

## Differentiating editing, post-editing, and revision

Félix do Carmo (University of Surrey)

Joss Moorkens (ADAPT Centre, CTTS, SALIS, DCU)

### **Abstract:**

*While several studies report translator resistance to post-editing, translators whose work has followed the evolution of translation technology may consider post-editing to be translation with just another input. The resistance narrative comes from translator perception and from an industry view of post-editing as a low-cost, low-skill revision task. This chapter challenges contemporary views of post-editing as revision, drawing from academic research, along with descriptions of practices and workflows from the industry, to argue that the addition of machine translation to translation workflows requires even more specialisation of translators. In arguing that post-editing should be viewed as a form of translation, we consider how translators are required to fulfil quality expectations and to perform higher order tasks than swapping words. The central focus of our analysis is the role played by editing, understood as four actions (deleting, inserting, moving and replacing), in distinguishing between the main processes performed by translators. We suggest a threshold that separates editing from translating, along a continuum of editing actions, and we propose that the development of editing support tools based on the four actions may increase efficiency, while also improving the usability and acceptability of the post-editing outcome.*

In her article on the transition from statistical to neural machine translation (NMT), Kenny (2018) argues that NMT is a sustaining rather than disrupting technology, a linear progression along a continuum. Similarly, we believe that for translators whose work processes have evolved alongside translation technology, post-editing (PE) may be just one more step in that progression, with machine translation (MT) suggestions acting as another contributory input to the translation decision process alongside translation memory (TM) matches, terms, and concordances. The aim of this chapter is to draw on the editing task that is present in translation, in revision, and in PE, to clarify the impact of translators' use of MT. The chapter critically analyses views and narratives about PE from Translation Studies, MT research, and the industry. The alternative view we propose calls for further discussion and study of the technical dimension of translators' work, and it draws on translation process research to recommend a re-understanding of PE as a translation process rather than a revision one. As a consequence of this new understanding, we claim that, for MT content to be used efficiently, specialised users with specialised tools are required.

The first part of the chapter (sections 1 to 3) draws on existing studies of PE, presenting the terms and assumptions upon which the subsequent sections are built, discussing the existing narratives in academia and industry (in section 2), and drawing out implications of these views for professional translators. In section 3, we set out reasons for opposing these narratives and for

considering PE to be a form of translation, informed by theoretical considerations and practical analyses of industrial workflows. The second part of the chapter (sections 4 to 6) details the role of editing in our alternative view of PE, analysing the consequences of this view in terms of the need for specialised tools and specialised users of technology. We consider in detail what editing is and its role in the study of PE, supported by translation process research. Two fundamental elements arise from our analysis: the threshold that separates editing from translating, and the description of editing as four actions. Section 5 applies these elements to the description of tools that specifically support editing, while in section 6, the views that were built up over the chapter converge into our conclusion that, although MT might seem to be a process that replaced translators in the translation process, in fact it requires even more specialisation by translators.

### **1. Setting the grounding notions**

The grammatical metaphor which allows the use of gerunds as nouns in English highlights the connection and differentiation between *words* and *what happens during the processes they refer to* (Halliday & Martin 1993). The use of the form ‘post-editing’ is preferred to ‘post-edition’ because this stresses that PE is eminently a *process*. Its analysis should therefore focus more on how it is done rather than on the results. In this chapter, we use gerunds to refer to the tasks and actions that are parts of processes, and nouns when we refer to a whole process from an external point of view.

The term ‘translation’ is used as a global technological process performed by one person, or by a group of people, each one working on a separate part of the same translation project. The word ‘translating’ will be used to refer to a specific type of writing task that implies generating and composing a sentence, writing it ‘from scratch’, as opposed to editing only parts of a translation suggestion.

‘Revision’ is also used to describe a whole process, performed by a person *different* from the one who created the translation (and thus not ‘self-revision’). ‘Revising’ is used to describe the activities performed while engaged in a revision process.

Translators read the source text and write a translation. If the text they read and the text they create in the target language are clear, the flow of writing is fluid. They only stop when they encounter a problem. Then they read the source text again, along with what they have just written. In this way, translating may be seen mostly as writing with pauses when problems occur. Translation is a process of creation, and it is by writing that the translation is created. Contrary to this, a reviser mostly reads when revising (Mossop 2020: 116). In fact, revisers hope to simply read, check and validate the work done by a good translator. They need only write when they encounter problems.

