Operational experience of a machine translation service

Ulla Magnusson-Murray

*Manager, ITT Technical Translations and Linguistics Centre,
ITT Europe Engineering Support Centre, Harlow, Essex,
UK*

This report on eighteen months of experience with a computer-assisted translation system covers the organisation of the ITT Technical Translations and Linguistics Centre, the service provided and the performance of the system. Cost-effectiveness is noted, as are improvements with increased experience. Other factors discussed are training, document handling and systems support.

PROLOGUE

Those of us who enjoyed last year's conference, <u>Term Banks for Tomorrow's World</u>, may have left with a picture of the traditional translator which looks something like this:

- a female language graduate
- equipped with a shoebox of terminology cards
- sweating over drawings of welding equipment (why didn't Daddy teach me how to weld?)
- scanning international standards and vocabularies which are five years out of date
- has never been asked to sit on any decision-making committee regarding terminology and documentation or translations
- is either a freelance, isolated from the companies she works for, or
- a staff translator, isolated from the decision-making process.

A different picture, however, soon emerged over dinner <u>after</u> the proceedings. A small group of people concerned

with the practical experience or the development of computer-assisted translation in systems such as ALPS, Logos, Systran and Weidner were seated around a table, together, quite by chance. The picture of the shoeboxes and obsolete dictionaries now gave place to a translator surrounded by computer terminals, data links, floppy disks and online dictionaries.

Although this may represent, to some of you, a somewhat futuristic translator, it is true of the situation at ITT Europe in Harlow today. Whether this brings translators any nearer to being recognised as part of the current technological evolution, though, is debatable and will not form part of my talk today.

INTRODUCTION

Within my present brief, I will discuss the organisation and work at the ITT Technical Translations and Linguistics Centre, and present a review of early experiences with MT. We use the Weidner computer-assisted translation system.

ORGANISATION AND WORK

We are a group of ten in-house translators/linguists turned MT editors, supplemented by a network of freelance technical translators. All translators have, of course, the language they translate into as their mother tongue. Although we have been established for six years, we have only recently been operating as a bureau service. We produce translations of customer documentation, tenders, contracts, research articles and some commercial material. Virtually all major European languages are dealt with by us in-house, whereas our freelance translators offer invaluable help with more obscure translation requests.

As you can see from Figure 1, we are not limited to translation services, but offer complete publication facilities. The maintenance of terminology, courses on English for special purposes, as well as the development of controlled technical language also come under the Centre's management.

We have three main areas of specialisation, viz. electronics, telecommunications and computer technology. The fact that our subjects are so specialised led us to seriously consider the use of computers in the translation process.

We have effectively been working with our MT system for some eighteen months now. Our present configuration is shown in Figure 2.

An MT editor is responsible for dictionary entry for his particular language pair. The translators and editors rely
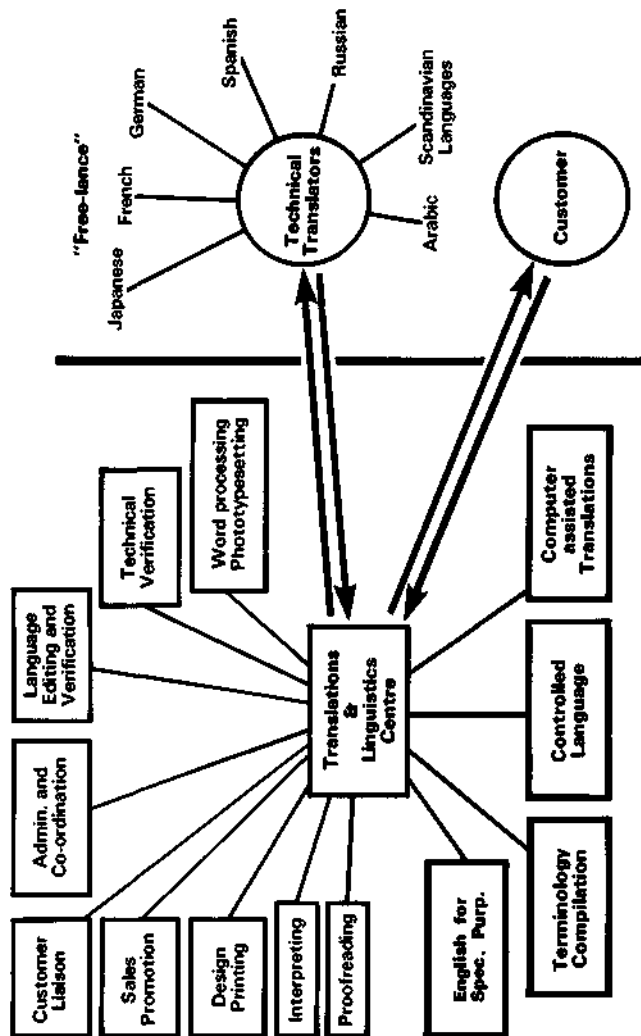
Figure 1.   Services offered by the ITT Technical Translations and   Linguistics Centre

