[*MT evaluation: basis for future directions*. Proceedings of a workshop...2-3 November 1992, San Diego]

ANNEX A

MACHINE TRANSLATION EVALUATION CRITERIA

Ralph B.Dessau *

1.0 INTRODUCTION

The last two years have yielded vast improvements for MT with regard to both hardware and software. Faster hardware and less expensive data storage have reduced costs, while software improvements have boosted its efficiency.

Precise translation is no longer enough. Today's MT users often spend vast amounts of time and money on Desk Top Publishing by converting ordinary texts to DTP formats before they have them translated. DTP is sophisticated word processing and a modern equivalent of typesetting. In order to produce elaborate composites of text, tables and graphics, DTP files contain hundreds of invisible formatting codes, which must remain intact to retain the format in the MT output and avoid DTP duplication. Some recent versions of MT systems can directly translate DTP files from word processors such as WordPerfect and MS Word.

MT is still primarily being used by international organizations for commercial and technical documentation. The number of MT products is limited, but costs vary considerably. Experienced MT users are rare, and few are familiar with more than one product. These conditions have created the need for suitable evaluation criteria. MT products claim different features and employ different techniques. Users need to know how to select the best system for their particular applications.

2.0 HARDWARE

2.1 HARDWARE TYPE

Hardware can be a major cost factor. Some systems require a mainframe. Others run on work stations or on dedicated Minicomputers. As far as I know, there is only one system available for the Macintosh. Most PC systems run under DOS, but can also be adapted to the PS/2. Keep in mind, however, that the hardware selection must be based on the requirements of the preferred MT software.

2.2 HARDWARE MEMORY

Memory requirements are crucial for PC compatible systems. A harddisk is normally a must. In addition, most programs require 640 K RAM and some are able to utilize expanded or extended memory (EMM) .

2.3 HARDWARE COST

Cost is very much a function of type selection. High-cost Mainframes and Minis are difficult to justify for most translation applications. PC costs are easier to justify, although prices vary significantly from advanced types, such as the PS/2 and 486 models down to the basic AT. A 286 PC is currently the lowest cost unit worth considering. ATs should no longer be considered adequate for MT applications.

3.0 SOFTWARE

3.1 AVAILABILITY OF LANGUAGE PAIRS

The availability of language pairs should be the first consideration in selecting an MT system. By using the same system to translate into several languages, the difficult words and phrases encountered in the first language need only be translated and entered in the dictionaries of the other ones to produce excellent result with minimum effort.

3.2 OPERATING SYSTEM

The next step is to identify which operating system is required by the preferred software. MS DOS is the most common for PCs, followed by PS/2, UNIX and Macintosh. MS Windows is another possibility but tends to reduce MT speed and is not in great demand. Digital Equipment's VAX system is dominant for Minicomputers. A few systems run on minis, such as SYMBOLICS or VAX and some require mainframes.

3.3 OPERATING MODE

Some systems operate on-line, while others run in batch mode. The latter allows operators to schedule translation tasks to be performed at night or at other times, when the computer is not in other use. Interactive mode requires the presence of an operator.

3.4 GRAMMATICAL CAPABILITIES

Grammar features receive the most scrutiny from people with a language oriented background. MT systems normally handle the different language features of syntax and grammar, either by means of algorithms in the program, Wildcard Phrases in the dictionaries or a combination of both. The algorithms or wildcards control the grammatical correctness of the output. In advanced systems, they provide contextual analysis, conjugate verbs and determine the gender of nouns and adjectives. By improving output quality they reduce the need for post editing.

3.5 SOURCE TEXT REQUIREMENTS

Some systems can directly process word processor files. Others require conversion to ASCII. This is a major drawback, as the output format is lost and must be restored. Other systems limit the text size, so large documents must be divided into fragments, or they impose restrictions on the vocabulary or grammatical complexity of the source text. Some use markers around text or tables to be left untouched, but error-free source text is always a must. No system can cope with bad spelling or typographical errors.

3.6 SOURCE TEXT ENTRY

MT systems can directly process text files in computer memory or on diskette. This eliminates keyboard entry and changes a translator's role to that of enhancing the MT dictionaries and editing the output. It also permits text entry by modem and OCR systems. When this feature is ignored by source text originators, and they provide source text in document form, they forego the savings of processing digitized files.

3.7 OUTPUT APPEARANCE

With DTP, output appearance is a major consideration. In ASCII format, even high quality output in terms of grammar and context will need considerable post editing to restore the source document format. In simple terms, a translated letter may no longer look like a letter and will require reformatting. Desk Top Publishing poses a major challenge to MT, as DTP documents contain hundreds of invisible formatting codes which the MT process must leave untouched to maintain the DTP format in the output. Important questions are:

- Is page formatting maintained ?
- Is upper/lower case lettering maintained ?
- Does the system recognize and generate all the required characters for source and target text (Ä, Æ, æ, Å, å, Ç, Ö, É, é, í, ó, Ñ, ñ, etc.) ?
- Are non-letter characters and punctuation marks left intact. (Numbers, &, @, #, (), $, %, etc.) ?
- What happens to under-lining, hyphenation and numerical tables ?

