

Technology as Translation Strategy

AMERICAN TRANSLATORS ASSOCIATION
SCHOLARLY MONOGRAPH *Series*

Volume II 1988

EDITED BY

Muriel Vasconcellos

State University of New York at Binghamton (SUNY)

Copyright © State University of New York at Binghamton (SUNY) 1988

ISSN 0890-4111

Printed in the United States of America

The Translator and the Postediting Experience

TERRENCE McELHANEY AND MURIEL VASCONCELLOS

In most circumstances, postediting is an indispensable element of MT, and for many applications it will probably continue to be so indefinitely. The translator can use his ingenuity to devise strategies that will reduce the level of effort expended, accelerate throughput, and at the same time improve the quality of the final product. Certainly the most critical factor in the optimization of daily operations is the translator's attitude toward the postediting task, together with the insights he brings to it.

The authors have walked many miles in the Indian's moccasins. Both coming to MT from extensive experience in traditional translation, they have between them postedited more than a million words of MT output at the Wang screen using SpanamTM, the Spanish-English system developed in-house at the Pan American Health Organization. Below they summarize some of the principles that have come to light as a result of this first-hand experience.

The Role of Postediting

The purpose of MT postediting, of course, is to adjust the machine output so that it reflects as accurately as possible the meaning of the original text.

Postediting of machine translations differs from the task of the traditional reviser in several respects (see Vasconcellos 1987b). While both processes are primarily concerned with the correction of errors, the types of errors may be quite different.

Shared in common is the fact that both the traditional reviser and the posteditor must be on the lookout for a host of different things;

both are expected to make corrections that range from the mechanical to the conceptual, from the stylistic to the political.

The traditional reviser has final responsibility, at one end of the spectrum, for the detection of lapses such as missing words, skipped passages, inadvertent repetitions, misspellings, mistakes in numerals, incorrect punctuation. Yet on the other hand it is also his job to identify and correct inappropriate glosses and misconstructions of meaning. In addition, in the event the translation is to be published, he will be looking at many more aspects of the text that are of importance to his client or employing organization.

With machine translation, the posteditor has the assurance, at the mechanical level, that nothing has been skipped or repeated, and also that there are not likely to be errors in spelling—although punctuation and capitalization, to the extent that practices differ between the two languages, may need to be watched. Errors in glossing will be present in the MT output, but they will tend to be different from the kind committed by human translators: while the machine may not always find the correct alternate translation for a word or phrase, neither will it make wild guesses—for instance, mistaking *células rotas* (Spanish ‘ruptured cells’) for *células rojas* (‘red cells’—an error once found in a human translation). In other words, what it does best is look up words in the automatic dictionary, and in this task it is tireless, thorough to the end, and completely accurate to the extent that the information supplied to the dictionary in the first place was correct.

As with traditional revision, misconstructions may need to be corrected, but their scope will be narrower: the human translator may miss the point of an entire sentence and recast it so that it is no longer faithful; the machine, on the other hand, when it “misunderstands,” will make a more local mistake, usually confined to a single lexical item; it leaves the tough passages in strictly literal form. The job of the posteditor is to *do* the interpretation, not to correct the interpretation of a colleague.

The two processes are most alike in the case of a text to be published or to otherwise undergo some type of close scrutiny—when cohesive and stylistic meaning become significant factors in the level of accuracy required. But even here there is a difference. The posteditor is conscious of working from left to right and of the need to minimize time and keystrokes. He will tend to reorder and reorganize less, attempting instead to improve the text in ways that are effective and

at the same time economical. Without prejudice to his commitment to accuracy, he seeks the minimum number of steps required in order to make a text acceptable for its intended purpose. The time factor is constantly in mind. The very *raison d'être* of machines is to save time, thereby reducing effort and costs. This premise applies especially to MT, where the key to profitability both for the client and for the translator is speed-with-accuracy. The translating posteditor, if he is to do his job effectively, cannot afford to lose sight of this goal.

The Posteditor

Given the very special nature of MT postediting, there is an ongoing debate over who should be assigned to this critical task. Of course, it is possible to dispense with postediting altogether and deliver raw output directly to the end-using scientist or technician on the assumption that his knowledge of the subject matter will enable him to deal with gaps, awkward or incomplete constructions, mistranslations, etc. While this option has proven to be viable in certain circumstances, with large and highly developed MT systems, the question at hand, rather, is who is the best person to perform the postediting task when it is called for. The lure of using MT to cut costs has given rise to a number of proposals: some budget-conscious managers would enlist secretaries; others have suggested junior editors; and a special new breed of "para-translators" has been proposed by Hutchins (*infra*). In contrast, there are those who would insist on the use of professionals. The choice at this level is between bilingual scientists and technicians, who offer knowledge of the source language and a familiarity with the text type, and translators or revisers, who bring with them long experience at problem-solving in the particular language combination and awareness of the thousands of pitfalls to be avoided.