- **Vax 11/780**

- **5 Hewlett-Packard terminals**

- **1 multi-lingual printer**

- **English/French/English**
  English/German/English   **Software**
  English/Spanish/English   **Packages**

- **Word processor link to VAX**

- **Word processor link to phototypesetter**

- **Access to the ITT Network**

- **Access to the X25 packet switching**
  network

- **A number of translators/editors with**
  square eyes, enthusiasm and a friendly
  disposition towards computers

Figure 2. ITT's configuration

on a terminologist to co-ordinate the vocabulary and to carry out research on company-specific and product-specific terms. Our first priority, when we start co-operating with a client, is to agree on the preferred vocabulary. It involves a great deal of effort, initially, and is costly, but it is absolutely necessary and will certainly avoid dialogues such as the following:

French Client: You 'ave translated 'hole generation' with 'production de trous'!

We: Well, yes, it seems a reasonable translation.

French Client: But, non! 'ow could you possibly manufacture a hole! ?

Until recently, our present departmental structure and work procedures have been sufficient to manage the work-load. Our procedures are shown schematically in Figure 3.

However, as a result of our experience with the MT system, we are continuously monitoring performance and modifying our way of operation. I shall deal with this topic in my conclusion.
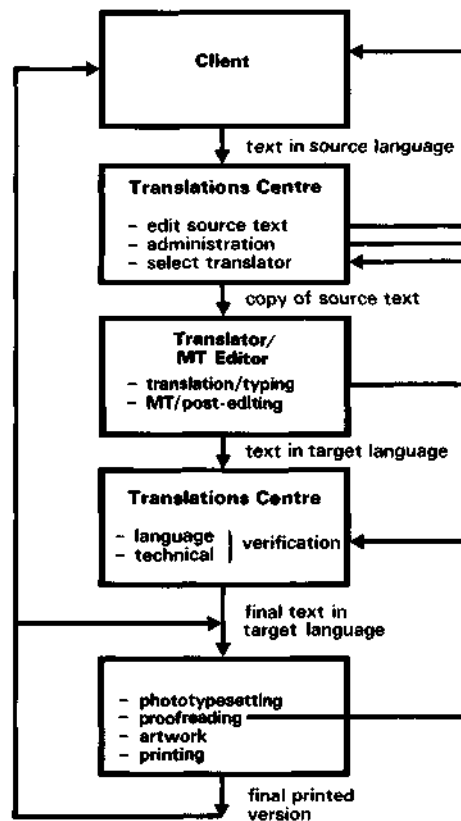
Figure 3. ITT's work procedures

Most of our clients look to us to solve their entire publication problem, i.e. to process the complete manual from translation through to printed copy. We are at present planning to install a fully integrated electronic documentation system (see Figure 4). The typesetting interface is already available and enables us to transmit translated copy electronically to the typesetting equipment.

