A Review of Translation Tools from a Post-Editing Perspective

Lucas Nunes Vieira
Université de Franche-Comté
Besançon, France
lucasnvieira@gmail.com

Lucia Specia
University of Wolverhampton
Wolverhampton, UK
l.specia@wlv.ac.uk

Abstract

Machine Translation (MT) has experienced remarkable improvements and consequently grown in popularity of late. It now functions not only as an end in itself but also as a valuable asset to be exploited by translators in the promising practice of post-editing the outcome of MT systems, which can yield faster and sometimes more accurate results. Most systems, however, were not originally designed having translators envisaged as potential users, which leaves a high demand for tools capable of catering for this new translation modality. With the purpose of showcasing what researchers and the industry have to offer in that respect, this study provides a review of a number of currently available translation tools from the perspective of translation post-editing. We have selected and described toolkits according to a set of criteria, highlighting main differences and similarities between them and also making mention of desirable features that have not been satisfactorily presented by any of the toolkits analysed.

1 Introduction

As an outcome of the extensive efforts that have been invested into the field of Natural Language Processing (NLP), MT has gone through remarkable improvement in recent years. The availability of systems capable of producing fairly accurate translations made MT grow in popularity in a number of areas. Its integration with Human Translation (HT), however, does not seem to have advanced in the same proportions. Even though there is a large quantity of systems available nowadays, the vast majority of them were not tailored to serve human translators, which seems to denote a disregard for how useful a resource it can be when incorporated in the HT process.

In an attempt to respond to this gap, we have lately witnessed a tendency for Computer Aided Translation (CAT) tools, originally based on Translation Memory (TM) technology, to include MT as an option in their array of functionalities, enabling the post-editing environment once used only for TMs to be used also for machine translated text. This new feature usually consists in automatically translating the source file and providing the result to the translator as a first version of the text to be post-edited. This practice is usually referred as *MT post-editing* and is considerably promising in enhancing the translation process both in terms of time and quality.

Trados™[1] and Wordfast™[2] are amongst the tools that went along the path of incorporating MT in an otherwise TM-based editing platform. Albeit less common, the reverse has also happened. Due to the acknowledged functionality of TMs, which keep translation options close at hand and maintain texts standardised, some MT systems have also enabled the use of TMs besides providing their own translation output. In this last group are systems such as Google Translate™[3]

[1] http://www.trados.com/en/
[2] http://www.wordfast.net/
[3] http://translate.google.com/

and Systran™[4]. The use of TMs has for a long time been a established reality to professional and hobbyist translators alike, and it has a number of particularities of its own that are not a direct object of this study. Nevertheless, we look into the new correlated use of both TM and MT, as well as into any functionalities TMs may render the translator whilst post-editing machine translated text.

In this paper we describe a number of tools that can serve the purpose of MT post-editing. Having in mind the high hopes for this new practice, we have analysed these tools based on the functionalities they already have and those that would be desirable, setting out to provide an overview of what the current picture of the market of MT post-editing tools is and what users can expect from it. Whilst our analysis is not exhaustive and does not focus entirely on providing a positive and/or negative account on the selected toolkits, it gives a wide-angled view of what they currently have to offer.

2 Related Work

Few studies have been undertaken towards an analysis of currently available tools that can serve MT post-editing. Google Translator Toolkit™[5] appears to be one of the resources that mostly attract attention from those engaged in showcasing the benefits and pitfalls of such tools. Ramos (2010) provides the point of view of the translation market, having run an experiment with a translation project. She gives an account of how this toolkit can be used in professional translation environments, and finally describes the experiment as worthwhile despite some of the downsides it presented.

Eisele et al. (2009) analyse online post-editing toolkits in view of the improvements their own toolkit project should experience. Focused on the advantages the Internet offers, they call attention to the poor profit that is made from the processing of the text, previous translations and extra resources that would help polish the outcome. They argue that more interaction should be assured between the translator and the MT systems as well as between translators themselves, which are exactly the inefficiencies their toolkit responds to.

Attempting to bridge the gap between tools related to the translation practice and actual problems translators usually find in their task, Désilets et al. (2009) conducted a study observing the work of eight professional translators, aimed at determining what kind of tools they resorted to and due to what kind of problem. What is particularly noteworthy in their findings is the richness of types and number of tools the translators consulted.

