AMERICAN TRANSLATORS ASSOCIATION CONFERENCE — 1987

Proceedings of the 28th Annual Conference of the American Translators Association

Albuquerque, New Mexico, October 8–11, 1987

Edited by Karl Kummer

**Published for the American Translators Association
by Learned Information, Inc.**

A COMPARISON OF MT POSTEDITING
AND TRADITIONAL REVISION

Muriel Vasconcellos, Pan American Health Organization

Abstract: Postediting of machine translations differs from the task of
traditional revision in several respects. In basic approach, revision is a
discovery process, while postediting is an ongoing exercise in adjustment.
      Both processes are primarily concerned with the correction of errors,
but the nature of error types is somewhat different, as is their distribu-
tion. With machine translation, the posteditor has the assurance that noth-
ing has been skipped or repeated. Lexical errors will be present in the MT
output, but they are more predictable than the kind committed by human
translators. As with traditional revision, misinterpretations may need to
be corrected, but they will tend to be narrower in scope. The two processes
are most alike in the case of a translation that is to be published.

1.   INTRODUCTION

      How similar is MT postediting to traditional human revision? The
answer to this often-asked question will range, depending on the aspect
considered, from "Actually, they are rather different animals" to "Very
similar indeed."
      A comparison of the two processes is timely, since machine-translated
(MT) output is being generated at a rapidly increasing pace and urgent de-
cisions await knowledge about how it should be handled. Confronted with
the new technology, managers are asking who should be assigned to the task
of making the output completely suitable for its intended purpose. Trans-
lators, in turn, are wondering what they are really expected to do--how
"far" they are supposed to go in modifying the raw product. And both
managers and translators would like to know whether the skills of tradi-
tional revision are transferrable to this new mode of work.
      While both processes are primarily concerned with the correction of
errors, there are differences in the types of errors and in the approach to
dealing with them. These differences are attributable to the nature of the
medium.

2.   TRADITIONAL REVISION

2.1   General Scope and Nature of the Task

      Literature describing and quantifying the process of traditional re-
vision is still relatively scant. Arthern (Ref. 1) has proposed a model

for assessing the quality of revision. He is concerned with rating the performance of revisers (or reviewers[1]) on the basis of the corrections they make, and he defines the process mainly by inference--i.e., by giving them a clean bill of health when they introduce a "necessary correction or improvement in readability." The need remains for a model of the process itself.

In the meantime, it seems reasonable to state that the purpose of traditional human revision is to catch and correct any errors that the first-stage translator may have made and, as required, to "polish" the text so that it is appropriate to the setting in which it is to be used.

With revision, the detection of errors is a discovery process. Typically, the reviewer starts from the assumption that the translator has made a conscientious effort to reproduce the original text in all the aspects of its meaning and that any errors are either unintentional lapses or a reflection of his lack of knowledge or experience. If the reviewer has to spend too much time correcting the translator's interpretation of the source text, throughput becomes too costly and measures have to be taken. The reviewer is expected to work twice to three times as fast as the translator, and any marked reduction in this ratio is unacceptable. Thus, by the time the translation reaches the revision stage, true substantive errors, limited naturally by these practical factors, can be expected to be rather infrequent. The reviewer sets out on an expedition to find them. The adventure is all the more exciting because they can be as creative and unpredictable as the human spirit itself.

2.2   Formal Corrections

There may be errors which, while they do not involve a judgment call or very much linguistic reasoning, nevertheless require close attention on the part of the reviewer: lapses such as missing phrases, skipped passages, inadvertent repetitions, misspellings, mistakes in numerals. The status of such errors as errors is obvious. They do not rely greatly on a fund of translation experience, but they are important to catch, and they are part of the reviewer's responsibility.

Moving up the scale of complexity, the reviewer may be expected to implement a particular "house style" in terms of format, punctuation, capitalization, and the like. Such a task, which in some operations is turned over to an editor, requires concentration and patience in ensuring that consistent standards are applied throughout the text.

