

# Deploying MT into a Localisation Workflow: Pains and Gains

Yanli Sun

Juan Liu

Yi Li

Symantec Software (Beijing) Co. Ltd

Yanli\_sun@symantec.com Linda\_liu@symantec.com Lily\_li@symantec.com

## Abstract

This user report demonstrates in detail the work involved in deploying MT into a typical enterprise localisation workflow. It analyzes the challenges that MT brings, taking the workflow of Symantec (a software company) as an example. This paper reports that the traditional workflow (i.e. a localisation process from English into Chinese) can be changed dramatically. Although localisation cost can be reduced, extra effort and skills are required from both internal content localisers and localisation vendors. When planning the deployment of an MT system into a workflow, both challenges and benefits of MT should be examined and be taken into consideration by a company.

## 1 Introduction

Machine Translation (MT) is blooming nowadays. Various free online MT engines or commercial engines are available, such as Google Translator, Microsoft Bing translator, Systran, etc. Case studies have shown that the deployment of an MT engine can be beneficial for all sides, including clients, translators and language service providers (LSP) in terms of cost and time saving (Ramos 2010). Localisation is the most common domain into which deployment of MT is reported (Karamanis, Luz and Doherty 2010).

Besides professional LSP or translation service companies, many large technology or software companies are also either developing their own MT system or licensing commercial MT engines for their own localisation purpose. For example, Microsoft has been using its internally developed MT engine for internal needs such as Knowledge-Base translation since 2002, and has made the MT

engine publically available since 2007 (Wendt 2010). Starting from 2003, Symantec has actively researched MT improvement (Roturier 2009). At the moment, the well-established Rule-Based MT (RBMT) engine – Systran is deployed in the production workflow of Symantec.

This paper talks about the localisation workflow for Chinese in the Beijing Branch of Symantec after MT was introduced by headquarters. It starts with an introduction of the challenges posed to internal linguists. Post-editing training for language service vendors is explained as well as research on MT refinement. The discount negotiated with vendors is also reported. Unlike the work of Ramos (2010) and many others, this paper is written not from the perspective of an LSP but from non-expert enterprise MT users. In short, this paper exhibits the whole process of adopting an MT engine into an enterprise workflow.

The remainder of the paper is structured as follows: Section 2 compares the translation workflow before and after MT is introduced. This section details the process of deploying MT, especially the preparation work involved. Section 3 talks about post-editing training for vendors organized by internal linguists. Training guidelines are listed together with the background of the post-editors. Section 4 shows some post-translation work done within the company in order to improve the performance of the MT engine. Finally, Section 5 concludes this paper by summarizing the major gains and losses.

## 2 Preparation Work

A lot of work is involved to include an MT engine into the production workflow. According to Roturier (2009), the following steps were taken when Symantec decided to deploy an MT into the localisation workflow:

- 1) Identify the right team: to use an MT technology effectively, several team members were involved to investigate, setup and maintain the MT engine.
- 2) Identify the right workflow: the workflow we used for deployment relies on a combination of MT and Translation Memory (TM) technologies.
- 3) Identify the right evaluation metric and tool: four evaluation criteria are set up to evaluate MT outputs internally, namely, Excellent MT Output, Good MT output, Medium MT output and Poor MT output.
- 4) Identify the right source content: domain-specific User Dictionaries (UD) were well-maintained. In addition, Controlled Language (CL) rules were employed during authoring stage. We will introduce more detail about CL in the following sections.

According to the content service team in Beijing branch, a documentation localisation workflow is as follows: a content provider (the client) raises a request for translation; content is forwarded to an LSP to be translated; translations are sent back for quality review and control by internal linguists before being returned to the client.

Once the MT engine is in use within Symantec, the new MT and TM integrated workflow differs from the traditional one in several ways: content will be pre-translated first by the MT engine; then this pre-translated content is sent to LSP to be post-edited to obtain final translation; translation will later be sent back to the client for quality review.

As can be seen, there are two major differences between the two types of workflow. First of all, a pre-translate step is added. Second, the LSP translation process is altered.

By adding one pre-translate step, the focus of internal linguists shifts from solely reviewing the final translation quality to managing how the translation is done as a whole. As a starting point,

source control was found to be a pre-requisite to MT success (Roturier 2009). If content or source material is not yet created, a set of Controlled Language (CL) rules (Huijsen 1998, O'Brien 2003) created and implemented at the authoring phase would help render better MT. CL is defined as “an explicitly defined restriction of a natural language that specifies constraints on lexicon, grammar, and style” (Huijsen 1998: 2). The mechanism of CL is to minimise ambiguities from the source instead of correcting errors after translation. O'Brien (2006) provided empirical evidence that controlling the input to an MT system could lead to faster post-editing speed indicating improvement in MT output. Specific rules for Symantec documents authoring include: “Avoid use of passive voice”; “The length of a sentence should not exceed 25 words.”, etc.

If the source text has been created, checking the source against the rules can help eliminate some source ambiguities. In order to ensure that documents are written in a way that conforms to the rules specified, CL checkers have been developed. An example of such a checker is acrolinx's acrolinx IQ<sup>1</sup> which was also used by Symantec. In order to improve the ease of localising a document for a global market, internal linguists based in various area were consulted during rule-construction and source control.