The conclusions these studies yielded hold a straight connection to the analysis here undertaken in that they denote the large number of functionalities MT post-editing toolkits would ideally have to comprise to thoroughly supply the needs of the market. However, these studies either consider only browser-based systems, or focus on how human translators perceive the toolkits, as opposed to giving a general overview of the functionalities in toolkits currently available in the translation industry.

3 The Toolkits

In order to choose which toolkits to consider we have taken into account criteria such as how popular the toolkits are amongst translators, and how much attention they have received in recent research, having used the survey conducted by Lagoudaki (2006) as a guideline for tools originally based on the use of TMs. Another factor that influenced our choice of toolkits was their availability to the general public. Due to practical reasons, toolkits whose use is restricted to a given enterprise, such as that of the Pan American Health Organization (PAHO), were not included in our study. For an overview of some toolkits that fall into this category, we refer the reader to the AMTA-2010 MT Postediting Showcase[6].

Bearing these factors in mind, we have selected the following toolkits for analysis: Google Translator Toolkit™, SDL Trados 2009™, Wordfast Classic 6™/Anywhere™[7], Caitra[8], Systran 7™,

[4] http://www.systran.co.uk/

[5] http://translate.google.com/toolkit

[6] http://amta2010.amtaweb.org/AMTA/papers/6-03-LgEmPEShowCase.pdf

[7] http://www.freetm.com

[8] http://www.caitra.org/

Lingotek™[9], Déjà Vu X2™[10], OmegaT™[11], and ProMT LSP 9™[12]. We have also constrained the analysis to the flagship, most traditional toolkit owned by each developer, having chosen the latest version of the programme made available to date. In the case of Wordfast™, which has an *Online*, a *Classic*, and a *Pro* versions, for the purpose of this study we opted to discard the Pro version since, as the company itself affirms, it is not the most popular one. As to the remaining two versions, both were considered, but not separately, since they share basically the same features, with minor differences as to the way they are represented.

4 The Criteria

We have established a set of ten criteria to meet the purpose of our analysis: interface intuitiveness, existence of a spell/grammar/style checker, use of MT outputs from multiple systems, integration between TM and MT, existence of an "auto-complete" function, preservation of source text formatting, existence of a Quality Assurance (QA) function, possible log from user's feedback, confidentiality of data, and existence of scores for TM fuzzy matches.

The criterion of "interface intuitiveness" regards the way functions are presented to the user: if they are easy to use and intuitively accessible, not requiring much previous knowledge or training. We acknowledge that this is a highly subjective criterion and in this study its judgement was based solely on the experience of a single translator attempting to use the toolkits for the first time. Extensive studies with multiple translators would be necessary to provide a more credible judgement for this criterion.

As to the criterion that looked into the existence of a spell/style/grammar checker, we considered that a tool had such a feature only if it were an integrant part of the toolkit or a default plug-in, not relying on an external application left to be included by the user.

Concerning the integration between MT and TM, we noted if there were any automatic facilities capable of combining the use of both resources, instead of merely displaying them, since all toolkits that enabled the use of TMs showed both MT output(s) and TM matches to the user.

Auto-complete functions provide the user with translation options as he or she types. This feature was taken into account due to its high optimisation property. It is worth noticing that by "auto-complete" we have in mind the function of providing the translator with a list of easily adoptable words or expressions that are looked up in uploaded data or MT output based on the first letters of a word typed by the translator and also on linguistic context. This criterion does not look into "interactive translation", where a MT system actually produces new translations as the user modifies the existing text.

QA is another feature that was included as a criterion due to its relevant role. Roughly speaking, it verifies the entire translation for problems such as terminology inaccuracies - based on any term bases or dictionaries - and non-translated segments.

The confidentiality criterion is relevant only for online systems. We have checked to see if developers assured or not the confidentiality of any data uploaded by the user.

Finally, regarding the criterion on log of user's feedback, we looked for any evidence that suggested the possibility of the user's judgment on the quality of a translation - either automatic or from a TM - being taken into account, either explicitly, through a rating scale, or implicitly, through the logging of the user's corrections for example.