The authors believe that postediting is and will continue to be for a long time the work of the professional translator. There are strong arguments in favor of using experienced translators for this task. The translator is the one best able to pick up errors in the machine translation (e.g., misparsed or unparsable ambiguities), he has a fund of knowledge about the cross-language transfer of concepts, and he has technical resources at his disposal which he knows how to use in the event of doubts. Moreover, for the very reason that translators are

best suited to the task, the more experienced they are, the more effective they will be. An inexperienced translator—to say nothing of the non-translator—is apt to waste precious time unnecessarily reworking passages or trying to deal with a problem whose solution would be obvious to a seasoned professional.

For the particular case of MT output, there are “tricks of the trade” that can be used to streamline and enhance the postediting process. These can be picked up by a motivated professional who already has a strong foundation of experience in traditional human translation and/or, better yet, in the revision or the editing of translated texts.

In addition to the professional expertise that a translator brings to the task, it is essential that he have good keyboard skills, quick reactions to the challenges that arise in the text, and a flair for solving problems creatively.

Mode of Operation

The strategies to be adopted in postediting will depend first on the mode that is used: interactive or batch.

Interactive systems (see Weaver in this volume) require that the translator/posteditor work directly on-screen throughout the process, providing responses to prompts generated automatically as the machine progresses from sentence to sentence. The prompts are triggered at linguistic decision points, and certain choices are presented. The translator's previous experience with similar contexts is an important factor in the effective use of this mode; he must be able to work quickly in order to offset the additional time spent in front of the screen while each sentence is being processed.

With batch MT systems, on the other hand, the posteditor goes to work only after the entire job has been processed and returned to the screen as a single running text. Again, the effectiveness of the system is directly related to the translator's experience in making decisions—although the types of decisions will be different. The choices are not offered on a platter; the translator is not presented with any guidelines on how he could modify the text. He must be able to identify the trouble spots himself and know how to fix them quickly and effectively.

The observations that follow are based on the authors' experience

with Spanam™, a batch system developed at the Pan American Health Organization (PAHO) for the translation of Spanish into English.

The Strategies

Working On-Screen

It is impossible to overemphasize the importance of working directly at the video display terminal. Besides being the most effective approach from the standpoint of translator strategy, working on-screen is also vastly more efficient in terms of the level of effort and staff involved. If the changes are entered by hand on hard copy, by the time this copy is picked up from the computer, the changes penned in, the corrections entered by a word-processing assistant, and then the final version reviewed again by the posteditor, the turnaround gain of MT will have been totally lost.

In contrast, the entry of changes directly on the screen is a one-step procedure; once the posteditor has registered the change, it needs no further processing except to be printed out. Moreover, the posteditor with good keyboard skills can enter changes much more rapidly at the terminal than he can write them by hand on hard copy. He can replace terms or expressions globally or selectively. Not only is the time of making each individual correction saved, but also, as he moves along in the output he is no longer disturbed by recurring errors; the text comes to look “cleaner,” approximating more closely the final product. If necessary, he can explore various versions of a particular passage in a matter of seconds, keeping the alternatives and comparing them one against the other. Always, the effect of the changes is immediately visible in final form without any need for the text to be retyped.

Saving Time by Minimizing Corrections

The question most frequently asked is: Just how much time should be spent on postediting? The only possible answer is “as little as possible”—and this will depend on a number of factors. In any event, the time spent will decrease as the posteditor becomes familiar with the system and as the system itself is updated and expanded—some of this expansion due to feedback from the translator himself. Like-

wise, as the posteditor gains experience, he will work more quickly, and he will learn how to produce an effective translation while actually doing less editing. (Unfortunately, time is a moot question in many cases, since clients often submit texts with nearly impossible deadlines. As users become more familiar with the capabilities of MT, however, the demands on the posteditor can be expected to become more realistic.)

The main determining factor is the user's need: What is the final purpose of the translation? The answer to this question will determine the degree of postediting required and hence the turnaround time. Texts submitted for information only allow the quickest turnaround. In this case, the posteditor's task will be focused on verifying the content of the document. A sentence such as:

S: Es indispensable la capacitación y la movilización de los recursos nacionales y el desarrollo de la voluntad de los países dirigida a fortalecer las instituciones y la capacidad nacional.

MT: It is indispensable the training and the mobilization of national resources and the development of the will of the countries aimed at strengthening the institutions and the national capacity.