Another preparation step which relies heavily on internal linguists is the construction of a domain-specific dictionary, especially for an RBMT engine such as Systran. Such a dictionary contains entries or terms that are unique to the content of a client company with its preferred target translation equivalents. In this way, translation consistency for these specific terms is ensured. However, like constructing authoring rules, building user dictionaries is both time and cost consuming, requiring much work from linguists. Symantec's Beijing branch is responsible for the construction and maintenance of an English to Chinese dictionary. Our internal linguists pointed out that to build a dictionary usually needs four steps: term

---

<sup>1</sup> acrolinx:  
[http://www.acrolinx.com/why\\_acrolinx\\_iq\\_en.html](http://www.acrolinx.com/why_acrolinx_iq_en.html)  
[last visited 2011-08-17]

harvesting, term filtering, term translation and term encoding.

Another change in the workflow with an integrated MT engine is that an LSP will post-edit the translations rather than translate directly. Depending on the purpose of the translation task or the quality of an MT engine, post-editing is divided into full post-editing (PE) (i.e. complete and detailed corrections leading to publishable quality output) and light post-editing (essential corrections only) (O'Brien, Roturier and Almeida 2009). In some cases, translators may even need to re-translate from scratch instead of post-editing an MT output. Since not all translators have experience dealing with MT output, training is necessary. To achieve this goal, Symantec conducted several training sessions to help their Chinese vendors understand post-editing. The next section explains the training guidelines in detail.

### 3 Post-editing Training

To obtain high-quality translation in a localisation context, human post-editing is today generally considered as a necessary step. As mentioned previously, when dealing with MT output, there are different levels of post-editing. To distinguish the level of post-editing required is necessary because efforts indicate different amount of work, and ultimately cost differences.

Up until this point, there have been no universal rules for post-editors training. One set of guidelines that readers can now refer to is the one posted by TAUS and CNGL<sup>2</sup>. In a workshop organized by O'Brien, Roturier and Almeida (2009), general guidelines of post-editing were summarized, such as: "Retain as much raw translation as possible"; "Don't hesitate too long over a problem"; "Don't worry if style is repetitive", etc. (2009:12). However, they also mentioned that different companies may use different rules to meet their own needs.

Based on the rules listed in the literature, together with the experience of internal linguists, in

---

<sup>2</sup> TAUS:

<http://www.translationautomation.com/machine-translation-post-editing-guidelines.html>

Symantec's pilot training course, the following training guidelines (or requirements to be more precise) were employed:

- The post-edited output quality should be the same as the traditional manual translation process.
- If the MT translation is acceptable, do minor changes to make it perfect.
- If the MT translation has some problems but could be understood, please do necessary editing to make it acceptable.
- If the MT translation is totally unacceptable, please just use the key terminology's translation and re-translate the sentence from scratch.
- Absolutely spend no time on evaluating MT output.
- If you are fairly sure that the terminology in the MT output is incorrect, do not spend too much time researching this. Ensure that you record these terminology inaccuracies for your language lead.
- Be careful not to post-edit word order in a sentence that does not violate semantic intelligibility rules. In other words, if the word order is correct but you would say it slightly differently, there is no need to post-edit.
- Avoid replacing a word with a synonym if the original word is correct. However, ensure that all forbidden words are edited in any MT output.

As for the background information of the post-editors who attended the training, on average, they have or are:

- Minimum 2 years' translation experience
- Minimum 1 year's editing experience
- 3800 words output per day;
- High language skills on both comprehension and expression
- Flexible with new process and technology
- Adapt to new environment quickly

Again, though there are no fixed rules with regard to the background of standard post-editors, some general skill sets that are required for post-editor are: "excellent knowledge of source language"; "perfect command of target language"; "word-processing skills", etc. (O'Brien, Roturier and

Almeida (2009: 29). Other necessary skills required for post-editor may include (as summarised by O'Brien (2002)): knowledge of MT, term management skills, term management (exchange formats, tools etc), basic programming skills (e.g. macros for automated text correction), etc.

We monitored the level of post-editing that the translators conducted, in order to ensure that the post-edited output quality was the same as the traditional manual translation. The pilot test also sought to inform the translators that neither over post-editing nor under post-editing were a desirable approach.

We selected a text sample with 28554 words and first translated it using the in-house MT engine. The MT output was then delivered to two professional translators to be post-edited. To measure the post-editing effort of the translators, we compared the original MT output with the post-edited version automatically using GTM (Turian, et al. 2003). General speaking, GTM is an automatic evaluation metric which calculates the similarities between an MT output and its standard translation and produces scores ranging from 0 to 1 indicating that the two translations are totally different to almost identical. Usually, the higher the score, the better an MT translation.

However, in the test, GTM was used as an indicator of the post-editing effort of a translator. If the GTM score is very high, that means the original MT output and the post-edited output are very similar. In other word, little changes were made to the original MT output; if the GTM score is very low, that mean a lot of edits are made to the MT output. Which one is better was then judged by the internal quality reviewers. It was concluded from the test that if GTM scores of a post-edited translation (compared to its original MT output) are in the range of 0.55 to 0.58, the quality of the post-edited translation is the best for the testing text.

