MultiLingual
Language | Technology | Business
MultiLingual Computing, Inc. • 319 North First Avenue, Suite 2 • Sandpoint, Idaho 83864-1495 USA • 208-263-8178 • Fax 208-263-6310

Industry Focus

Project management and machine translation

Ana Guerberof Arenas

The use of machine translation (MT) together with human post-editing seems to be discussed in almost every training, publication or conference in the localization industry these days. In a recent survey provided by the Translation Automation User Society (TAUS), 52% of the 67 companies consulted were providing post-editing services to their customers. One important reason for this increase is the pressing need to lower costs in the localization cycle when internationalizing new products.

If the service is so broadly provided, it is logical to assume that project managers will sooner or later be confronted with a project using MT. The customer, on the other hand, will want to obtain a reasonable price discount for completing the project. It is, therefore, of crucial importance for managers to be familiar with an array of new variables. Precisely because there are so many variables to consider when dealing with this type of projects, the risk of possible issues and errors increases.

Of course, some might argue that using MT and post-editing is exactly the same as using translation memories (TMs) and reviewing fuzzy matches. While it might be true that after a certain learning curve managing this type of project becomes easier, there are many initial considerations.

Variables in MT

It is useful, albeit not essential, to find out whether the writers of the source documents or software have used specific writing guidelines or controlled language — in other words, if the source language is written in a standard way to avoid ambiguity and complex grammatical structures. If this is the case, the chances of having a decent quality output will increase, and this will be valuable information when negotiating discounts with customers or writing guidelines for post-editors. When available, managers should use these guidelines to better understand the structure of the source text and to create, together with their team, sound post-editing guidelines. If project managers work in the same company as the writing team, they might provide feedback to the writers to optimize the text for MT usage.

There are basically two types of MT engines: rule-based (RBMT) and statistical (SMT). There are also hybrids that combine both technologies. The European Association for Machine Translation website (www.eamt.org) offers related articles, and it is also a good source of information on this topic.

In general, with RBMT engines there is an initial process of dictionary creation that a linguistic team would need to implement in order to train the engine. If project managers work in a company that owns the engine, they might have to coordinate or liaise with this linguistic team at different stages of the dictionary creation/update process. If they work in a company without a proprietary engine, they might be asked to make suggestions for engine performance improvement.

With SMT, the process involves training the engine with large TMs and/or glossaries, normally from one million words upwards. This data should be as clean as possible in order to increase the quality of the output. The manager might need to provide and prepare this type of content and make sure it is up to standard. Feedback and changes are provided thereafter.

Managers need to fully understand the customer's quality expectations for the final project, as post-editing can be superficial or thorough depending on the purpose. As in general revision terms, there are different types of expected quality levels. Post-editing is typically classified as full post-editing (post-edition to human quality) or rapid post-editing (post-edition with minimal corrections for text "gisting"). Between these two options, there is a wide range of alternatives. Establishing the quality expected by the customer will help determine the price as will writing

Ana Guerberof Arenas has over 18 years' experience in the localization industry and is currently operations manager at Logoscript.

specific instructions to post-editors. If this is not done, some might correct only major errors, thinking that they are obliged to utilize the MT proposal as much as possible, while others will correct major, minor and even acceptable errors because they feel the text has to be as human as possible. The best way to discuss this issue with the customer is using the actual MT output with samples. In general terms, customers know their readers and the type of text they want to produce. If they haven't thought about it, the discussion will help them define the quality criteria to follow.

More often than not, the manager will not receive clear information about the quality of the output. Depending on who is the source of information, the quality feedback could be overly enthusiastic or extremely negative. It is rare to receive a serious analysis of the output with samples and scores. Some MT output providers might send project managers an automatic score (BLEU, Meteor, NIST or TER) that gives information on how close the output is to human quality with a single number. Unfortunately, this number might mean little in practical terms, and generally a human evaluation with scores and samples is the best possible quality assessment to establish prices and transfer information to post-editors. However, it involves time and is rarely done. My advice would be to assess the output for each language combination using different parameters — for example, grammar, terminology, format — in a randomly selected set of strings extracted from the overall content that could be classified according to segment length and where a post-editor can then classify the quality of the segment (Excellent, Good, Poor, or even from 4 to 0, or any other classification). Even though time is required for this assessment, it will give a clear idea of the productivity savings the team of post-editors might be expected to obtain during the project. If managers do not have this information, they are working pretty much in the dark in terms of prices and might be overwhelmed by the number of e-mails sent by post-editors complaining about the quality of the output with little data available to discuss the matter.

Post-editors and training

The first two questions that spring to mind are "What is a post-editor?" and "How is this job really different from translation?" In the TAUS survey, 74% of the resources doing post-editing were "regular" translators, and 26% were specially trained. The data does not explain if the 74% were somewhat "trained" beforehand in this activity nor does it reveal the final result when not using trained post-editors.

