

February, 2016

# A brief introduction to writing a post-editing guide

Uwe Muegge



This work is licensed under a [Creative Commons CC BY International License](https://creativecommons.org/licenses/by/4.0/).

# A brief introduction to writing a post-editing guide

As post-editing of machine translation (PEMT) is no less trivial than human translation from scratch, it is generally a good idea to provide linguists engaged in this activity with guidance. Here are some practical tips to help you create your own post-editing guide.

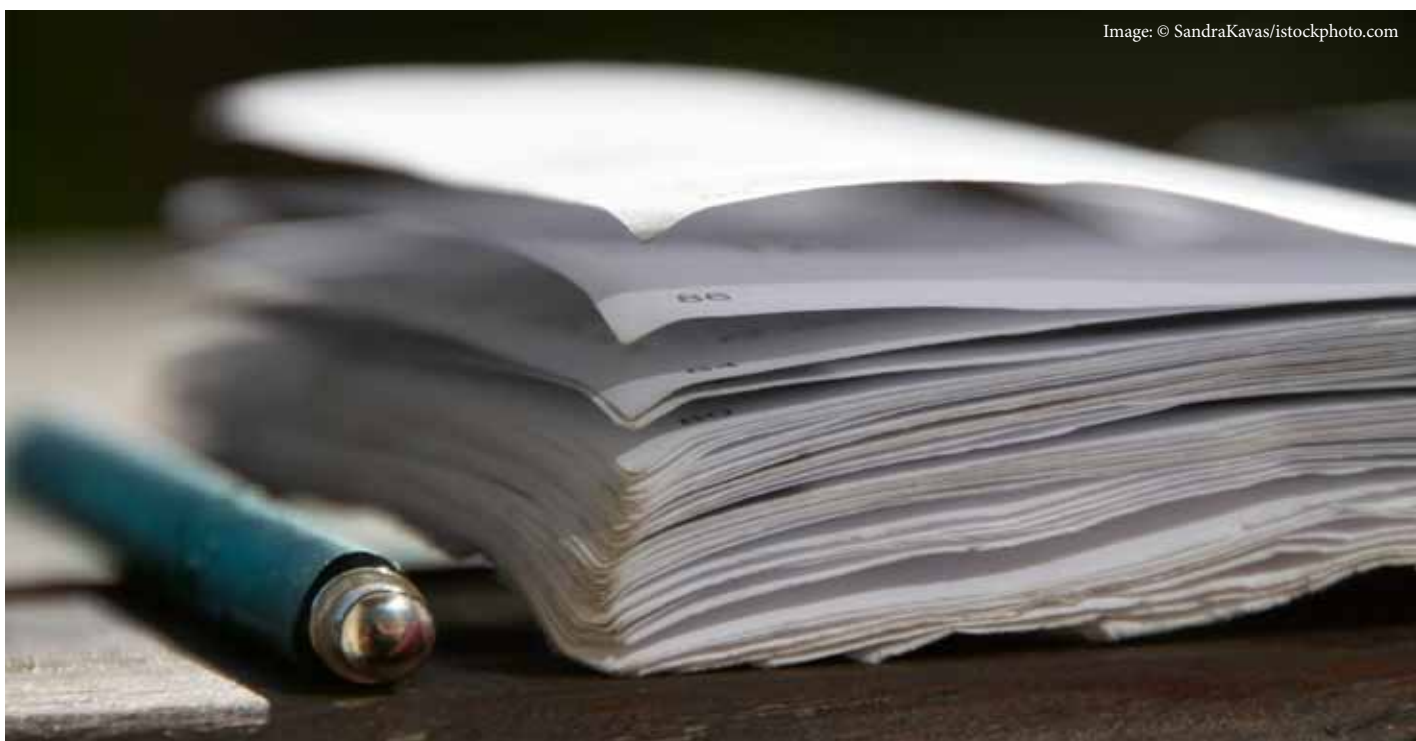


Image: © SandraKavas/istockphoto.com

By Uwe Muegge

A post-editing guide is particularly helpful if a consistently high level of output quality is required, and/or if multiple and/or different linguists are working on a post-editing project in the same language for the same client. Of course, different clients have different quality requirements, and different MT systems/engines produce different outputs. All of the above makes writing a post-editing guide a difficult task. While every PEMT environment is unique, here are a few general tips for creating an effective post-editing guide quickly and efficiently.