House style also extends to the harmonization of usage and terminology, which is more the work of translation proper. The requirement for consistency will vary depending on the institutional setting and the purpose of the particular document.

2.3   Substantive Corrections

Most of the reviewer's attention, however, will tend to be focused on substantive corrections. At the lexical level, these are mostly replacements for glosses that are inappropriate, but the reviewer can also find clear lexical mistakes. In addition, there are misconstructions of meaning that may extend over a phrase, a clause, a sentence, or even a whole passage. It is possible for the translator to miss the point of an entire sentence and cast it from a perspective that is unfaithful to its intended meaning. Such misunderstandings can arise when the source text has problems of expression in the first place. The job of the reviewer, if the original has been misinterpreted, is to ferret out the meaning and provide a rendition that is as close as possible to the author's original intention.

2.4 Discourse Organization

Good translators and revisers pay attention to the fabric of the overall text as well as to the lexical and syntactic pieces of which it is made. They understand, both intuitively and from observation, about the ways in which this fabric is manifested (Refs. 2, 3, 4, 5, 6, 7 and others): through the factors that ensure uniformity of perspective--e.g. register; through the signposts that link the various pieces together in temporal, spatial, or logical order; through the cohesive ties that weave the threads of reference; through the strategies that make it possible to subordinate a concept, on the one hand, or, on the other, to focus on new or highlighted information; through the devices for establishing the author's viewpoint and expressing the relationship to the thought in the author's mind; and through the sets of cues which together constitute the nature of the particular genre--those features that distinguish a play from a novel from an editorial in the New York Times. All these aspects of discourse carry meaning, and in a translation for publication the reviewer is concerned with capturing the nuances in order to ensure that the translation is a text and not just a congeries of isolated phrases and sentences.

3. MT POSTEDITING

3.1 General Scope and Nature of the Task

Despite a number of differences between traditional revision and MT postediting, in the two processes the scope of responsibility is essentially the same. Both the traditional reviewer and the MT posteditor must be alert to a staggering variety of possibilities; both are expected to make corrections that range from the mechanical to the conceptual, and, for more important texts, from the stylistic to the political. In both cases, their responsibility increases as the nuances of the text become more important because of the use to which it is to be put. And as the responsibility increases, the two tasks become more similar.

With MT postediting, the focus is on adjusting the machine output so that it reflects as accurately as possible the meaning of the original text. In other words, the emphasis is on an ongoing exercise of adjusting relatively predictable difficulties, rather than on the discovery of an inadvertent lapse or error. The passages that clearly require corrections, though many of them are minor and local, are more frequent than in traditional revision.

3.2 Formal Corrections

As for the detection of ommissions and the like, the MT posteditor has an easier time of it. One can be confident that nothing has been skipped or repeated. Also, errors in numerals and spelling are unlikely, since they will occur only if there has been a mistake at the level of the machine dictionary. Punctuation and capitalization will be copied from the source language unless there is a specific instruction to the contrary in the dictionary record (for example, capitalized names of the months in English vs. lower case in Spanish). To the extent that practices differ between the two languages, these areas may need to be watched.

3.3 Lexical Corrections

In the MT output there are apt to be unfortunate lexical choices. As every translator knows, in the transfer between two languages--any two

411

languages--most of the source words will suggest more than one option in
the target. The one-to-one equivalent is the exception, and it is found
mainly in special-purpose technical text.

Before the output is generated, it is possible in an MT system to
control the selection of target glosses at the level of dictionary lookup,
especially with technical texts, by means of subject-oriented microglos-
saries. In addition, it is possible to specify idioms and rule-based col-
locations--i.e. combinations of words that tend to occur together under
certain conditions. But there will be cases that the system cannot handle;
these will come through in the raw output and the posteditor will have to
deal with them: the Spanish word núcleo will have two different transla-
tions in an atomic energy text; some of the contexts can be anticipated,
but not all of them. Or derecho in a legal text can mean both 'right' and
'law' within the same paragraph. There are many examples of this kind.