Writing and reading seem to set the boundary between translation and revision, though this is of course a simplification (translation process research such as Carl, Dragsted & Jakobsen (2011) has shown how interconnected these activities are). The difficulty in setting them apart stems from the fact that translating (mostly writing) occurs during revision and revising (mostly reading) occurs during translation.

As will be described below, the term ‘post-editing’ entered the industrial world of translation when MT systems started to produce content in target languages that was deemed to be of good enough quality to be edited and improved by translators. So, it is *as if* MT has replaced the translation process, and PE occupies the place of the revision process.

As for the term ‘editing’, it is used here to describe a type of writing task that is different from translating. In editing, translators act on a segment of text, be it a suggestion already rendered in the target language, or a segment still written in the source language, with either one requiring only a few changes here and there to be ready for validation. The term may be used to describe the actions performed during PE, and it may also be used to describe the actions carried out to update the translation of a fuzzy match from a TM. Editing is a writing action that happens after ‘checking’, which is a reading task that has the purpose of identifying whether the segment should be validated or edited.

In the film or publishing industries, editing is part of the creation process, but it is only performed once all parts of the intended result have been produced. Editors apply surgical actions to remove bad sections from a sequence, they decide what to fit into a specific place, they move sections to their optimum positions, and they try to select the best options for the different parts of a film or a book. Likewise, deleting, inserting, moving and replacing are the four actions that compose ‘editing’. These will be referred to as ‘editing actions’, but one may also talk about ‘edits’.

## **2. Narratives of resistance to post-editing**

Descriptions of translation in an industrial environment highlight efficiency and quality as two key requirements (Sager 1993). In this context, any technological advance may be characterised as progress, with any form of resistance inviting criticism. But are the narratives of translators’ resistance to PE realistic or is PE already a common practice, at least for translators working in the localisation industry? We should stress that this industry (consisting of companies that offer translation, language and technology support services to global corporations) is characterised by long workflows and supply chains in which work methods are standardised. Translators who work in these supply chains use tools and follow specifications and instructions that are defined in the upper portion of the chain in response to the technological needs of corporations and global management processes. In the early 1990's, these translators started working with computer-aided translation (CAT) tools because that allowed for more control over the processes and outputs, and they soon became recognised as keen adopters of advanced technologies. The volume of work produced by the localisation industry implies that a large and growing proportion of professional translators currently work with the aid of technology. In this section, we analyse the foundations of the narratives this industry creates and their consequences, and we suggest that these narratives should be based on more realistic descriptions of the production processes.

Many research publications either describe or assume a strong resistance from translators to PE. Allen, for example, says that PE introduced a new challenge to professional translators, because of the pressure towards “the acceptance and use of half-finished texts” (2003: 297).

Austermühl refers to two major attitudes towards this type of challenge: “satisficers”, for whom the correct strategy is not looking for optimal solutions but meeting adequacy criteria, and “optimisers”, who commit to “the art of finding the best choice among all choices” and who may face an identity crisis because of industry pressures (2013: 331).

Two examples of studies that have focused on the causes of translators' resistance to PE are Cadwell, O'Brien & Teixeira (2018) and Moorkens & Way (2016). Both studies are based on data from real users and interviews with practicing translators. They suggest that use of MT by translators may increase if they feel a greater sense of agency and have greater confidence in the utility of the MT suggestions.

In addition to translators, we can analyse the voice of the industry as expressed in the documents produced to standardise procedures. This is how the ISO 18587 standard explains the reasons for PE adoption:

*The use of machine translation (MT) systems to meet the needs of an increasingly demanding translation and localization industry has been gaining ground. Many translation service providers (TSPs) and clients have come to realize that the use of such systems is a viable solution for translating projects that need to be completed within a very tight time frame and/or with a reduced budget. (ISO 2017: v)*

The justification for such a budget reduction – Lommel (2018) suggests that PE is, on average, paid at 60-65% of the full word rate – is that PE is a form of revision rather than translation (see section 3). The advantages for translators are presented as “to improve productivity”, “to reduce turn-around times” and “to remain competitive in an environment where clients show an increasing demand for using MT in translation” (ISO 2017: v). We can easily intuit that these factors add pressure rather than satisfaction. The message to translators is: ‘post-edit or perish’. Furthermore, the standard guidelines for post-editing extend further control over translators’ work, with the need for results to be comprehensible, readable or indistinguishable from human translation, and the additional requirement that this must be achieved using as much of the MT output as possible (TAUS 2010).