## Don't reinvent the wheel

I have some good news for anyone interested in writing a guide for post-editors: You do not have to start from scratch – and in fact, you shouldn't! Here are two easily available guideline documents that will help you jump-start your own PEMT guide project.

### TAUS MT Post-editing Guidelines

In August 2010, the Translation Automation User Society (TAUS) devoted one of their round-table events

to post-editing of machine translation. Over the course of several months, the conversation among round-table participants, as well as members of academic, business and government communities engaged in post-editing resulted in what is arguably the first publicly available guide on post-editing best practices. The *TAUS MT Post-editing Guidelines* actually provide advice on both pre- and post-editing, and they include clear and concise recommendations for light post-editing and full post-editing. It deserves mentioning that this guide is one of more



www.sdl.com

# Market Leading Software for Translation Professionals

SDL Language Solutions offer a unique language technology platform

From translation memory productivity tools for the individual translator to project management software for translator teams, from translation management solutions to cloud-based machine translation for enterprises and LSPs.

You are not just investing in a market-leading translation productivity tool when you buy SDL Trados Studio, you are investing in a CAT tool that integrates with the full SDL language technology platform including the new innovative Language Cloud.

Find out more: [www.translationzone.com](http://www.translationzone.com)



Follow us on twitter @sdltrados, YouTube "SDL Trados" or LinkedIn "SDL Trados Group".

## Definitions

### Machine translation (MT)

An automated translation from one natural language into another that is created by a computer application.

### Post-edit (PEMT)

Edit and correct machine translation output.

### Translate

Render source language content into target language content in written form.

### Translation memory system

A type of translation software that supports linguists by automating many translation-related tasks, including making previous translations easily available for re-use.

than a dozen best-practice guides related to machine translation that TAUS makes available to anyone for free at <http://ow.ly/Xt5MB>.

### ISO/DIS 18587 Translation services – Post-editing of machine translation output – Requirements

ISO 18587 has been under development for several years, and as the "DIS" in its title indicates, this standard is currently in draft status where public input is sought and significant changes are still possible. The title also states that this is a requirements (as opposed to a guidelines) standard. Consequently, the 13 pages of this standard are full of definitions, specifications and requirements: from qualifications and competencies of post-editors to quality requirements to activities before and after post-editing production. While ISO 18587 does not cover every aspect of post-editing – e.g. it focuses mainly on full post-editing, covering light post-editing only in a two-paragraph section of the annex – this standard is as comprehensive as it is concise.

## Specify the quality level

Just like in traditional human translation, there are no generally accepted quality standards for post-edited machine translation output. Therefore, it is of critical importance that any post-editing guide clearly specifies the characteristics of the desired quality level(s) for post-edited machine translation. Without such guidance, the risk is high that the post-editing project fails: Post-editors might either edit too little and deliver poor quality or edit too much, which might cause schedule delays and budget overruns. Here are three possible quality levels for post-editing of machine translation.

### Light post-editing

In my experience, if you ask three post-editors what the term "light post-editing" means, you get four different opinions. Experts seem to agree on only one feature of light post-editing: In a light post-editing project, post-editors should try to preserve as much of the original raw machine translation as possible. But does light post-editing include researching specialized terminology or is it okay to use "generic" terms? And do post-editors in a light post-editing project check for meaning errors? In fact, does light post-editing even involve checking the machine translation against the source? A useful post-editing guide should address all of these questions – especially the last one.

### Full post-editing

There is typically less controversy about the level of quality expected in the final deliverable of a full post-editing project. The general consensus seems to be that after full post-editing, at a minimum, a machine translation should be:

- Complete (i.e. no omissions or additions);
- Free of errors in meaning and terminology;
- Free of grammar, spelling and punctuation errors.

One area that is still hotly debated is the matter of style: Does full post-editing require stylistic changes to the point that make a machine translation indistinguishable from a human translation? Clarifying this particular point will add a lot of value to any post-editing guide.

## Transcreation

The term “transcreation” refers to the process of adapting content to a local market to a much higher degree than a typical (human) translation does. Therefore, it may seem counterintuitive to post-edit a machine translation as part of a transcreation project. The primary objective of PEMT typically is to make the translation process more efficient. Transcreation, on the other hand, with its focus on producing the most compelling messaging in a foreign language, is a quality-driven process. Does it make sense to post-edit a machine translation to transcreation level? One of the clients I am currently working for answers this question in the affirmative. To ensure the success of post-editing projects at the transcreation level, the following resources should be available:

- Post-editing guide with comprehensive specifications regarding tone and voice of messaging;
- Linguists with deep subject-matter (brand and product/service) expertise.