The criteria listed here are by no means the only relevant aspects to be considered when choosing a post-editing tool. Amongst a number of other important criteria, they were selected because they allow distinguishing one toolkit from another. Other general criteria found in all or most toolkits are presented in Section 5.1, while desirable criteria which are not present in any of the toolkits studied are discussed in Section 6.

[9] http://www.lingotek.com/

[10] http://www.atril.com/en/software/deja-vu-x-professional

[11] http://www.omegat.org/

[12] http://www.promt.com/

5 The Analysis

5.1 The Baseline

We have noticed that there are some functionalities that are shared, to a lesser or greater extent, by all or most toolkits analysed. In this section we describe these functionalities, which could be considered the baseline for any MT post-editing tool.

It is relevant to mention that even though a considerable proportion of the functions included in this section were not present in Caitra, we have still decided to take them into account as part of the baseline. We have made this choice due to the fact that Caitra is not exactly a toolkit designed to serve the translation market. As affirmed by its developer in Koehn (2009), the main objective behind Caitra is that of having a testbed for research engaged in analysing the interaction between human translators and a MT system, and the different ways in which it can serve them. The research is particularly interested in interactive translation, a feature provided by Caitra and that has already been briefly described in Section 4. Since exploring this feature is one of the main motivations behind the tool, Caitra was the only of the analysed toolkits that did not provide the user with the possibility of uploading TMs.

All toolkits put some effort into assigning intuitive meaning to the interface of the system, i.e. to exploit features of its interface that can contribute to making the translation process easier and faster. These include colour codes and other graphical resources aimed at representing information that can be useful to the translator. The toolkits including the use of TMs achieved this mostly by representing the degree of matching between the source text and a TM. The majority of toolkits profited from the traditional meaning associated with certain colours, representing full matches in green and fuzzy matches in yellow, orange or lighter shades of green. Some systems also exposed the differences in a fuzzy match of the TM by either highlighting divergent or missing terms or striking them through with a line. The only system that does not utilise any colour scheme to represent matches with the TM is OmegaT™, which presents percentage figures for matching instead. Different colours were also used to guide the translator as to which segment of the text is under translation and/or edition.

Another characteristic shared by all systems is the choice of displaying both source and target texts on screen, being it either in a horizontal or vertical layout - allowing the user to make this decision in most cases - or having both texts in the worksheet, intertwining source and target segments.

Basic and yet extremely useful features, such as the possibility of employing keyboard shortcuts and customising the disposition of panes on the screen, were also present in the majority of systems, with the exception of Caitra, which lacked a "find and replace" function. OmegaT™ has a weakness in this respect since, in spite of its detailed search feature, it does not provide the option of automatically replacing a searched term in the entire text at once.

Caitra was the only system not offering a concordance search option. The other toolkits either had a function specifically designed for this purpose or showed a searched term in both source and target files, allowing the user to see how the term had been translated throughout the document, which was shown either at once or with one occurrence at a time.

The possibility of uploading dictionaries and/or glossaries that can be user-made and user-augmented was also found in the majority of programmes, Caitra being the only exception.

The existence of a function or supplementary application specifically designed for the compiling of TMs and/or term bases was also found in most tools, with the exception of Google Translator Toolkit™ and Caitra.

A function aimed at displaying information related to the translation process, such as figures representing the proportion of TM matches, fuzzy matches, and machine translation text adopted, was also found in all systems but Caitra. This type of information can reveal translation trends that could be useful either for research purposes or for the translator's own record.

5.2 Additional Features

In regard to features that were not common to all or most toolkits, we have used the criteria described in Section 4 to analyse the systems either on a yes/no basis, as in bearing or not bearing a given function, or in a more detailed evaluation of to what level they achieve a given criterion, or what exactly they have to offer. The product of our analysis can be seen in Table 1. Complementary to Table 1, what follows is a summary of the main assets of each tool as well as relevant differences that were noted between them.

Google Translator Toolkit™, the online platform developed by Google for translation and/or MT post-editing, has as a frequently referred downside from the perspective of commercial translation the fact that there is no strict assurance from the developers that the data uploaded by the user as well as his/her translation options will be kept confidential. Also in its disfavour, it lacks a QA checker and poses the stated risk of the formatting of the source file not being fully preserved. From the viewpoint of its usability and interface, however, the tool proved fairly easy to manipulate. As to its TM facility, it does not display the score of fuzzy matches on screen, but it has a colour code to indicate the differences in the text.

