

A Comparative Study of Post-editing Guidelines

Ke HU, Patrick CADWELL

ADAPT Centre, Dublin City University, Glasnevin, Dublin, Ireland

Ke.hu2@mail.dcu.ie, Patrick.cadwell12@mail.dcu.ie

Abstract: With the popular use of machine translation technology in the translation industry, post-editing has been widely adopted with the aim of improving target text quality. Every post-editing project needs to have specific guidelines for translators to comply with, since the guidelines may help clients and LSPs to set clear expectations, and save time and effort for translators. Different organizations make their own rules according to their needs. In this paper, we focus on comparing five sources of post-editing guidelines, and point out their overlaps and differences.

Keywords: translation, light post-editing, full post-editing, post-editing guidelines

1. Introduction

Post-editing has been increasingly researched and implemented by Language Service Providers (LSPs) in recent years as a result of the productivity gains it can bring to translators (Guerberof, 2009; Federico et al., 2012; WEB, a). However, it has been noted that there are no widely accepted general or standard post-editing (PE) guidelines (DePalma, 2013; TAUS, 2016). Since needs vary, it seems that guidelines will never be general or standard. Therefore, this paper is not going to set a general standard to post-editing guidelines (hereafter abbreviated as PE guidelines), but select, review and compare different PE guidelines which are representative (one set of guidelines produced by a resource centre for the translation industry, one by a LSP, and three by scholars). The research mainly focuses on the comparison of five proposals (O'Brien, 2010; Mesa-Lao, 2013; Flanagan and Christensen, 2014; Densmer, 2014; TAUS, 2016).

Since most organizations prefer to keep their PE guidelines for internal use only, we just have access to the ones that have been published, which are not many. Among them, we select the five proposals above as our focus because they have been published recently, are relatively complete and are proposed in terms of two categories: light (rapid or fast) post-editing and full (or heavy) post-editing. For the convenience of comparison, the five selected sets of PE guidelines are general rather than language dependent or aiming at specific contents.

2. Different Levels of Post-editing

According to ISO 17100:2015, post-editing means to “edit and correct machine translation output (ISO, 2015)”. Allen (2003) pointed out the distinction between different levels of post-editing. He first explained the determinant factors of the post-editing level and proposed using inbound and outbound translation to categorize the

types and levels of post-editing. For the inbound one, there are two levels: MT with no post-editing (for browsing or gisting), and rapid post-editing. For the outbound one, which means the translation is for publication or wide dissemination, the three levels are MT with no post-editing, minimal post-editing and full post-editing. Apart from rapid and full post-editing, the two popular categories, the intermediate category of minimal post-editing was qualified as “fuzzy and wide-ranging (Allen, 2003:304)”. He then provided a number of case studies on post-editing as well as the PE guidelines of the European Commission Translation Service (ECTS), some of which were written by Wagner (1985). Wagner’s guidelines are general and apply to projects with severe time constraints. Her PE guidelines have been mentioned in the research of O’Brien (2010) and Mesa-Lao (2013). Belam (2003) proposed her “do’s and don’ts” PE guidelines under the categories of rapid and minimal post-editing.

Rather than differentiating between guidelines for light and full post-editing, the Translation Automation User Society (TAUS) differentiated between two levels of expected quality, including “good enough” quality, and “human translation quality” (TAUS, 2016). However in this paper, for comparison purposes, we will still regard them as light and full PE guidelines, which are the two most popular post-editing levels.

3. Definitions of Light and Full Post-editing

It can be seen clearly that most people or organizations dealing with translation have very similar views about the two levels of post-editing. For light post-editing, it usually means the quality is good enough or understandable, while for full post-editing, “human-like” is usually the key word. According to TAUS (2016), full post-editing should reach quality similar to “high-quality human translation and revision” or “publishable quality”, while light post-editing should reach a lower quality, often referred to as “good enough” or “fit for purpose”. As DePalma (2013), founder of Common Sense Advisory, put it:

“Light post-editing converts raw MT output into understandable and usable, but not linguistically or stylistically perfect, text... A reader can usually determine that the text was machine-translated and touched up by a human... Full post-editing, on the other hand, is meant to produce human-quality output. The goal is to produce stylistically appropriate, linguistically correct output that is indistinguishable from what a good human translator can produce.” (DePalma, 2013, Online)

Iconic, a MT company based in Dublin, categorizes light and full post-editing by answering three questions: what, when and result (WEB, a). It suggests that light post-editing is for internal dissemination while full post-editing is for wide dissemination or certified documentation.