THE MT EXPERIENCE

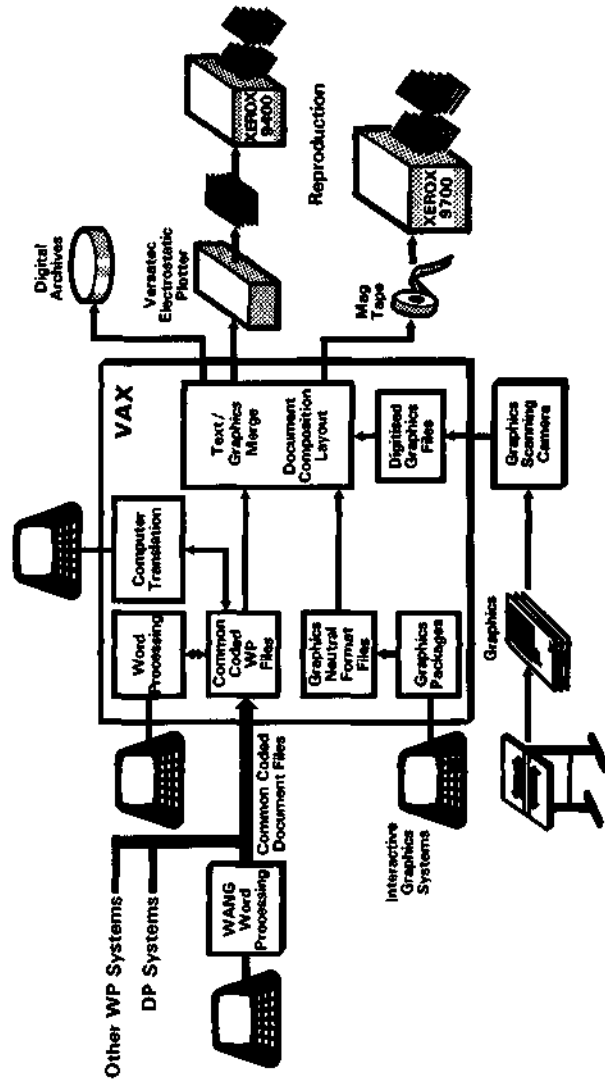'Enough', I hear you say, 'of that commercial schmaltz - does the system work?' Yes, it does - not to say that we

Figure 4.   Proposed integrated electronic documentation system

do not have problems.   Here are some examples:

- some clients' word processing formatting procedures do not match ours,   although we have identical word processing systems.
- by inputting ae, oe and ue in the German source text, instead of ä, ö and ü, a client delayed the completion of a translation quite considerably.
- misspelt words (and there are many of those) in a source text, received on diskette or electronically,   are thrown out by the system as unrecognised words.
- incompatibility between word processing systems at present limits our range of clients.

Compatibility and standardisation in these areas are now becoming vital issues and cannot be stressed enough.

It could be argued that what I have described above has little to do with the actual translation. However, the change in emphasis in the translating domain now demands that we can master anything from word processing routines to trouble-shooting to backing-up of disks.

During the first six months of the fully operational installation, we conducted an evaluation and feasibility study, covering not only quality of output, but also communication facilities, system support, ergonomic aspects, usefulness of raw translations and so on. The study was for the benefit of ITT and the prime objective was to investigate the practical and economic aspects, in order to assess the merits of installing such a system. For the purposes of this paper, only a few aspects of the study will be considered.

One of the prime tasks was to prove that it would be more economical to translate by computer than by conventional means. During the experiment, documents amounting to 13,000 words in total were translated manually and by computer. The texts were different in length, subject-matter (all electronics, however), syntax and style. Some of the MT editors had many years' experience of technical translation, while others were less familiar with the subjects. The manual translators were all experienced technical translators. The figures we arrived at for <u>manual</u> translation times were revealing although not surprising. As a result of the analysis, we discovered that the various sub-tasks involved in producing a <u>manual</u> translation were time-consuming and resulted in a more expensive final translation than we had previously realised. So much for the good news!

Our expectations of computer-assisted translations were obviously high. However, when the early results were

analysed, we found that we could not achieve these expectations overnight. While the MT editor was gaining experience with all aspects of the system, there was a gradual improvement in the ratio of MT to manual translation times. This improvement was not as significant as originally expected and eventually levelled off at 2:1. There is a variety of reasons for this:

(1) the difference in subject-matters in each document meant that the editors had to familiarise themselves with new terminology before each translation.

(2) the source text was often badly structured, resulting in bad raw translation which required much post-editing.

(3) the timings registered for computer-assisted translations included dictionary update, idiom entry and CPU time, in addition to straight post-editing.

(4) the editors were still at a stage where they were learning about all aspects of computer-assisted translations, from theory of grammar to man-machine interaction.