The problem of polysemous source words is compounded in everyday
language, where the wear and tear of evolution makes it more common to have
differences between the source and target languages in the range of seman-
tic fields. Whole concepts and distinctions may exist in the target but
not the source. A classic example is Spanish esperar, which can mean
either 'hope' or 'wait' in English. While the reviewer must be alert to
these possibilities, the first-stage translator is also aware of them, and
usually the right choice has been made before the translation reaches the
revision stage. In some MT systems (Systran at Wright-Patterson AFB, for
one), such cases can be dealt with using "slashed entries": the raw output
will generate, for example, the dual item hope/expect, and the posteditor
will have an expedient way of eliminating the unwanted option. Passages
with such words can make for "sleepers" in both human and machine transla-
tion. The text seems perfectly natural at first glance. For example, the
English sentence
                Time was one of the factors to be considered.

was produced by Spanam[2] in a text on Legionnaire's disease. The Spanish
word tiempo had been translated as time in the absence of sufficient context
to show that the intended meaning was 'weather'. A human translator would
probably not have made that mistake. MT, on the other hand, as yet unable
to follow complex trains of thought, is more prone to produce this kind of
problem for the posteditor.

On the other hand, some types of lexical errors will be less frequent
in the MT output. While the machine may not always find the correct alter-
nate translation for a word or phrase, it will not guess wrong. It will
not assume--as more than one human translator has--that the Spanish phrase
tuberculosis miliar is a typographical error and should be rendered in
English as military tuberculosis! What it does best is look up words in
the automatic dictionary, and in this task it is tireless and totally
thorough. It can be trusted to the extent that the information supplied to
the dictionary in the first place is correct, and that information will be
retrieved with complete accuracy.

3.4   Coping with Misconstructions

As with traditional revision, there may be misconstructions that need
to be corrected, but they will tend to be narrower in scope. While the
human translator may give the wrong slant to a whole sentence and even
longer passages, the machine, when it "misunderstands," will make a more
local mistake, usually confined to a single lexical item; the passages that
cannot be analyzed are clear in their need for intervention. The interpre-
tation itself is left to the posteditor, as opposed to the case of the re-
viewer, who is correcting the interpretation of the first-stage translator.

3.5   General Patterns and their Predictability

One of the main differences in the types of errors to be corrected is
that in machine translation a lot of the errors are predictable. There is
the inventory of polysemous source words that need to be kept in mind, or
possibly included in the dictionary as slashed entries; there are the
source-language constructions that defy analysis; and there are the con-
structions that typically give problems at the interface between the par-
ticular pair of languages (for example, VSO vs. SVO with Spanish into
English--Ref. 5) and which can yield results ranging from unacceptable to
awkward but grammatically correct--needing to be polished in some cases but
passable in others, depending on the purpose of the translation. The post-
editor becomes familiar with what to expect and develops techniques for
dealing with the recurring glosses and syntactic patterns.

If such errors and patterns can be predicted, it may be asked why
they cannot be dealt with by the computer program in the first place. The
answer is that the decisions the posteditor makes are based on his accumu-
lated extralinguistic knowledge of the subject area and other worlds, to-
gether with his innate understanding of discourse organization.

The two processes are most similar when the translation is to be pub-
lished or for some other reason is going to be subjected to close scrutiny.
The posteditor has the same level of responsibility as the traditional re-
viewer and therefore has to bear in mind all the same criteria.

Cressey (Ref. 8) has come up with some interesting statistics on
types of corrections made in the course of postediting. Based on a review
of an Engspan[3] text postedited by a professional translator, he calculated
an average of 12 changes per 100 words. The percentage breakdown was as
follows:

	%	Less than 5%:
Word order change	20	Agreement
Article added or deleted	18	Minor stylistic adjustment
New word substituted	17.5	Passive construction
Preposition changed	12	Verb problem
Major reconstruction	6	Syntactic category problem
Topic-comment reversed	5	Relative clause problem
		Number Problem
		Ungrammatical input

To illustrate "ungrammatical input" he cites the following sequence, repre-
sented in the original text as a complete sentence:

> While the tranquilizer-using population is skewed towards females
> and the elderly, both of whom do less driving than the average.

