EXPERIENCE IN ENGLISH-FRENCH POST-EDITING

Bernard Lavorel,
Commission of the European Communities, Luxembourg

A survey of post-editing work undertaken by translators from the
Commission's French translation section in Luxembourg is presented in
two parts, one based on the experience of one translator (the author)
who handled some 150 pages of M.T. output over a two-year period, the
other covering the reactions of six translators who post-edited about
100 pages over a period of ten weeks.

The translators were asked to list the M.T. errors they found the most
troublesome.

On the whole the same errors were listed in both parts but in a
different order of gravity. Irritation appeared to be the main
criterion used by the group of six, who listed the lack of common
terminology as the most troublesome source of error whereas the author,
having gained some understanding of typical M.T. performance, was more
concerned by problems of a syntactical nature.

It would appear that some knowledge of the limits and capabilities of
machine translation as well as experience of revision or post-editing
itself could lead to a more pragmatic approach to translation as a joint
man-machine effort.

INTRODUCTION

In order to gain some insight into those problems which cause translators the
most trouble in the post-editing of machine-translated texts, a two-part
survey of translators' reactions was recently conducted in the French section
of the Commission's translation division in Luxembourg.

The first part of the survey related to the author's own experience of
post-editing over a two-year period (1979-1981) during which he handled 17
documents averaging nine pages each from the field of information science and
communications.

The second part concerned the reactions of six translators who had post-edited
nine documents over a period of two months in 1981. These covered various
subjects, mainly of a technical nature (mining, cables, pollution) and
averaged about 11 pages each.

PART I OF THE SURVEY

In the part of the survey relating to the author's experience, error
classification was based on the comments made after each piece of
post-editing. The error types found fall into four categories, ranging from
the most troublesome to the least troublesome.

1. Verb forms. The translation of verbs often produced renderings which were
unacceptable in French.

Examples.

A number of new languages are expected to acquire the status of Community language.	Un certain nombre de nouvelles langues <u>se sont attendues</u> à ce que acquièrent le statut...

Concerns were expressed that no mention had been made...	<u>Des soucis ont été exprimés</u> qu'aucune mention n'avait été faite...

This type of error was frequent and could only be corrected by rewriting the sentence using another construction such as "On s'attend à ce que de nouvelles langues..." in the first example and "On a deploré que..." in the second.

Very often the use of a passive form in English required a different verb form in French, sometimes involving minor changes as in

They will be asked to...	Ils seront demandés de...
(altered to)	Ils seront invités à...

However, in some cases a fair amount of work was involved as in

A strong recommendation was expressed in favour of...	Une recommandation forte a été exprimé en faveur de...
(altered to)	Il est vivement recommandé de...

2. <u>Mistranslation of prepositions.</u> Numerous examples were to be found here.

The substitution of the computer for the conventional calculating machine	La substitution de l'ordinateur <u>pour</u> (instead of à) la machine...

The delegation agreed to send...	La délégation a consenti <u>d'envoyer</u> (instead of <u>à envoyer)</u>

3. <u>Common idioms.</u> These come quite naturally to the experienced translator but were often rendered literally in machine translation.

Mr X gave a progress report on the project...	Mr X a donné un rapport d'avancement sur le projet...
(altered to)	Mr X rend compte des travaux réalisés dans le cadre du projet...

The Chairman welcomed the participants...	Le président a accueilli les participants
(altered to)	Le président souhaite la bienvenue aux participants...

4. <u>One translation for one word.</u> Variety of vocabulary, as required by the context, was often lacking in M.T. For example "further" was translated as "davantage" in the following phrases, requiring a number of changes.

further work	<u>la poursuite des</u> travaux
further consideration	<u>complément</u> d'examen
without further delay	<u>sans plus tarder</u>

PART II OF THE SURVEY

In the second part of the survey, which covered the experience of six translators over a relatively short period, discussions with those involved provided the following errors, again listed in order of the trouble they caused.

1. Lack of common terminology. This was a major concern. The following reactions were typical of many more.

"The machine does not use standard, well-known formulations. It does not use the correct expression for the context."

The post-editor does not seem to be able to "forgive" the machine for such ignorance. He has the impression that it only "knows" basic equivalents and a few idioms. (Note this personification of the computer.)