(5) there was a tendency, among most of the editors, to over-polish the translations.

CONCLUSIONS

During this year we have continued to measure computer-assisted translations of documents of a similar nature. The conclusion to be drawn from this is that, for a given text type, there is a marked improvement as the editor becomes familiar with the subject. Also, the larger the text, the less significant the time taken for dictionary update - hence faster post-editing.

It is evident that the efficiency of machine translation is heavily dependent on the proficiency of the MT editor, the capability of electronic transmission of text and, to a significant extent, the quality of the source input. To maximise the editor's performance and to render the whole translation process more cost-effective, new approaches, however, must be considered. Figure 5 illustrates my immediate plans.

In practice, this means that pre-editing (if any), dictionary entry and the translation process will take place on our premises, and post-editing will be handled either by our freelance staff in their homes or in-house.
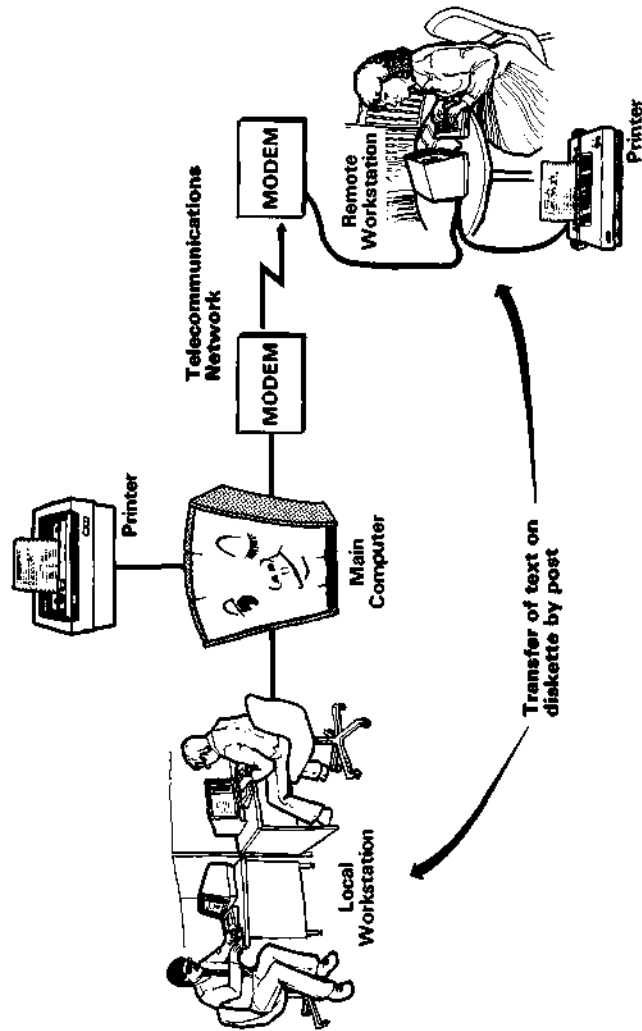
Figure 5. Proposed change of approach for ITT

FINAL REMARKS

As a result of our experiences with MT, I feel confident in stating that it is a technique which will find more and more application in the future. In this context, it may be useful to end with a few crucial points:

1. Be sure the system you invest in matches your requirements .

2. Make sure training, maintenance and good documentation are provided by the suppliers.

3. Select your translator/editors meticulously, since MT work is very specialised.

4. The investment in establishing a workable department is quite significant in terms of time, money and effort - quite apart from the cost of the system.

5. The editor's proficiency will increase with time until eventually a plateau is reached. However, one editor will always be more proficient than another.

6. As the size of your dictionary increases, the translating and editing time is significantly reduced and the system becomes more cost-effective.

7. It is important to appreciate the linguistic limitations of computer-assisted translations.

8. The cost of manual translation is increasing all the time - data processing costs are more static.

9. Although you can expect to at least <u>double</u> your translator's output, the real cost-saving in MT lies in complete electronic transfer of information and the integration into a fully electronic publishing system.

AUTHOR

Ulla Magnusson-Murray, Manager, ITT Technical Translations & Linguistics Centre, ITT Europe Engineering Support Centre, Great Eastern House, Edinburgh Way, Harlow, Essex CM20 2BN, UK.