This example is typical of the fare dealt with routinely by the traditional
human translator. The difference is that the first-stage translator usu-
ally handles the problem before it is passed up to the reviewer.

The Cressey study may not be a true snapshot of a posteditor's inter-
ventions with raw MT output because some dictionary work had been done on
the text in question, which reduced the possibility of substantive errors.
On the other hand, it is also reasonable to expect that when a given sub-
ject is translated by the system on a regular basis, much of the needed
vocabulary will be already in the dictionary. For this reason, a test done
on random raw MT output would not necessarily present for the posteditor
the conditions to be expected in a real working environment.

Cressey's breakdown does, however, give a picture of the types of
corrections that a posteditor is apt to make, and the distribution appears

to be somewhat different from that of traditional revision. To begin with, the proportion of word-order changes seems to be quite high relative to the other mode. This is also true of the problems characterized by Cressey as having to do with a verb, syntactic category, relative clause, number, and agreement, which in a first-stage human translation would be rare or non-existent. The treatment of articles and prepositions, while also part of the work of human revision, is probably more frequent in postediting, since the first-stage human translator can be expected to be sensitive to prag-matic issues. Experience would suggest that the 6% for major reconstruc-tion is actually rather low for MT, and even in human revision it could be higher. For the category "new word substituted" there is no way of knowing the severity of the original problems based on the information supplied by Cressey; they could have been substantive errors or simply word choices that were improved upon. In any case, lexical substitution may be consid-ered to be one of the most frequent tasks in human revision and is probably nearly as frequent in MT postediting.

A classification of postediting corrections has also been developed by Riabsteva (Ref. 9), in a study of English-Russian postediting, and she has come up with a list very similar to Cressey's. She excludes the task of foreign-word replacement because of its "routine character" and its lack of relevance for a linguistic study of postediting. Also, she calls more attention than Cressey does to the problem of homographs being rendered as the wrong part of speech. But basically her categories, with finer discrim-inations in two cases, are the same as his.

3.6   Discourse Organization

In the typical and recommended scenario, the posteditor works directly on-screen (Refs. 6, 10). The reviewer, on the other hand, is more apt to write by hand on hard copy. Because of this difference in the working medium, the posteditor becomes conscious of a left-to-right progression which corresponds to the natural communication channel and to the delivery of information. Aware of this progression, and also conscious of the need to save time and keystrokes, the posteditor will intentionally avoid reor-dering the text, focusing rather on improvements that are effective and at the same time economical.

The ongoing adjustments that are required in postediting help to main-tain awareness of the factors of discourse.

3.7   Unnecessary and Incorrect Changes

Many translators approach MT with the feeling that the output has to be turned into a version they would have produced themselves from scratch. Extensive changes make the process too costly and negate the advantage of MT. A safe motto is "When in doubt, don't." It may take time for the post-editor to figure out ways of salvaging as much of the output as possible.

If unnecessary changes are costly, incorrect changes are dangerous. It often happens that a beginning posteditor, especially one who is unfamil-iar with the particular special-purpose technical language, will change valid output and make it wrong. This is a hazard that does not exist with human translation or revision. A posteditor prone to such "counter-corrections" will have to be revised, which neutralizes the advantage of machine translation. One way of dealing with the problem is to supervise the posteditor very closely in the beginning and to provide background reading material written originally in the target language.

4. THE MEDIUM AND THE MODE

The fact of working on-screen (Ref. 6) gives the posteditor an immediate view of the result of his changes. Replacements can be made throughout the text either globally or on a selective basis, so that it begins to take on improved shape even before it is reviewed.