(actual raw Spanam output) is easily intelligible to any speaker of English and would not require editing in an information-only postedit, e.g., a translation of the interventions at a meeting for use by the rapporteur in preparing the final report by the end of the day.

Texts for publication, the other extreme, will obviously require the closest attention. Some aspects of the postediting job—research in the case of technical terminology, attention to house style, policy questions—will hinge on whether or not the text is to be further edited by the client. In any case, however, the postediting translator will probably assume linguistic responsibility, making sure that all shades of meaning are accurately reflected, that major ideas are in strongest focus and minor ones subordinated, that the necessary cohesive links are present, that the register is appropriate, that deixis and footing are consistent, that anaphoric references are clear, that the style is smooth. These extra considerations add to the time spent on the text, but as in any translation process, efficiency is gained with experience.

Linguistic Strategies—“Quick Fixes”

For publication purposes, the example above, *It is indispensable the training and the mobilization . . .*, is grammatically unacceptable, and the posteditor is faced with the decision of what to do with it. In the interest of saving time and effort, he learns how to work from left to right and to avoid, if possible, any major rearrangement of the sentence. This approach may call for changing the *function* of some of the words while still retaining their *information content*. Thus, in the example, inversion of the sentence would be time-consuming and also *unfaithful to the order of information in the original sentence* (Vasconcellos 1986), giving for this reason a result that appears heavy and “awkward”:

TT: The training and the mobilization of national resources and the development of the will of the countries aimed at strengthening the institutions and the national capacity *are indispensable*.

On the other hand, a functional rendition, or “quick-fix,” in which the nominalizations are returned to verbal form, could be handled with a minimum of keystrokes as the posteditor advances naturally from left to right (i.e., without a MOVE operation, which takes much longer) and would at the same time be a *better* translation for the fact that it preserved the information structure:

QF: It is indispensable *to* train and mobilize national resources and *to* develop the will of the countries *to* strengthen the[ir] institutions and the[ir] national capacity.

This type of solution is especially useful for translation into English from one of the Romance languages, which have a high frequency of sentences that begin with a verb. With practice, the posteditor can become adept at making corrections such as these. They challenge his creativity.

Mechanical Aids

Trivial as it may seem, much time can also be saved if the posteditor knows how to move the cursor quickly from place to place in the text. For example, the SEARCH function, combined with related strategies, is much faster than manual manipulation of the directional keys

(see Kingscott in this volume). The mouse, on systems that offer it, is also quite effective. In both cases, speed and dexterity are acquired with practice.

Feedback

During the course of his work, the posteditor should provide glosses for words that the machine did not find or that should be translated differently. Even though he may not be the one who will make the actual changes in the dictionary, it is important that he participate in the process. The translator's contribution, captured while the text is fresh in his mind, will help to improve the quality of future MT output—to his own benefit and that of his colleagues.

Conclusion

In the beginning, translators may feel somewhat awed by the array of novelties that they face in postediting: new keyboard techniques to master, new editing strategies to learn, new problems to watch for, new approaches to the process itself of translation—and all of this at once. But not surprisingly, many of them are able to adapt within a short period of time.

More than twenty translators have tried postediting at the Pan American Health Organization. Many have found that they were at home in their new milieu after about a month of full-time practice at the screen (about 100,000 words of “wearing-in”). They have reported that they are more relaxed at the end of the day, that traditional translation now seems tedious. Their average volume is around 6,000 words a day—of finished copy, ready to deliver to the requesting office. In some instances they have registered a daily output of 12,000. For those who have become involved in the system—in building the dictionaries and identifying linguistic patterns that can be codified—interest increases steadily with the passing of time.

The following strategies would appear to be at the heart of their success:

- Keeping revision to a minimum—the idea is to edit, not to rewrite;
- Working at the screen; using linguistic “quick fixes” and automated word-processing support whenever possible;
- Providing feedback to the system.

REFERENCES

- McElhaney, Terrence. "The Postediting Experience." *Jerome Quarterly* 1.4 (August-September 1986): 6-8.
- Vasconcellos, Muriel. "Functional Considerations in the Postediting of MT Output." *Computers and Translation* 1.1 (1986): 21-38.
- . "Postediting On-Screen." Paper presented at Translating and the Computer 8 (London, November 1986). To appear in the Proceedings. London: Aslib, in press. Reported in *Language Monthly* 39 (1986): 3-7.
- . "A Comparison of MT Postediting and Traditional Revision." *Proceedings of the 28th Annual Conference of the American Translators Association* (Medford, NJ: Learned Information, 1987).