With a long tradition in the market of CAT tools, SDL Trados 2009™ was one of the toolkits that demonstrated the widest array of functions, being on a very similar level to Déjà Vu X2™ in that respect. SDL Trados 2009™ allows the user to upload TMs and also provides the MT output of two systems, Google Translate™ and Language Weaver. Showing a certain degree of interaction between the TM and MT, it allows populating with MT segments not matched in the TM. Its interface is extremely informative with a number of panes that include the translation environment itself as well as other functions and details related to the project. Due to its large quantity of functions it could prove slightly complex, especially for inexperienced users, which also applies for Déjà Vu X2™. For that reason both these tools were described as having medium interface intuitiveness, as shown in Table 1.

SDL Trados 2009™ also has a very helpful QA device that warns the user of problems in the translation signalling problematic segments with an exclamation mark. The QA function checks for untranslated segments, problems with spacing and also terminology - whenever a term does not receive the translation that is expected for it according to an uploaded term base, the QA function displays a warning. Wordfast Classic 6™ and Déjà Vu X2™ have similar devices for QA. Wordfast Classic 6™ finds untranslatable terms in both source and target texts, checks for typos in proper names and provides warnings about differences in final punctuation between source and target and also when the number of characters between them is too discrepant. Déjà Vu X2™ assures translation quality by means of running consistency checks in a number of levels. It looks for discrepancies in terminology and also for different translations that might have been given to identical sentences in the source text. Systran™, in turn, has reviewing panes that draw the user's attention to problems in the translation. It also highlights untranslatable terms.

Déjà Vu X2™ showed a particularly good integration of MT and TM, allowing the user to repair fuzzy matches with the use of an Example-based Machine Translation system. Whenever the system encounters a fuzzy match, it can be set to look for the appropriate translation of non-matched terms in any uploaded terminology base. If the search is successful the fuzzy match may be turned into a perfect match due to the automatic corrections made by the system. It has a very informative interface, with an extremely clear representation of TM matches and MT output, making use of colours, bars and percentage figures. It has an "auto-complete" function that suggests terms and phrases based on uploaded content as the translator types. SDL Trados 2009™, Caitra and Wordfast Classic 6™ also bear such function, although in a slightly different way. In SDL Trados 2009™, an "auto-suggesting dictionary" has to be uploaded specifically for this purpose. In Caitra, the suggestions come from the phrase table of the Statistical Machine Translation (SMT) system Moses (Koehn et al., 2007). In Wordfast, suggestions can be based either on uploaded content, on MT system or on the Web. Also, Déjà Vu X2™

Criteria	**Google Translator Toolkit** (online-free)	**SDL Trados 2009** (trial)	**Wordfast Classic 6 / Anywhere** (demo / online-free)	**Déjà Vu X2** (trial)	**Systran 7** (trial)	**Lingotek** (online-trial)	**Caitra** (online-trial)	**OmegaT** (free)	**ProMT LSP 9** (trial)
Interface Intuitiveness	High	Medium	Medium	Medium	Medium	High	Low	Medium	High
Spell / grammar / style checker	Yes: spelling	Yes: spelling. Hunspell[12] as default plug-in	Relies on text editor/ browser	Yes: spelling	Yes: spelling	Relies on browser	Relies on browser	Relies on external source, no default plug-in	Yes: spelling. ORFO[13]as default plug-in
Multiple MT entries	No: just Google Translate	Yes: Google Translate, SDL ATS and Language Weaver	Yes: Google Translate and Microsoft Translator	No: just Google Translate	No: just Systran	Yes: Google Translate and Microsoft Translator	No: just Moses[14]	Yes: Google Translate, Apertium[15] and Belazar[16]	No: just ProMT
TM + MT	No automatic interaction	Interactive: populates non matched segments with MT	No automatic interaction	Interactive: repairs fuzzy matches with Example-based MT	No automatic interaction	No automatic interaction	No (no TM)	No automatic interaction	Interactive: populates non matched segments with MT
Autocomplete function	No	Yes	Classic 6 - Yes / Anywhere - No	Yes	No	No	Yes	No	No
Preservation of formatting	Not fully assured	Yes: shown on screen/ formatting tags	Yes: formatting tags	Yes: embedded codes	Yes: shown on screen	Yes: formatting tags	No	Yes: formatting tags	Yes: shown on screen
QA	No	Yes	Classic: Yes/ Anywhere: No	Yes	Yes	No	No	No	No
Possible log of user's feedback	Yes: rating with stars	No evidence found	No evidence found	No evidence found	No evidence found	Yes: rating with stars	Yes: data collection	No evidence found	No evidence found
Confidentiality	Not assured	Assured	Assured on both versions	Assured	Assured	Up to user	Not assured	Assured	Assured
Score for fuzzy match level	No	Yes: bar, colour and percentage	Yes: colour and percentage	Yes: bar, colour and percentage	Yes: three levels. Not shown on screen	Yes, but not shown on screen	No	Yes: percentage	Yes: colour and percentage