If the posteditor has good keyboard skills, the corrections are entered much faster than if they were made by hand. At the same time, however, the nature of the on-screen mode, together with the usual pressure for fast turnaround (clients think: "After all, the machine did most of the work!)" makes the posteditor more conscious the effort involved in making changes. Without relenting in his commitment to accuracy, he tries at the same time to minimize the number of steps required in order to make a text acceptable for its intended purpose.

5. THE HUMAN RESOURCE

From the foregoing discussion, it can be seen that MT postediting is a highly specialized activity. As long as postediting is required (i.e. the application does not make it possible to deliver raw output directly to the end-using scientist or technician), there are strong reasons for assigning the work to an experienced translator, or better yet, a reviewer. These professionals bring with them long experience at problem-solving in the particular language combination and awareness of the myriad pitfalls to be avoided (Ref. 10).

The translator is the one best able to pick up errors in the machine translation (e.g., misparsed or unparsable ambiguities); he has extensive experience with the cross-language transfer of concepts; and he has 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.

The strategies that help to streamline the postediting process are most easily learned by a 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 this person have good keyboard skills and quick reactions to the challenges that arise in the text.

6. CONCLUSION

In summary, then, MT postediting differs from traditional revision in terms of both the overall approach and the type of errors that need to be corrected. It is an ongoing process in which the MT output is adjusted by means of small local fixes which include mainly modifications in word order, the addition or deletion of articles, the substitution of new words (especially prepositions), and minor modifications in structure,. Less locally, it can also involve the repair of larger constructions which for one reason or another are inappropriate or incorrect. And the posteditor, like the traditional translator, can be faced with the problem of ungrammatical input. The two processes are most similar when they are concerned with preparing a text for publication, in which case the devices of discourse organization should be mastered.

NOTES

¹The gentler appellation reviewer, used by the U.S. Deparment of State, some of the international organizations, and in this article, avoids the implication that the process inevitably involves changing the work of the first-stage translator.

²Spanish-into-English MT system developed in-house by the Pan American Health Organization.

³English-into-Spanish MT system developed in-house by the Pan American Health Organization.

REFERENCES

1. Arthern, Peter. "Four Eyes are Better than Two." Paper presented at Translating and the Computer 8 (London, November 1986). To appear in the Proceedings. London: Aslib, 1987 (in press). Reported in Language Monthly 39: 3-7.

2. Halliday, M.A.K., and Ruqaiya Hasan. Cohesion in English. London: Longman, 1976.

3. Halliday, M.A.K. "Notes on Transitivity and Theme in English." Journal of Linguistics 3: 37-81, 199-244, 1967, and 4: 179-215, 1968.

4. Vasconcellos, Muriel. Theme and Focus: Cross-Language Comparison via Translations from Extended Discourse. Ph.D. dissertation, Georgetown University. Washington, D.C., 1985. Abstracted in Jerome Quarterly 2 (2): 2, 1987.

5. Vasconcellos, Muriel. "Functional Considerations in the Postediting of MT Output. I. Dealing with V(S)O versus SVO." Computers and Translation 1 (1): 21-38, 1986.

6. Vasconcellos, Muriel. "Postediting On-Screen." Paper presented at Translating and the Computer 8 (London, November 1986). To appear in the Proceedings. London: Aslib, 1987 (in press). Reported in Language Monthly 39: 3-7.

7. Nida, Eugene, and Charles R. Taber. The Theory and Practice of Translation. London: E.J. Brill, 1974.

8. Cressey, William W. Review Article: "Project for Integrated Development of English-Spanish Machine Translation." Computers and Translation 2 (1): 45-53, 1987.

9. Riabtseva, Nadesda K. "Machine Translation Output and Translation Theory." Computers and Translation 2 (1): 37-43, 1987.

10. McElhaney, Terrence, and Muriel Vasconcellos. "The Translator and the Postediting Experience." In: Technology as Translation Strategy, edited by Muriel Vasconcellos. Binghamton (N.Y.): University Center (SUNY), 1987 (in press). ATA Scholarly Monograph Series, vol. 2.