4. Comparative Studies of PE Guidelines

TAUS established PE guidelines in partnership with CNGL (Centre for Next Generation Localization) in 2010 with the hope that organizations could use the guidelines as a baseline and tailor them for their own purposes as they required. This is the first attempt at publicly available industry-focused PE guidelines. The guidelines start with some recommendations on reducing the level of post-editing required. TAUS highlighted two main criteria that determined the effort involved in post-editing: the quality of the MT raw output and the expected end quality of the content. They then proposed the

guidelines according to the different levels of expected quality. Flanagan and Christensen (2014) carried out a research project and tested the TAUS PE guidelines (2010) among translation trainees. Based on the result, they developed their own set of PE guidelines for use in class. They adopted the TAUS guidelines for light post-editing and proposed their tailored guidelines for full post-editing according to the TAUS baseline for translator training purposes. Recently in 2016, TAUS updated their PE guidelines to include a greater amount of detail than the previous set. The updated guidelines have been divided into five parts. In addition to an updated version of the previous guidelines, which constitutes its second part, the other four parts are as follows: evaluating post-editor performance, post-editing productivity, pricing machine translation PE guidelines and about the MT guidelines. For the purposes of this paper, we will only discuss the second part that elaborates on the PE guidelines of different levels. This part is almost a copy of the previous guidelines, but there is one specific difference in that it says “human translation quality” in the caption for the high level post-editing (although it still uses “quality similar or equal to human translation” in the body of the text).

At the 2010 AMTA conference, O’Brien presented a tutorial on post-editing. She first introduced the general PE guidelines of Wagner (1985), then the guidelines on light and full post-editing respectively. Mesa-Lao (2013) restated O’Brien’s general PE guidelines in his study. He reported his suggestions on how to decide whether a MT output should be recycled in post-editing or not. He also mentioned the rules of Microsoft (the “5-10 second evaluation” rule and the “high 5 and low 5” rule) on making these decisions in his research.

Although LSPs possess their own tailored PE guidelines, very few have been released online. Lee Densmer, senior manager at Moravia, wrote down her PE guidelines in her blog at the website of Moravia. The guidelines may be her personal opinion but can represent the attitudes of Moravia to some extent. Similarly to Allen (2003), Densmer (2014) listed the determinant factors of post-editing levels. They both believed that the client and the expectation to the level of quality played important roles. Based on their date of publication, we could argue that determinant factors listed by Densmer are more related to modern technology. Let us take TM as an example. While the factors listed by Allen are more traditional, including the time of translation, the life expectancy and perishability of the information, Densmer pointed out that the key phrases for light post-editing were “factual correctness” and “good enough”, which are in line with TAUS. She argued that light post-editing was not an easy job for linguists, due to the fact that linguists had to try their best to turn a blind eye to those ‘minor’ errors. With reference to full post-editing, she indicated that “the effort to achieve human level quality from MT output may exceed the effort to have it translated by a linguist in the first place (Densmer, 2014)”, and Iconic (WEB, a) supports this assertion. In the end, she exposed the “shades of grey” which referred to the fact that many clients want the quality of full post-editing with the price and speed of light post-editing.

Inspired by the categories used in the LISA QA Model (Localization Industry Standards Association Quality Assurance Model) and SAE (Society of Automotive Engineers) J2450 translation quality metric, we created Tables 1 and 2 as follows to compare the five proposals of PE guidelines. According to the variables in the left column, we listed all the corresponding requirements of the five proposals. There are some differences in terminology used by authors on PE, but these terms appear to refer to roughly the same concept, such as “accurate” and “correct”. If the guidelines did not mention the variable, the cell was left blank.

Table 1. Comparative study of light PE guidelines

LIGHT POST-EDITING	TAUS (2016) (FIANAGAN & CHRISTENSEN, 2014)	O'BRIEN (2010)	MESA-LAO (2013)	DENSMER (2014)
Accuracy	TT communicates the same meaning as ST	Important	Important	Factually accurate
Terminology		No need to research	No need to spend too much time researching if incorrect	Be consistent
Grammar	May not be perfect	Not a big concern	No need to correct unless the information has not been fully delivered	Correct only the most obvious errors
Semantics	Correct			Correct
Spelling	Apply basic rules	Apply basic rules		
Syntax	Might be unusual	Can be ignored	Do not change	
Style	No need		No need	
Restructure	No need if the sentence is correct		No need if can be understood	Rewrite confusing sentences
Culture	Edit if necessary	Edit if necessary		
Information	Fully delivered			
Others	Use as much raw MT output as possible	Textual standards are not important; very high throughput expectation; low quality expectations	No need to change a word if correct	Fix machine-induced mistakes; delete unnecessary or extra machine-generated translation alternatives

From Table 1, it can be seen that all proposals value the accuracy of the message and correctness of semantics by light post-editing, while grammar, syntax and style are not a big concern. O'Brien and Mesa-Lao believe that there is no need to spend too much time researching incorrect terminology, while Densmer contends that terminology should be consistent. TAUS, Flanagan and Christensen, and O'Brien hold that the spelling fixes should be applied with basic rules, and the text should adapt to the target culture. If the sentence is understandable or correct, most proposals express that it should not be restructured. O'Brien clearly points out the quality expectation for light post-editing is low. Densmer emphasizes machine-induced errors and translation alternatives in her guidelines.