Post-editing training and quality reviewing is not the last step for internal linguists; another new task related to MT is to log the repetitive errors and to propose methods to improve the MT engine.

## 4 Quality Review and MT Refinement

Once the post-editing task is done by the vendors, the next step is to log the errors found in the MT output. The linguists will not only focus on examining the quality of the final translation but also talk with translators and track the errors of the MT engine. Once the errors are logged, then various approaches could be proposed or taken to improve the overall quality of the MT engine in order to obtain further cost-savings.

In the same pilot test mentioned in Section 3, during their post-editing process, the vendors were also asked to record MT errors. The errors could be summarized as follows:

- 1) Errors in translations of domain-specific word/term. They need to be put into User Dictionary or be put into a Do-Not-Translate list.
- 2) Errors in word order. Translations of clauses are incorrect for most of cases.
- 3) Errors in tag handling.
- 4) Errors in prepositions.

Researchers in Symantec have been endeavouring to reduce the number of errors and to obtain better MT output. Besides working closely with the MT engine provider, new or novel approaches keep being proposed and tested by Symantec researchers for different language pairs. For example, Roturier introduced the automated and statistical post-editing tests within Symantec in 2009. Aranberri (2009) reported her work on improving the translation of English to French, German, Spanish and Japanese. Sun (2011) focused her research on improving English to Chinese translation, translation of prepositions in particular.

## 5 Conclusion

So far, this paper has presented the full process of deploying an MT engine into a localisation workflow. After post-editing training and testing, vendors agreed to provided 20%-50% discount on future translation tasks.

The focus of this paper, however, is on reporting the challenges or "pains" introduced by an MT engine. Adding an MT element into the workflow requires additional internal resources. An internal

linguist reports: “for MT projects, we need to do more work on preparation such as source checking, term harvesting, term encoding, analyzing, etc.” Moreover, post-editing level has to be determined and training is required.

In summary, if any company wants to deploy an MT engine into its enterprise workflow, both the gains and cost should be taken into consideration. With fast development in this research area, we can also expect that MT will be more useful in the future.

## Reference

- Aranberri, M. Nora. 2009. -ing Words in RBMT: Multilingual Evaluation and Exploration of Pre- and Post-processing Solutions. PhD thesis. Dublin City University.
- Huijsen, W.O. 1998. Controlled language – an introduction. In *Proceedings of the 2nd Controlled Language Applications Workshop (CLAW1998)*, 21-22 May, Pittsburgh, Pennsylvania, pp.1-15.
- Karamanis, Nikiforos, Luz, Saturnino and Doherty, Gavin. 2010. Translation practice in the workplace and machine translation. In *Proceedings of the 14th Annual Conference of the European Association for Machine Translation*, 27-28 May 2010, Saint-Raphaël, France, 8pp.
- O'Brien, S. 2006. Machine-Translatability and Post-Editing Effort: an Empirical Study Using Translog and Choice Network Analysis. PhD thesis. Dublin City University.
- O'Brien, Sharon, Roturier, Johann and Almeida, Giselle. 2009. Post-editing MT output: views from the researcher, trainer, publisher and practitioner. [Tutorial at] In *Proceedings of the Twelfth Machine Translation Summit*, August 26, 2009, Ottawa, Ontario, Canada, 77pp.
- O'Brien, Sharon. 2002. Teaching post-editing: a proposal for course content. In *Proceedings of Sixth EAMT Workshop "Teaching Machine Translation"*, November 14-15, 2002, UMIST, Manchester, England, pp.99-106.
- O'Brien, Sharon. 2003. Controlling controlled English: an analysis of several controlled language rules sets. In *Proceedings of EAMT-CLAW-03*, Dublin city university, Dublin, Ireland, pp.105-114.
- Ramos, C. Luciana. 2010. Post-editing free machine translation: from a language vendor's perspective. In *Proceedings of the Ninth Conference of the Association for Machine Translation in the Americas*, October 31 – November 4, 2010, Denver, Colorado, 5pp.
- Roturier, Johann. 2009. Deploying novel MT technology to raise the bar for quality: A review of key advantages and challenges. In *Proceedings of the 12th Machine Translation Summit*, 26-30 August, Ottawa, Canada, pp.1-8.
- Sun, Yanli. 2011. An Investigation into Automatic Translation of Prepositions in IT Technical Documentation from English to Chinese. PhD thesis. Dublin City University.
- Wendt, Chris. 2010. Better translations with user collaboration – integrated MT at Microsoft. In *Proceedings of the Ninth Conference of the Association for Machine Translation in the Americas*, October 31 – November 4, 2010, Denver, Colorado, 5pp.
- Turian, P.Joseph, Shen, Luke and Melamed, I.Dan. 2003. Evaluation of machine translation and its evaluation. In *Proceedings of the MT Summit IX*, 23-27 September, New Orleans, USA, pp.386-393.