[12] http://hunspell.sourceforge.net/
[13] http://www.orfo.ru/
[14] http://www.statmt.org/Moses/
[15] http://www.apertium.org/
[16] http://belazar.belinter.net/

Table 1: Analysis of eight selected toolkits based on the set of established criteria

offers the user the option of choosing the subject of the text, which leads the system to a better and more adequate choice of TM matches. This is particularly useful for users with huge amounts of uploaded data. The system has a number of algorithms that, whenever in face of more than one match, choose the best option taking into account, apart from the level of fuzziness, details such as the client and subject of the translation. A similar function is shared by Systran 7™ and by ProMT™. Both systems apply different translation rules to the input document depending on the genre it has been labeled with.

As affirmed previously, two versions of Wordfast™ were considered in this study: Wordfast Anywhere™ and Wordfast Classic 6™, both with very similar features, but working on different platforms. Wordfast Classic 6™ works as an add-in tool for Microsoft Word™, whereas its Anywhere version can be accessed through the internet. Concerning the online version, it is worth noticing that the vendor affirms that any uploaded data will remain confidential unless opted otherwise by the user. Also very traditional in the market of CAT tools, Wordfast™ was initially designed as a TM tool only, but its newer versions also count on MT for post-editing, providing the outputs of Google Translate™ and Microsoft Translator™. Wordfast™ offers the translator the possibility of using its own TM, called VLTM - Very Large Translation Memory -, which is accessible over the internet and is available in both versions here considered. Similarly to TMs uploaded on the Anywhere version, vendors claim that any content or TM uploaded by the translator whilst in use of the VLTM will only be retrieved if he/she sets the programme to do so. The existence of an ”auto-complete” function as well as of a QA checker were the most relevant differences we have noted between the versions of Wordfast™ here analysed: Wordfast Anywhere™ does not have these functions.

Despite being essentially a rule-based MT system, Systran™ also presents an interface for post-editing. It displays both source and target texts side by side on screen, with the latter being editable. It also allows for the upload of TMs, leaving for the user to decide what match level would make the system perform the translation based on the TM instead of on its own set of rules. While most systems allow the user to indicate the fuzzy match threshold by establishing the exact percentage or score, Systran™ has only three levels to choose from: strict, normal, and flexible. It also proved extremely useful in maintaining the formatting of the source file, being Systran™ itself and ProMT™ the only tools from the ones analysed that fully display the formatting of both files on screen.

Lingotek™ is not only a translation and MT post-editing tool but also a platform where translators can exchange information and revise each other's text. It differs from the other analysed toolkits in that it is also used as an intermediary between clients and translators, who can be assigned with projects through the web site, stating on their personal profile how much they charge for their services. As an MT post-editing tool, Lingotek™ proved fairly intuitive, allowing the translator to upload his/her own TM or use a public one provided by the web site. It has a function that looks up desired terms in the Google Index, showing the search engine embedded in one of the panes of the system.

Lingotek™ and Google Translator Toolkit™ were the only tools to show clear evidence of collecting feedback from translators. Apart from the rating scale that comes attached to each translation segment, in the form of stars, which Google Translator Toolkit™ also has, Lingotek™ gives the translator the possibility of reporting a translation as incorrect and forwarding it for revision. Besides the online platform, it has other versions that were not considered here for practical purposes.