2. Need for extensive rewriting. Resentment was expressed at the necessity of completely rewriting entire clauses or sentences in cases where phrases were misplaced or misused. "The machine makes no effort to work out translation problems. It just identifies the various items and puts them into some order, not always the correct one, as when it translates 'the other four' by 'les autres quatre'."

3. Repetitive style. This was sometimes a source of strain on the post-editor who had to read the same standard translations time and time again and felt he had to adopt a style alien to his own. Comments here ranged from, "The post-editor has no freedom of style. His work is tedious and frustrating," to, "It is a laborious, exacting and unrewarding activity."

4. Miscellaneous problems. Other troublesome errors mentioned by those interviewed included non-translated terms, deleted sentences, spelling mistakes and incorrect inflexion of nouns, adjectives and verbs.

CONCLUSIONS DRAWN FROM THE SURVEY

On the whole, the same error types are listed in both parts of the survey but in a different order of gravity.

The fact that translators from Part II stress the lack of common terminology while there is no mention of this in Part I, is perhaps an indication of the different approaches to M.T. used. The second group appears to be annoyed that the machine cannot reproduce the kind of everyday language any translator would be expected to know. Yet experience of machine translation over a lengthy period (more than two years) provides a better idea of what one should expect and lessens the feeling one might have that the machine ought to be able to do certain things. Realization that the machine does have certain limitations convinces one that it is nothing more than a machine, and this in turn tends to reduce exasperation.

Similarly, it is interesting to note that repetitive style is placed third in Part II but only fourth in Part I. The irritation experienced on seeing the same turns of phrase again and again in cases where other equivalents could be used is typical of reactions expressed in connection with human translation as every translator is expected to make efforts to vary his style.

The post-editor appears to become most irritated when he tackles his work as if he were revising a human translation. He therefore tends to blame the machine for making numerous elementary errors requiring his correction and may even feel somewhat embarrassed at having to make so many changes. With experience this feeling slowly ebbs away and the post-editor simply makes the best possible use of the rough translation.

Irritation is, then, the basis used by the translator in evaluating M.T. He approaches his correction of raw output using the same criteria a teacher would when marking a translation examination. This is perhaps not the best or most pragmatic attitude as it is likely to lead to considerable frustration. The printout is, after all, no more than a tool to be used as an aid.

Finally, some translators are upset at the thought that M.T. could perhaps make inroads into their own individual style and competence. They therefore refuse to have anything to do with M.T., losing sight of the benefits to be gained in carrying out post-editing work.

It is, then, not only necessary to make every effort to improve the quality of machine translation, it is equally important to persuade prospective post-editors that they should adopt a completely new approach to their work. And this can only be achieved by means of appropriate training.

TRAINING OF POST-EDITORS

From the comments made, it appears that translators tend to regard the machine as a living being. They expect it to know everything the average translator knows, they find the mistakes it makes are sometimes unacceptable and often irritating, particularly when repeated, and they feel embarrassed about correcting its work.

Most post-editors are translators rather than revisors and therefore have no experience of revision. In their normal work they are expected to translate as best they can. Indeed, their work is assessed on the basis of accuracy and quality, perfection of style being one of the main criteria. For this very reason, they are not prepared to produce any work which is likely to put their reputation at stake. More specifically, they are unwilling to hand in work written in any kind of standardized style since it is their own particular style which differentiates them from others. As one translator commented, "I am not prepared to put my name to any work below my usual standard."

Quite clearly then, if translators are to be expected to take on post-editing work, they should be given appropriate training beforehand.

Potential post-editors should therefore be given an objective explanation of what the machine can and cannot (yet) do and why, perhaps including basic training in M.T. programming and the difficulties encountered. They should also be trained to adopt a more pragmatic approach, using the machine as a tool to assist them in their work.

Furthermore they should be instructed in the rules and practice of revision work and encouraged to adapt their approach to the use which is to be made of the particular translation work they are doing, adjusting their standards accordingly.

FINAL RECOMMENDATION

Style clearly seems to be the main problem in post-editing. If it is not
overcome, a distinction will have to be made between the quality of machine
translation and that of human translation and consequently between translators
who make use of M.T. and those who do without.