In this context, resistance by translators to the increased pressure from their usual work practice seems quite reasonable. Translators may feel that the limits of job satisfaction are being stretched in professional contexts where translation or textual work is only one of an increasing array of skills and competences that they are expected to embody (European Master's in Translation Network 2017). Unilateral imposition of PE from an unpredictable MT source, paid at a lower word rate, makes PE feasible only if the processes involved are much more efficient than normal translation. Such changes may also bring about the feeling that translators are the sole guardians of quality, the ones with the role of identifying problems of an unknown nature in fleeting details.

There are enough indications, however, that the resistance narrative may not correspond to reality. A close look at the production processes in a typical translation workflow in the localisation industry shows that PE is almost always done using modern CAT tools, designed for optimising work with TM rather than MT (Moorkens & O'Brien 2017). In these environments,

MT suggestions appear intermingled with TM matches as resources for translators to check and edit. In addition, CAT tools show suggestions for terminology, allow for searches of words in context, show predictive writing suggestions and repaired fuzzy matches. The inclusion of suggestions from MT is not a major disruption to the growing complexity of CAT tools, but a natural evolution. When MT becomes an added resource, the distinction between editing TM suggestions and post-editing MT suggestions becomes less obtrusive. This is borne out by the similarities between high-quality TM matches and NMT output, in terms of editing effort and the comparative perceptions of their usefulness, as reported in Sánchez-Gijón, Moorkens & Way (2019).

In translator training, PE has already become a standard practice. Many translation programmes have, in one way or another, incorporated PE into their curricula since at least 2009 (Guerberof & Moorkens 2019) and teaching of PE is now expected for all courses included in the European Master's in Translation Network (2017). Researchers have even presented suggestions for PE to be taught as extensions of translators' skills (Kenny & Doherty 2014). We may thus say that there is a whole generation of graduate translators for whom PE is expected as part of their jobs.

Furthermore, industry statistics reveal that PE is growing steadily, at least since TAUS published their guidelines (2010). Lommel & DePalma (2016) refer to an estimate by 56 enterprise clients that the percentage of PE in the global translation industry could reach 10% in 2019. If we add to these numbers 'unofficial PE' (when translators are given source language text in the target window and they decide to machine translate it, instead of typing it over with their own translation), the numbers show that, notwithstanding natural variations between local realities, PE is much more widely used than the resistance narratives lead us to think.

There may be good reasons why the discourse on translators' resistance caught on so steadily. In the current state of affairs, any type of empirical research will oppose the narrative of resistance, thus serving as reinforcers of the reasonableness of PE, but in the process it contributes to acceptance with no criticism. In fact, it is easier to argue against unreasonable fears than to answer difficult questions, such as: 'Have the technological advances really helped translators become more efficient and produce better quality, or is the industry simply selecting the satisficers who are happy to quickly select the least bad of all options?' and 'Does the industry want translators to be trained to be satisficers, or is it ideal to train them to be flexible optimisers?' Whether misgivings are reasonable or not, the choice of whether to post-edit should always be consultative, as imposition of any process on the translator will lead them to "feel that the material agent gets precedence and is inevitable, no matter how unfitting it might be for the task at hand" (Cadwell, O'Brien & Teixeira 2018: 17).

We propose that new views of PE should be brought forward, so that questions like these are more frequently studied and debated. In the next sections, we present a few arguments towards the claim that PE needs to be studied in more detail, beyond the acceptance/resistance divide.

### 3. Post-editing as a specialised form of translation

Seeing PE as a form of revision assumes that the MT system completed a full translation. This is the prevailing notion behind the views of PE as an undemanding task in terms of language skills, one that can even be performed by monolinguals without access to the source text (Krings 2001, Schwartz 2014).

The arguments in favour of viewing PE as a form of revision are strong. For example, several papers show that most of the time spent on PE is spent on pauses rather than on keyboard actions (Koehn 2009, Ortiz-Martínez *et al.* 2016). This suggests that PE is, like revision, more associated with reading than with writing. However, we argue that this is not a sufficient reason to classify PE as revision, and that the current context in which PE is performed challenges that view.

The first claim against PE being identified as a type of revision is the argument presented by translators when they are confronted with a translation that needs to be revised but then realise that it has actually been produced by an MT system. Translators, quite properly, refuse the classification of such assignments as revision on the basis that these jobs imply more translation than revision (they involve more writing than reading, or at least the ratio between the two is not consistent with that expected of revision work).

To be a form of revision, the only aim of PE would be to eliminate errors from a finalised translation, the difference being that this finalised translation had been produced by an MT system rather than a human translator. But in a professional translation workflow, this fundamental condition is not met by PE, because MT does not produce a finalised translation. MT text is only an ‘output’, or a set of ‘suggestions’ or ‘hypotheses’ for the translation of a text. Instead, it is the post-editor – the translator – who is responsible for the final translation.