Caitra is a tool that was designed with the main objective of collecting translation data for research, as previously stated. It works integrated with Moses and is freely available online. Briefly put, the tool displays the MT output suggested by Moses for the source text uploaded by the user. The only options to enter a source text are by either pasting it or typing it in a text box. Apart from the suggested translation phrases, the user also has access to a number of alternative translation options found by Moses, which are dis-

played in different colour shades proportionaly to their probability of occurring in that context - the darker the shade, the more likely the translated phrase was deemed by the system. The fact that all these translations are displayed in this informative colour code implies that Caitra provides some degree of information regarding the quality of the translation proposed for post-editing, a trace found solely in Caitra and ProMT. Caitra is also the only tool to perform "interactive translation".

From the three tools that present evidence of collecting users' feedback, Caitra is the only one whose feedback is provided implicitly, as the user's choice in adopting or not a translation suggested by Moses. The text entered by the translator and keystrokes are also stored by the tool.

OmegaT™ and Wordfast Classic 6™ were the only non-online toolkits that offered versions for operational systems other than Windows. Although easy to use, OmegaT™ was deemed to have low interface informativeness. On the other hand it was the only tool to offer by default MT outputs from three different systems: Google Translate™, Apertium and Belazar.

ProMT LSP 9™, similarly to Systran™, essentially consists in a MT system that provides the possibility of post-editing machine translated text. It comes along with a package of applications that allow the use of the tool in association with other programmes such as text editors and instant messengers. The tool also performs translation of entire documents, supporting a number of different formats such as .pdf, .doc, and .ppt, keeping also the text formatting of the source file. ProMT™ is the only system that allows the user to make decisions whenever its set of translation rules comes across a case of ambiguity. ProMT LSP 9™ also offers the option of editing the translation rules through a rule editor application, making it possible to tailor the system to the user's needs. As a MT system itself, the tool does not provide MT outcomes from other sources. It was classified as having high interface intuitiveness due mostly to the relation between its numerous features and applications and the reasonably little effort it takes to employ them.

6 Desirable Features

A number of functions that from translators' point of view would be desirable to MT post-editing were not found in any of the analysed toolkits. A better integration between MT and TM, for example, could be highly useful. Déjà Vu X2™ was the only system that showed some development in this respect, allowing for the possibility of repairing fuzzy matches of the TM with translations from a MT system. Whilst most systems can be set to automatically translate segments that were not matched with the TM, Déjà Vu X2™ allows the user to do so on a sub-segment level, transforming fuzzy matches into perfect matches. Nevertheless, in our analysis we have not encountered the option of making sub-segment corrections of fuzzy matches through MT systems that are not Example-Based.

Even though some systems keep track of keystrokes and other information related to the translation process, a satisfactory change tracking function was not found in our analysis of the toolkits. The vendors of Trados™ claim to have included this function in a version of the programme that is yet to be released. However, it is to date an asset the market seems to lack. Such a feature would be extremely useful for the occasion of having more than one professional working on the same text. It would ideally expose the changes performed by each of them in a clear way so that the changes could be easily discarded or further modified in a later stage of the process.

Also not met by any of the tools is the capability of providing an indication of the quality of a MT output to be post-edited at the sentence or sub-sentence level. This is believed to be a highly valuable feature since the translator would be able to filter out translations that are not worth post-editing. For translators, such quality indicator could be interpreted in a way that is similar to a fuzzy match score in TMs. As previously remarked, Caitra and ProMT were the only tools to show some initiative in this respect. In Caitra, however, this information refers solely to the likelihood of the translations provided by the system, which could be seen as a form of MT system confidence estimation, as opposed to a global indi-

cator of translation quality. ProMT LSP 9™, in turn, has an application that offers the possibility of performing a comparison between two different translations of the same source text, providing the number and proportion of mismatches - by character or by word - between the translations. The percentage provided by ProMT™ could be considered indicative of translation quality as it highlights convergences between two different translation versions, and calls the user's attention to differences that might indicate inaccuracies from one of the parts. Nevertheless, the system is in fact simply performing the comparison of two texts and exposing the differences between them, without providing more specific information based on the estimated quality of the translations. Moreover, if both translations are incorrect, this comparison will not be informative. Specifically engaged in contributing to the use of quality indicators of MT for post-editing is the study carried out by Specia (2011), which shows that quality estimation models learnt from straightforward annotations of translation quality are particularly promising in enhancing post-editing. However, such quality estimation models have not yet been incorporated in any of the translation toolkits studied here.