In my experience, most of the post-editors currently working in localization are "regular" translators. However, that does not mean that all the translators make efficient post-editors or that they know, initially, how to perform the activity. By an efficient post-editor, I mean a post-editor who can edit the text according to the quality standard set in the post-editing guidelines in a desirable time frame. Why are all professional translators not efficient post-editors? There are different reasons. One that is frequently mentioned is the positive versus negative attitude towards using MT. Another is how technical the translator might be. The information and training translators receive to complete the task as well as the amount of time given can also affect performance. Even if translators are extremely dedicated professionals, we cannot forget that all these factors, together with the price paid to perform the task, might influence the final quality.

In brief, the manager of a localization project using MT could potentially contact the existing pool of professional translators. However, in order to minimize the risks, complaints and e-mails, translators should receive an initial training and detailed post-editing guidelines.

Training in localization does not tend to be extensive because of obvious time constraints. Consequently, the manager should consider a kick-off training of approximately three to four hours in a live training session or an on-line presentation covering the expected quality, output quality, description of the post-editing guidelines, project-specific characteristics, and a time for comments and questions. This should help minimize the anxiety the translators might experience when working on this new task. Sometimes managers might not be involved in the training, but they should make sure that translators working in the project have received basic training and instructions before starting.

There is a pronounced learning curve when post-editing, so translators might initially show a low productivity that will rapidly increase to acceptable levels. They should be informed that this learning curve exists and that their productivity will increase as they get familiar with the type of errors and the process involved in MT projects.

It is crucial for translators to feel they are involved in the project and that their comments and suggestions are considered. This will improve their attitude towards MT, and therefore quality and productivity will increase. The manager does not need to have all possible solutions to translators' problems, but should be able to provide answers according to what can and cannot be expected in a project involving MT.

There cannot be a post-editing project without specific post-editing guidelines. These guidelines are not the same as style guides, project briefs or localization kits. The manager needs to send post-editors language-specific guidelines created for the actual post-editing task.

What should the document cover? Obviously, it is difficult to answer this question, as it will depend again on the quality of output, language combination and the usual variables. Besides, we need to consider that post-editors cannot be burdened with a whole book on post-editing, as time is of essence and their work needs to be profitable. The guidelines should be short and precise, and they should include:

- description of the type of engine
- description of the source text (type and structure of source text)
- brief description of the quality of output for that language combination
- quality expected by the customer
- scenarios when to discard a segment (post-editors should have an idea of how much time to spend in order to "recycle" a segment or discard it altogether)
- typical errors to be corrected for that language combination, including reference to tagging and links
- changes to be avoided according to the customer's expectations (for example, certain stylistic changes)
- how to deal with terminology (according to output analysis and customer's expectations — the terminology provided by MT could be perfect or it could be obsolete)

Even though time is needed to create guidelines, managers will need to invest, in turn, less time in answering e-mails or dealing with frustrated post-editors who were not told how to proceed with a given segment, error or terminology issue. Again, managers do not need to create the guidelines on their own, but they do need to make sure that they are in place before starting the project.

Productivity measurements

Productivity constitutes one of the big unknown factors in projects involving

MT and post-editing. This is partially due to the fact that using MT in localization projects is relatively new, and, therefore, standard metrics do not exist yet. This is mainly due to the amount of variables to consider, however. At any rate, we have little information on the productivity of translators' work in general. The industry uses standards – for example, 2,000 to 2,500 translated words per day – but we all know these standards are hardly applicable to all translators. Moreover, there are also agreed metrics on TM editing – percentages paid according to fuzzy match level – but most translators would agree that these percentages hardly represent the amount of work they need to perform on each proposed segment. Studies dealing with productivity when post-editing MT segments – such as Krings 2001 and O'Brien 2007 – do not show pronounced productivity increases when using MT. Frequently, however, MT developers will claim that their engine dramatically increases the translator's productivity without necessarily making their methodology available.

There is definitely uncertainty about the gains when using MT and post-editing. A figure that is normally used when discussing productivity in post-editing is 5,000 words per day, but the reality is that each project will have different productivity according to the different variables.

Nonetheless, it might be necessary, if the customer requests it, to measure productivity during the project. I would not advise doing this in a live project, as all the factors involved will distort the results. Time constraints, using TM mixed with MT segments, different files, or having translators measure their own time can be factors of distortion. However, many customers will ask for some kind of productivity estimation. If this is the case, I would recommend the following:

- The group of post-editors should be representative.
- It is better if MT post-editing is compared with human translation to measure the post-editing productivity of a translator against his or her own productivity when translating from scratch and not against an average figure.
- TM fuzzy matches should not be mixed with MT segments and new segments, as it will be impossible to know where the productivity increase or decrease is coming from.
- Different files with different fields of content will affect the result, as it would be hard to determine if errors are due to MT or to the different types of content.
- Post-editors will need specific instructions on how to carry out the measurement and specific templates to keep a record of their activity.
- Averages can be misleading. Remember, if I drink two beers and you drink zero beers, the average is one beer, but you are still very thirsty!
- When providing data to the customer, it should be explained clearly giving possible reasons for the resulting productivity.
- The quality of the post-edited text should be considered when calculating productivity. A fast post-editor might leave many errors uncorrected, and a slow post-editor might overcorrect the text.