PEMT at the transcreation level may involve a second round of post-editing, as is the case in my client’s project, but it doesn’t have to.

## Specify the tools

Unfortunately, one aspect that is absent from most post-editing guides that I have seen, including ISO 18587 and the TAUS PEMT guide, is guidance on the tools environment for post-editing machine translations. It is a lesson that I have learned the hard way: Failing to specify what types of translation tools to use in a post-editing project

typically leads, at the very least, to less-than-optimal linguist productivity and output quality.

## Translation memory system

Many linguists seem to prefer a word processor over a translation memory system for post-editing. I have explained elsewhere that translation memory systems have many benefits over word processors for translation (see Resources). And the same advantages apply to post-editing. Here are a few examples of the benefits of post-editing in a translation memory system:

- Simplified editing (source and target are in sync and in close proximity);
- Simplified completeness check (it is hard to accidentally skip a sentence);
- Automatic terminology lookup (more on that below).

For these reasons, I strongly recommend including a section in your post-editing guide that specifies that, as a general rule, linguists use a translation memory system for post-editing projects.

## Terminology management system

As in human translation, handling terminology efficiently can be a challenge in post-editing of machine translation. Of course, the most efficient way of managing terminology in PEMT is to get the terminology right during the MT stage of the project. So how do you ensure that the post-editors use the same (and ideally correct) terms in an efficient manner throughout the project? By creating a project-specific dictionary at the beginning of the project, and by using a terminology management system that is part of a translation memory system. If a comprehensive, project-specific termbase is available in a built-in terminology management module, the translation memory system automatically

recognizes these terms. In this type of post-editing environment, linguists not only use terms more consistently, their productivity can also be much higher (can’t beat automation when it comes to efficiency)!

## Automatic quality checker

When post-editing a machine translation in a translation memory system, the system links each source sentence to the corresponding target sentence. Of course, when post-editing in a word processor, there is no such connection. By working with a repository of linked source and target sentences, post-editors using a translation memory system can conduct automatic quality checks. Here are a few examples of the types of automatic quality checks post-editors can perform:

- Consistency of terminology;
- Integrity of numbers;
- Integrity of formatting and tags.

Today, many translation memory systems offer some kind of automatic quality checking as a standard feature. In addition to built-in quality

checking, there are a number of specialized tools such as ApSIC Xbench and D.O.G. ErrorSpy that can be used for this task.

## ABOUT THE AUTHOR

### Uwe Muegge

is the Director of Solution Architecture at Z-Axis, a global provider of consulting, localization, and IT services, where he is responsible for controlled authoring, machine translation, post-editing and terminology management. Uwe Muegge has more than 15 years of experience in translation and localization, having worked on both the client and vendor side of the industry, as well as in localization training and standardization.



- @ info@zaxistech.com
- www.zaxistech.com
- @UweMuegge

## Resources

- ISO. 2015. *ISO/DIS 18587 Translation services – Post-editing of machine translation output – Requirements*. International Organization for Standardization. [Online] 5 19, 2015. [www.iso.org/iso/catalogue\\_detail.htm?csnumber=62970](http://www.iso.org/iso/catalogue_detail.htm?csnumber=62970).
- Muegge, Uwe. 2015. *Do translation standards encourage effective terminology management?* Revista Tradumàtica. 1 2015, Vol. 13
- Muegge, Uwe. 2010. *Ten good reasons for using a translation memory*. tcworld. [Online] 2010. [www.tcworld.info/e-magazine/translation-and-localization/article/ten-good-reasons-for-using-a-translation-memory](http://www.tcworld.info/e-magazine/translation-and-localization/article/ten-good-reasons-for-using-a-translation-memory)
- Muegge, Uwe. 2014. *The great leap forward that never was*. The ATA Chronicle. [Online] 2014. [http://works.bepress.com/uwe\\_muegge/92](http://works.bepress.com/uwe_muegge/92)
- TAUS. 2010. *MT Post-editing Guidelines*. TAUS website. [Online] 12 30, 2010. [www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines](http://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines).