

# Is MT Right for You?

MURIEL VASCONCELLOS

**Y**ou need to keep several points in mind when deciding if MT is right for you. First, you must determine if you have an application for which MT is appropriate. It's important to pick your application and then decide on your system rather than vice versa.

Costs soar when the input isn't in machine-readable form, and an OCR (optical character recognition) device, while helpful, isn't a panacea. If your documents aren't in electronic form, you may want to think twice about using MT. In addition, there should be a large volume of material (e.g., 100,000 words per month) to be translated, with the expectation that more will be coming from the same source. In the beginning, there should be only one domain (i.e., subject matter); you can branch out later as you become more familiar with all the ins and outs.

Your decision to use MT will hinge, in large part, on the format, volume, and linguistic characteristics of the source-language text. The text should contain no ambiguities.

## Hardware and Human Factors

You need a hardware platform that an MT system will run on. In a multiuser setting, you must be sure that there is

good word processing support and that all the users have strong word processing skills. Multitasking workstations designed for translators are helpful.

Be sure to recruit people who have a positive attitude about using MT. This is especially important during the first few months while you are getting your system up and running. This stage involves customizing the dictionaries and gaining proficiency in postediting.

Your choice of a system will depend on the characteristics of your application, so it's important to identify criteria that are specific to your needs. And don't be tempted to buy software just because it's inexpensive. As with a house pet, the price you pay up front is a drop in the bucket compared to the cost of the care and feeding for the rest of your mutual lives. For example, a less expensive system might cost more in terms of support personnel and customer support (some MT companies charge you for it). Also, if your time frame for translations is tight and if your budget for human intervention is limited, it's crucial that you test the system's performance on randomly selected texts.

Despite decades of scientific study, the evaluation of translations is an uncertain exercise. The definition of an

error will vary, depending on the purpose of the translation and the values of the end users. Errors in raw output are important mainly to the analyst, who knows the inner workings of MT systems and can classify the error types according to their causes. Such an analysis can tell something about the system's potential and the effort that will be required to fix and to maintain it. You should make sure that you compare outputs from different systems produced under the same conditions.

The value of a system depends on its potential to grow and to improve its performance, as well as how easy it is to use and to maintain. It's important to know the language combinations that have been developed for the system, the size of the dictionaries or knowledge bases, the ease with which you can add to the dictionaries, and the possibilities of extending the system to include the domain that you are interested in.

---

*Muriel Vasconcellos is president of the Association for Machine Translation in the Americas and is a Washington, D.C.-based consultant in translation and machine translation. You can reach her on BIX c/o "editors" or on Compu-Serve at 71024,123.*

is essential with any MT product for personal computers, because it lets you add your own terminology to the program. PC-Translator simplifies the creation of your own dictionaries by importing lists of terms in ASCII format directly into the software. In addition, MT software generates lists of words not found in a given text to help you customize your software. You decide which words and phrases to add to the dictionaries. The ability to add, delete, or modify dictionary entries dramatically improves the quality of a translation and reduces the time spent postediting an output. You'll find that it can take from two to four weeks to customize a system.

All these systems ask you to insert the part of speech of the word you are adding and to provide its translation. With exten-

sive dictionary coding, the system can deal with ambiguities that arise from the use of words that can take the form of multiple parts of speech. For example, the program will recognize the different translations of a homograph (i.e., a word that is spelled like another but has a different meaning or pronunciation) used as a verb and as a noun in the same sentence (e.g., "The can can explode"). Because Globalink's GTS-Professional can classify *can* as both a verb and a noun, it's better able to translate the sentence than a product that requires less dictionary coding.

## Workstation-Based MT Products

MT workstation products are designed to handle heavy volume—when you have to translate 2000 or more pages of text per

year. Translation speeds range from 20,000 to 1 million words per hour.

A workstation MT system is a large investment. Software prices start at \$10,000. A system can cost several hundred thousand dollars, and pricing structures are as diverse as the possible configurations.

Socatra (Quebec, Canada) spent over 12 years preparing its XLT computer-assisted translation system for the commercial market. Access to XLT is uniquely controlled by the company. Socatra rents software for a specific number of words. After you pay an initial subscription, Socatra provides you with the software and an access card, which resembles a credit card. The card contains a microprocessor that counts the words translated and acts as a security device. You can obtain an XLT