A word of caution: the productivity data obtained in one particular project cannot be extrapolated to the productivity gain for all translators, all language combinations, or all MT projects. However, it can serve as an indication for future projects with similar characteristics.

If the project manager has planned the project correctly, the post-editors should start working on the text after a brief training. However, post-editors will most likely face challenges not reflected in the training or guidelines as it is virtually impossible to foresee all possible issues in such short periods of preparation time. Therefore, it is advisable to open communication channels specific for post-editing. This could be achieved through an on-line query management tool or through e-mail. It is of the essence, therefore, to expect queries and to have one individual assigned to solve these fairly quickly.

Towards the end of the project it is also advisable to gather general impressions on the post-editing experience in the form of a questionnaire. This is not always done, again for time and budget constraints, but I believe it is essential if dealing with the first post-editing experience. The questionnaire could cover the perceived quality of MT – excellent, good, medium and poor, for example; most frequent errors found; perceived productivity when using MT (increase, decrease or equal); and if the price matched the work carried out. The more information managers obtain from post-editors, the more data they will have to work on future projects and to negotiate with their customers.

As we have seen earlier, MT could be introduced in the cycle to save costs. It would not make much sense then to review post-edited text. However, on certain occasions revision might be necessary, either to obtain very high quality or to determine post-editors' competence. The reviewer should receive the same information as the post-editor. Most localization companies use review forms that comply with LISA, J240 standards or similar ones. However, my suggestion is to create specific post-editing categories encompassing accuracy (to what extent the post-edited version contains the same information as the source text), language and terminology. There could be a rating (from 0 to 4) or an error count in order to provide a final result. Post-editors should receive this feedback to accelerate their learning curve.

This is only a sample, and other reviewing aspects can be considered (readability, clarity, transfer, and logic) as there are multiple ways of assessing quality. It is crucial, nonetheless, to establish a relation between the speed and the quality of post-editing because, as I mentioned earlier, there is no point in having post-editors who report high productivity gains but who do not correct up to the expected level of quality.

Communication and pricing

During the lifetime of the project, it is advisable to inform the customer of the general issues encountered during the post-editing phase rather than waiting for a post-project review to give a long list of complaints from the team. For example, certain language combinations or types of documents could perform worse than initially planned. This communication will help the manager establish a good channel to discuss pricing or productivity at a later stage. It is necessary, however, to give some figures and point to possible reasons when discussing post-editing issues. It is not advisable to communicate to the customer how unhappy all translators are with the task of post-editing without providing hard data.

Prices will normally be agreed either by establishing a fuzzy match equivalent for MT or an hourly rate. According to the TAUS survey, 40% of the companies apply a fuzzy match rate, 34% an hourly rate and 26% apply "other."

The level of fuzzy match assigned to MT segments could vary between 50% to 90% of the full word-rate. That is, the cost savings could be between 50% to 10%. Sometimes even the full rate might apply if the quality of the output is poor. This is too wide to be helpful, but the truth is that the different variables do affect the price considerably and each project can be dramatically different. The hourly rate could assume a productivity of 5,000 words per day or more, or simply just pay for the hours done, which could prove to be risky when updating the budget.

If the manager works in a company that owns the engine, establishing a price is relatively easy, as she or he will have access to previous data on quality of the output and productivity. However, if this is not the case, it is advisable to assess the quality of the output and agree on the expected quality before agreeing on a definite discount with the customer. This is not always possible, and the manager will need to gamble. If this happens, being conservative will help, offering up to a 10% discount over the word rate. It is also advisable to transfer the discount to post-editors. Bear in mind that training, creating guidelines and providing feedback also constitute cost.

With a project that involves MT and post-editing, the manager faces new challenges. The success of such projects will depend on how aware the manager and indeed the whole production team is of the constraints surrounding MT and post-editing. The best advice is to obtain as much information as possible; to prepare and train the team before the project starts in order to avoid unnecessary risks; and to communicate with the customer on a regular basis to establish common expectations and feasible outcomes. Finally, it is critical to gather as much data as possible in order to prepare and learn for future projects. **M**

References

Krings, Hans. *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes.* Kent, Ohio: Kent State University Press, 2001.

O'Brien, Sharon. "Eye-tracking and Translation Memory Matches." *Perspectives: Studies in Translatology* 14.3 (January 2007): 185-205.

Logoscript

Translation & Software localization

Logoscript is a supplier of multilingual solutions to global enterprises. Our mission is to enable companies to generate value from markets around the world. For over 10 years, we have been providing translation and software localization services into European and Asian languages.

As member of LCJ, LOGOSCRIPT was awarded the Services Award for its contribution to the Microsoft's Windows Vista® release. We are also a Certified SAP translation partner and have a veteran production team experienced in using industry technology to improve the operational efficiency of our customers' globalization teams.

We invite you to contact us at +34 93 343 6380 or e-mail us at **sales@logoscript.com**
For further information, visit **http://www.logoscript.com**

© 2010 MultiLingual Computing, Inc., 319 North First Avenue, Suite 2, Sandpoint, Idaho 83864-1495 USA, 208-263-8178, www.multilingual.com