Post-editors must recognise the special status of the suggestions presented by MT systems when accepting a PE job. They need to be aware of their likely error types and unpredictable quality level. In our view, MT processes are what Catford calls ‘transference’ rather than translation (1965: 43): the output is solely based in source language meaning, without cognisance of how a concept ought to be expressed in the target language. House, for example, comments on how such a strategy is inadequate for translation between German, which requires “informational explicitness”, and English, where inference is based on context (2003: 169). If translators fail to consider reader expectations in their target language, they will most likely do a bad job and miss most of the detailed errors that the MT text contains.

A similar situation exists at the end of PE: it is not proven that a PE process produces a revised text. Due to specifications that restrain production time and limit expected quality, post-editors are under pressure not to revise their work before delivery. The perspective of PE as revision considers ‘self-revision’ of PE a redundant task. This also takes the form of a pressure not to over-correct; the aim is for post-editors to be efficient satisficers, not stalling optimisers. However, in current industrial workflows it is not uncommon for post-edited texts to go through a full revision by a different translator before delivery to clients, because that is the only way to guarantee fitness for purpose. In DARPA’s handbook for the Global Autonomous Language

Exploitation program, the description of PE production is of a two-stage process that includes a revision pass by a second translator (Dorr 2010). In a broad study of interactive PE carried out with universities and translation companies in collaboration, the outputs of PE were revised by different translators, not for evaluation purposes, but for quality assurance (Sanchis-Trilles *et al.* 2014). This revision of PE could not be accepted in commercial contexts, which are so opposed to redundancy, were it not that PE is in fact a form of translation.

The current context in which PE is performed, within CAT tools, also raises questions about classifying PE as revision: does it make sense to say that if translators edit a TM fuzzy match they are translating, but if they edit an MT suggestion they are revising? Sánchez-Gijón, Moorkens & Way (2019) found both processes to be quite similar. There are ongoing discussions and tests to identify the threshold at which a suggestion from the MT system is more useful than a fuzzy match from TM, and setting this arbitrarily may harm performance (Moorkens & Way 2016, Zaretskaya 2019). More importantly, this shows how close translating and revising are to each other in PE.

In professional settings, PE is determined by specifications, requirements of style guides, client terminology and consistency, among many other external factors that change from project to project, but which are also updated from assignment to assignment. However, some of the most cited studies on PE productivity exclude most or all of these factors, testing only one of two scenarios: very limited test conditions, employing students who focus on language issues without considering any external factors, or difficult-to-reproduce ideal lab conditions with professional translators who are experienced and work daily in the same conditions used for the tests (Vasconcellos 1987, Zhechev 2014). One study even suggests that PE does not require as much research as translation or revision (Wagner 1985), a claim which can only be valid in those ideal scenarios in which the expertise of translators makes that research redundant. Moreover, many of these studies are done with simplified interfaces, in which sentences appear in isolated text boxes with no TMs or terminological support (Plitt & Masselot 2010). Often in these studies all segments need to be post-edited, excluding the need for the translator to decide whether to validate, to edit, or to retranslate a segment from scratch because of conflicting support resources, as so often happens in real scenarios. When studies are done in actual production settings, including TMs, termbases and MT content, productivity gains are not as high as other studies claim (Läubli *et al.* 2013).

Claims about increased productivity have, nevertheless, been generalised to all work environments. This generalisation is particularly problematic when we know that it is impossible to test and measure in the lab all external and internal factors involved in professional PE. To mention just one internal factor, types of errors produced by MT systems are not reproducible if one uses a different system, a different language pair, different training data, or simply changes the test data. This unpredictability of MT output and errors has been especially evident since the advent of neural MT (Castilho *et al.* 2017, Daems & Macken 2019).

The description of PE, based on these lab tests, is therefore incomplete and of limited validity. Outside of lab conditions, PE may, for example, require more complex reading and

writing than revision, or even than translation from scratch. Krings (2001) has shown that reading slows down when MT content is added, particularly when it is medium-quality content, which requires a careful analysis to decide whether it is best deleted or retained to be adequately edited (as was also found with statistical MT by Moorkens & Way in 2016). This also reveals that, although one may consider that any content already in the target language will be a welcome help for translators, even ‘satisficers’ may not like to have all the segments processed by MT. One needs to accept that the decision to delete and retranslate a whole MT sentence may be the most efficient one, in view of contents that require either extensive reading or extensive writing.