Table 2. Comparative study of full PE guidelines

FULL POST-EDITING	TAUS (2016)	O'BRIEN (2010)	FLANAGAN & CHRISTENSEN (2014)	MESA-LAO (2013)	DENSMER (2014)
Accuracy	TT communicates same meaning as ST	Important	Important		Absolutely accurate
Terminology	Key terminology is correct	Key terminology is correct	Key terminology is correct	Apply the term as used in the term database for any incorrect terminology	Consistent and appropriate
Grammar	Correct	Accurate	Correct	Correct	Correct
Semantics	Correct		Correct	Correct	Correct
Punctuation	Correct	Apply basic rules	Apply basic rules		Correct
Spelling	Apply basic rules	Apply basic rules	Apply basic rules		Correct
Syntax	Normal		Correct		Make modifications in accordance with practices for the TL
Style	Fine	Ignore stylistic and textual problems		Not important	Consistent, appropriate and fluent
Restructure			No need if the language is appropriate	No need if the sentence is semantically correct	Rewrite confusing sentences
Culture	Edit if necessary	Edit if necessary	Edit if necessary		Adapt all cultural references
Information	Fully delivered	Fully delivered	Fully delivered		
Formatting	Correct	All tags are present and in the correct positions	Ensure the same ST tags are present and in the correct positions;		Correct (including tagging)
Others	Basic rules apply to hyphenation; human translation quality	Apply basic rules to hyphenation; high throughput expectation; medium quality expectations	Use as much raw MT output as possible; ensure the untranslated terms belong to the client's list of 'Do not translate' terms	No need to change a word if it is correct; accept the repetitive MT output	Perfect faithfulness to the source text; fix machine-induced mistakes; delete unnecessary or extra machine-generated translation alternatives; cross-reference translations against other resources; human translation quality

Regarding full post-editing, TAUS and Densmer expect that the quality should have no difference with human translation, and they emphasize the significance of fine style. However, O'Brien and Mesa-Lao do not agree with a need to pay much attention to the style. They expect the quality after full post-editing be medium rather than equal to translation from scratch. Should the quality after full post-editing be the same as human translation or maintain the traces of machine translation? We can see from Table 2, especially the "Others" row that the resource centre and LSP are more inclined to human translation quality than the scholars. If full post-editing should reach human translation quality, it still remains a question whether full post-editing is more pragmatic than translation from scratch in terms of cost. It is even debatable if post-editing can actually bring productivity gains, which leads to scepticism toward the benefits of post-editing. Guerberof (2009) and Federico et al. (2012) reported productivity gains in their research, while Gaspari et al. (2014) found that post-editing could lead to productivity losses over translation from scratch.

The requirements of the full PE guidelines surpass the considerations of the light PE guidelines in terms of accuracy, semantics and culture in particular. Different from light PE guidelines, most full PE guidelines require the correctness of terminology, grammar, punctuation, syntax and formatting.

5. Conclusions

From this comparative study, we can see that the existing PE guidelines have many overlaps, especially for light post-editing. The main differences lie in the full PE guidelines and concern the requirement for style and the expected quality of the target text, which we believe depends on the use and type of the text.

As we mentioned before, there are no standard PE guidelines. DePalma (2013) contends that clients should share with LSPs exactly what light and full post-editing is to be included before contracting for a job. Densmer (2014) also asserts that the quality levels, throughputs, and expectations must be defined in advance. We agree with their ideas and advise LSPs and their clients to discuss and create their own tailored PE guidelines together beforehand.

In addition to the general PE guidelines above, there are other sources of PE guidelines which are either language-dependent or aim-specific. Such guidelines include, for example, the GALE PE guidelines (WEB, b), PE guidelines with a focus on Japanese (Tatsumi, 2010), ACCEPT's guidelines for monolingual and bilingual post-editing (ACCEPT, 2011), language dependent (English-Spanish) PE guidelines (Rico and Ariano, 2014), PE guidelines for BOLT Machine Translation Evaluation (WEB, c), and PE guidelines for lay post-editors in an online community (Mitchell, 2015).

