national policy-makers, scientists, health professionals, students, patients, and the media. Microglossaries can be defined to handle terms that belong to different subject areas or registers.

**Standard Syntax**

The system expects standard syntax. No pre-editing of the source text is required. Before running a translation, we check the input file for formatting problems, verify that the electronic copy matches the hard copy, run a spelling check, and block sections of the text that should not be translated. If appropriate, the operator can specify a text type in order to tell the program to expect a special type of syntax.

**2 Cost Savings for PAHO**

This year we were asked to participate in a survey of translation costs in international organizations. Our average cost per 1,000 words in 1998, including all salaries, benefits, contract costs, and overhead, came to US\$ 332. While this may sound high to an

independent translator, it was the lowest of all the organizations that participated in the survey.

Recently our Director asked us how much money machine translation had saved the Organization in 1998. Using a conservative approach, we came up with the following figures:

Number of words outsourced	2,331.000
Average cost of human translation per 1,000 words	US\$ 110
Potential cost without MT	US\$ 256,410
Actual cost	189,130
Saving on outsourced work	<u>67,280</u>

Number of words done by staff	646,763
Estimated increase in productivity using MT (33%)	215,590
Average cost of outsourced HT	110
Potential cost without MT	23,715
Actual expenditure	0
Savings from increased productivity	<u>23,715</u>

Savings on professional translation in 1998 90,995

These figures are conservative because the actual cost of outsourced human translation is closer to US\$ 120 per 1,000 words, and we sometimes must pay another 15% for rush work. The current rate for postediting the raw MT output is either US\$75 or US\$84 per 1,000 words

**3 Empowering the Nontranslator**

Another reason why these figures are conservative is that they do not include the value of the machine translations done by nontranslators in technical and administrative units at PAHO Headquarters and in our field offices. We know that staff at Headquarters used the system to translate over 2 million words last year. This year they are on a pace to translate more than 3 million.

Fully automatic machine translation can be of great assistance to the nontranslator, especially in an organization like PAHO where most of the staff have at least a working knowledge of both Spanish and English. Correspondence and reports often need to be prepared in both languages. Whether we like it or not, secretaries and professionals often do their own translations.

By giving these individuals access to the translation software, we are giving them instant access to the terminology that we have worked so hard to feed into the MT dictionaries. If we do a good job, we can get a multiplier effect that would never be possible with other tools that require a professional translator in attendance.

We can't be sure how nontranslators use the MT output, but the high volume and number of repeat users

would indicate that the software is helping people save either time or money, or both.

#### **4 The Productivity Issue**

Attempts to introduce new computer tools for translators have generally met with strong resistance in international organizations. Managers are trying to cut costs either by increasing the overall productivity of the translation process or by outsourcing translations instead of hiring in-house staff. New demands are being placed on salaried translators.

Staff translators who use dictation and transcription have welcomed online access to terminology and background references. On the other hand, they have not embraced translation memory systems or machine translation, both of which require more time in front of the computer screen and more use of the keyboard. On the other hand, translators who are accustomed to typing their own draft translations find that machine translation can make them more productive.

Why is it, then, that some translators claim that it is faster to translate from scratch than to postedit the machine output? Perhaps they need to improve their word-processing skills. Or maybe they need to give the machine dictionaries an injection of their favorite terminology.

#### **5 Views from other PAHO Users**

We have several different categories of users in addition to our staff translators. Some of our contract translators report that the MT output has helped them complete the work twice as fast. Others report a 20-30% gain in productivity. Occasionally, someone will complain that the MT output slowed them down. If this happens, we usually adjust the rate and have the computational linguist take a good look at the side-by-side to discover the cause of the problem.

PAHO staff who access the system directly are our most enthusiastic users and our best advocates. They have spread the word about the system to their colleagues in the field and in other institutions. If they lose the MT icon or toolbar for any reason, they call us immediately to help them find it

We know that most of our outside licensees are using the software, because they call us to get an upgrade when they migrate to Windows NT. And we continue to add new clients who find out about us by word of mouth. A private consulting firm licensed ENGSPAN to translate its project proposals in a big hurry. We installed both systems for the Ministry of Health of Mexico and trained their staff so that they can use it to translate the documents for a global conference on health promotion that will be held in the year 2000. A freelance translator in the Washington DC area recently licensed SPANAM, took the dictionary update training, and has begun to add specialized terminology to the dictionaries.

#### **6 Views from PAHO's MT Developers**

What has made the PAHO MT system work is the skill, dedication, and hard work of our computational linguists, translators, and programmers. Our challenge this year has been to migrate the system to the 32-bit Windows platform and make it as user-friendly as the newest packages on the market. We have designed a completely new dictionary update capability, and we are confident that our users will be able to learn to introduce their terminology quickly and accurately. The conversion effort represents a considerable investment on the part of the Organization. We are proud that SPANAM and ENGSPAN have earned their place on the PAHO desktop.