An easy way to test whether an MT system maintains WordPerfect format is to make it translate the WordPerfect printer test file. Most Word processing software provide such files. A WordPerfect user needs only identify the file called PRINTER.TST, order the MT system to translate it, and compare the printed output to the original.

3.8 MT DICTIONARIES

Dictionaries are a vital ingredient. Many believe the number of words is important, but this can lead to the adaptation of huge vocabularies with useless words, which slow the system down. Instead, MT systems need the right words, especially for professional and even company-specific terms. MT suppliers can't provide them, so users must create their own specialized glossaries of words and phrases. The ease of enhancing MT dictionaries is therefore more important than their initial size.

MT Dictionaries must cover both single words and phrases, where Phrases are sequences of two or more words with a specific meaning. Users must control the glossaries for each topic and be able to define the search order of different glossaries. MT systems employ different methods for storing, sorting and searching their dictionaries. Without going overboard, it is useful to examine their features and possible limitations with regard to:

- Capacity for additional words and phrases
- Dictionary Format (ASCII, Coded, Compressed, etc.)
- Sorting, Indexing and Compression Methods and Duration
- Number of User Selective Dictionaries
- Wildcard Phrase Capability (Single/Multiple)
- Ease of Dictionary Enhancement
- Search Utilities (Word Search, Phrase Search)
- Editing Utilities (List, Edit, Add, Delete)

3.9 MEMORY REQUIREMENTS

Dictionary size and format impact on the system's memory requirements and limit the selection of computer hardware. Most systems call for at least 640 K of RAM and 2-3 MB of harddisk space. Some may further be able to employ extended or expanded memory (EMS Compatibility) .

4.0 SPEED

The speed of MT systems is often an unnecessary concern. Computers don't sleep and can work at night. At only ten type-written pages per hour, the output from eight hours processing will amount to eighty pages. This by far exceeds anybody's output by conventional methods as well as the post-editing capacity of an expert translator. It is again worth noting, that output quality is far more important than quantity.

4.1 COST

The price of an MT system is important but should not be the only cost consideration. Some software requires costlier hardware. Training costs must be added. The higher post-editing needs of a less expensive system may obviate a price advantage. The cost of future updates and discount savings for multiple users also affect the overall cost.

5.0 BENEFITS

The benefits of MT systems are easy to define and quantify. When used correctly, a good system will:

- Improve Productivity - Reduce Typing
- Organize Dictionary Work
- Reduce Translation Delays
- Ensure Adherence to Terminology and output style
- Prioritize Dictionaries by Topics
- Eliminate word search duplication
- Maintain DTP formats - eliminate DTP duplication
- Work unattended - Improve Computer Utilization

5.1 COST SAVINGS

MT Cost savings depends on several factors:

- Hardware and Software Costs
- Training and Operating Costs
- Revenue loss during the learning period
- Increased Productivity - Fewer Delays

A PC-based system should double the output of average users and may triple that of experienced users. The resultant productivity increase may, in turn, eliminate the need for additional translators to cope with higher work loads. MT has the immediate advantage of eliminating a laborious initial keyboard entry, when source text is available in digitized form (in memory, on diskette or by modem).

When translation costs are charged at fixed rates per word or page, it is easy to calculate the effect on revenues from higher productivity. If a translator, making ten cents per word, boosts his daily output by 1,000 words after switching to MT, his added weekly earnings will amount to $500. Even allowing for a $5000 revenue loss during the inevitable learning period, he or she would easily recover the total $7,500 cost of a complete system, including a PC, printer, OCR scanner and MT software within a year.

The larger, more expensive systems can only be justified for very high volumes of translation, when sufficient savings are achieved by lower labor costs and greater productivity. Their cost effectiveness will therefore largely depend on their ability to reduce post-editing.

When translation bottlenecks delay the completion of technical manuals or proposals, shipping documents and invoices, such delays can produce serious losses. Higher productivity reduces delays and helps prevent such losses.

5 . 2 TERMINOLOGY

Another MT advantage is strict adherence to established terminology. Many users translated by conventional methods before switching to MT, and used glossaries of foreign terms and phrases to ensure consistent output. MT has changed the translator's role to that of enhancing the glossaries and post-editing the output. Consistent Terminology and output is ensured by the MT system.

October 1992.

* Ralph Dessau
Linguistic Products
P.O. Box 8263
The Woodlands, TX 77387