Acknowledgements

This work is supported by the Science Foundation of Ireland (SFI) ADAPT project (Grant No.: P31021). The authors would also like to thank Dr. Sharon O'Brien for her helpful comments and suggestions.

References

- ACCEPT. (2012). *Seminar Material on Post-editing – Edition 2*, available at <http://cordis.europa.eu/docs/projects/cnect/9/288769/080/deliverables/001-D622SeminarMaterialonPostEditingEdition2.pdf>
- Allen, J. (2003). Post-editing. *Computers and Translations: A Translator's Guide*. 35, 297-317.
- Belam, J. (2003). "Buying up to falling down": a deductive approach to teaching post-editing. In: *Proceedings of MT Summit IX*, Workshop on Teaching Translation Technologies and Tools (27 Sept. 2003, New Orleans, USA), pp 1-10.
- Densmer, L. (2014). *Light and Full MT Post-Editing Explained*, available at <http://info.moravia.com/blog/bid/353532/Light-and-Full-MT-Post-Editing-Explained>
- DePalma, D. (2013). *Post-editing in practice*, available at <http://www.tcworld.info/e-magazine/translation-and-localization/article/post-editing-in-practice/>
- Federico, M., Cattelan, A., Trombetti, M. (2012), Measuring user productivity in machine translation enhanced computer assisted translation, In: *Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas*, AMTA 2012 (28 Oct. – 1 Nov. 2012, San Diego, USA).
- Flanagan, M., Christensen, T.P. (2014). Testing post-editing guidelines: how translation trainees interpret them and how to tailor them for translator training purposes. *The Interpreter and Translator Trainer*, vol. 8, no. 2, pp. 257-275.
- Gaspari, F. 2014, Perception vs Reality: Measuring Machine Translation Post-Editing Productivity, In: *Proceedings of the Third Workshop on Post-Editing Technology and Practice at the 11th Conference of the Association for Machine Translation in the Americas*, AMTA 2014 (22-26 Oct. 2014, Vancouver, Canada), pp. 60-72.
- Guerberof, A. (2009). Productivity and quality in MT post-editing, In: *Proceedings of MT Summit XII-Workshop: Beyond Translation Memories: New Tools for Translators MT*, AMTA 2009 (26-30 Aug. 2009, Ottawa, Canada).
- ISO, *ISO 17100:2015: Translation services – Requirements for translation services*, available at http://www.iso.org/iso/catalogue_detail.htm?csnumber=59149
- Mesa-Lao, B. (2013). *Introduction to post-editing – The CasMaCat GUI*, available at http://bridge.cbs.dk/projects/seecat/material/hand-out_post-editing_bmesa-lao.pdf
- Mitchell, L. (2015). The potential and limits of lay post-editing in an online community, In: *Proceedings of the 18th Annual Conference of the European Association for Machine Translation*, EAMT 2015 (11-13 May. 2015, Antalya, Turkey).
- O'Brien, S. (2010). *Introduction to Post-Editing: Who, What, How and Where to Next?* Available at <http://amta2010.amtaweb.org/AMTA/papers/6-01-ObrienPostEdit.pdf>
- Rico-Pérez, C., Ariano-Gahn, M. (2014). Defining language dependent post-editing rules: the case of the language pair English-Spanish. In O'Brien, S., Balling, L.M., Carl, M., Simard, M., Specia, L., (eds.), *Post-editing of machine translation: processes and applications*. Newcastle, pp. 299-322.

- Tatsumi, M. (2010). *Post-editing machine translated text in a commercial setting: Observation and statistical analysis*. PhD thesis, Dublin City University, Dublin, Ireland.
- TAUS. (2010). *MT Post-editing Guidelines*, available at <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines>
- TAUS. (2016). *TAUS Post-Editing Guidelines*, available at <https://www.taus.net/think-tank/articles/postedit-articles/taus-post-editing-guidelines>
- Wagner, E. (1985). Post-editing Systran - A challenge for Commission Translators, *Terminologie et Traduction*, no. 3, pp. 1-7.
- WEB (a). *Post-Edited Machine Translation*. http://iconictranslation.com/solutions/post_edited_mt/
- WEB (b). *Post Editing Guidelines for GALE Machine Translation Evaluation*. http://projects ldc.upenn.edu/gale/Translation/Editors/GALEpostedit_guidelines-3.0.2.pdf
- WEB (c). *Post Editing Guidelines for BOLT Machine Translation Evaluation*. http://www.nist.gov/itl/iad/mig/upload/BOLT_P3_PostEditingGuidelinesV1_3_3.pdf

Received May 2, 2016, accepted May 8, 2016