More sophisticated on-screen alignment resources are also desirable. Systran™ and ProMT™ were the only systems to keep source and target texts connected by highlighting in both the exact part of the text where the pointer was placed. In both systems, this "on-the-fly alignment" was done either on a word or on a phrase basis, which mostly resulted from expressions that held a higher degree of fixedness. These expressions were highlighted as a single unit, which not necessarily corresponded to a counterpart with the same lenght, the alignment being done either on a 1-n or on a n-1 basis between both languages. This feature was considered extremely useful and could perhaps go through further improvements and be incorporated by other toolkits. In ProMT, we have noticed that any edition performed in the text incurred in a loss of alignment in the entire document. The same occurred in Systran in some occasions when a term or expression was edited or deleted, but in the case of Systran™ the loss of alignment was limited to the edited part of the text. This perhaps could be improved so as to keep the alignment despite the modifications and/or necessary deletions that can be performed.

7 Conclusion and Future Work

This review has shown at first that MT post-editing is already a functionality that is offered by most relevant translation tools currently available, both for those originally designed as a TM tool and those originally designed as MT systems.

We have discussed a number of features that are shared by all or most toolkits selected, suggesting a starting point for new tools aiming to enter the market. Furthermore, our study has also highlighted characteristics that are specific to a subset of the toolkits, which can be used as a guideline by translators and language service providers looking for post-editing tools.

Despite the fact that MT post-editing is already a reality for the translation market, our review has shown that there is still considerable room for improvement, as a number of features deemed desirable for the work of a translator were not satisfactorily found in any of the tools analysed.

As future work, we intend to update this review to include the analysis of other tools. Post-editing is a popular topic and we expect new tools and adaptations of existing tools to be constantly released. One of such new tools is SmartMATE, by Applied Language Solutions (Way et al., 2011), which integrates translation memories, glossary management and multiple, customisable machine translation systems for concurrent editing, proofreading and reviewing.

Additionally, we would like to conduct an enquire with professional translators about the negative and positive aspects of each tool so as to draw conclusions on the tools that mostly meet translators' needs.

References

Désilets, Alain, Christine Melanon, Geneviève Patenaude, and Louise Brunette. 2009. How translators use tools and resources to resolve translation problems: an ethnographic study. In *MT Summit XII: Proceedings of the twelfth Machine Translation Summit*, Ontario, Canada, August.

Eisele, Andreas, Christian Federmann, and James Hodson. 2009. Towards an effective toolkit for translators. In *Proceedings of the ASLIB International Conference Translating and the Computer 31*, London, UK, November.

Koehn, Philipp, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, Chris Dyer, Ondřej Bojar, Alexandra Constantin, and Evan Herbst. 2007. Moses: open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, pages 177–180, Prague, Czech Republic.

Koehn, Philipp. 2009. A web-based interactive computer aided translation tool. In *Proceedings of ACL/AFNLP (Software Demonstrations)'09*, pages 17–20.

Lagoudaki, Elina. 2006. Translation Memories Survey 2006: Users perceptions around TM use . In *Proceedings of the ASLIB International Conference Translating and the Computer 28*, London, UK, November.

Ramos, 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 (AMTA)*, Colorado, USA, November.

Specia, Lucia. 2011. Exploiting objective annotations for minimising translation post-editing effort. In *Proceedings of the 15th Annual Conference of the European Association for Machine Translation (EAMT)*, Leuven, Belgium, May.

Way, Andy, Kenny Holden, Lee Ball, and Gavin Wheeldon. 2011. SmartMATE: Online Self-Serve Access to State-of-the-Art SMT. In *Proceedings of the 3rd JEC Workshop Bringing MT to the User: Research Meets Translators*, Luxembourg.