Taking into consideration all that happens at professional translators’ workbenches during PE, we propose that it should be considered a type of translation. Not only because PE represents an evolution of industrial translation processes and because it fulfils the same purpose as translation (to produce a good target text in an efficient and effective way), but also because it requires advanced writing and reading skills in two different languages. In the next section, we focus on one of these skills: editing.

#### **4. What editing looks like**

The technical dimension of translation has often been described in terms of editing actions. Nida considers “additions, subtractions, and alterations” to be “techniques of adjustment” (1964: 226), and Toury presents “omissions, additions, changes of location and manipulations of segmentation” as indications of what he calls “matricial norms” (1995: 95).<sup>1</sup> Pym refers to “actions” being “what we actually observe translators doing (e.g. typing, correcting typographical mistakes, looking up terms in glossaries, etc.)” (2011a: 95), and Angelone classifies the editing activities of “additions, deletions and revisions” as “production-behaviours” (2010: 21). Finally, Alves and colleagues use the term “production segments” as being composed of “revisions, deletions, substitutions, etc.” (2010: 125). Consequently, it seems difficult to find a consensual classification of the micro-procedures translators carry out when they are producing their translations, although these lists all incorporate similar action terms.

Krings (2001) mentions that keyboard actions may be good subjects for research because they are easily observed. He analyses linear writing processes and non-linear ones. In linear processes, translators add elements in a sequence with pauses, but without interrupting the sequence. We call this ‘translating’, because the translation is produced by following a syntactical generation procedure. Non-linear actions are mostly deletions and insertions, with translators primarily engaging in replacement or overwriting operations, especially when making short edits. We call these non-linear actions ‘editing’, since the translator is reading and only intervenes in the text to delete, insert, move or replace units. Both forms of writing (translating and editing) occur in translation, in revision, and in PE.

---

<sup>1</sup> Matricial norms govern the presence of target-language material reflecting source-language content, the location and distribution in the text of this material, and choices in textual segmentation.

It is not only in Translation Studies that editing is described in terms of actions. Early studies of edit distance by Damerau (1964) and Levenshtein (1966) estimate the fewest number of operations necessary to transform one segment into a different one. The purpose of those early estimates was to correct spelling errors in typed text or errors in computer code. This approach was later adopted for TER – Translation Edit Rate (Snover *et al.* 2006), one of the metrics most used by the MT community for tasks such as automatic PE and quality estimation, or to compare the quality of MT output from different systems. These metrics assume the shortest distance between an unedited and an edited string, but inevitably, translators do not take the shortest possible route from raw MT output to post-edited segment (as described in the section below on complex editing). Consequently, TER is not a fair description of what actually happens during the process, but more a description of translation products (Daems & Macken 2019, do Carmo 2017).

### *The editing threshold*

One of the purposes of quality estimation has been the identification of a threshold to enable the filtering out of MT segments that are of low quality and thus reduce productivity (Specia, Scarton & Paetzold 2018). This may be called the ‘editing threshold’, a boundary that sets the point from which a sentence requires much more writing than reading, a point from which editing becomes translating.

Let us say that the threshold is placed at a commonly-used lower limit for fuzzy matches in CAT such as 75%, which corresponds to an editing rate of 25% of all words in a sentence. In CAT, sentences with less than 75% matching words in the source language are still considered to require translating (and are thus paid at a full word rate), since it is unpredictable whether the edits in the target language will affect only this proportion of words. But let us concentrate on the level at which the TM does not provide any suggestions: the so-called ‘new words’ or ‘no matches’. This is the band in which MT output is considered to be most valuable. This output may be perfect and require no editing, but it may also include segments that need to be entirely rewritten, giving us an editing range that goes from 0 to 100%. The editing threshold we suggest re-establishes the CAT bands, but now in the context of PE by determining that, above 25% editing, the work should be considered as translating, implying a reading and writing effort that is above what is expected in simple editing.

An argument for the usefulness of the editing threshold is the fact that previous studies have shown particular behaviours and gains in productivity, not close to lower or upper bounds, but in medium quality or editing ranges (Krings 2001, Moorkens & Way 2016). Do Carmo (2017) carried out an experiment with 50 translators all working in the same language pair, using two work modes: one which relied on an auto-complete feature, and another involving an interface which constrained work to the four editing actions. In that experiment, the global average editing was also close to the threshold (26% for the autocomplete mode and 24% for the mode with the four editing actions). The study showed how detailed analyses of results show high variation in editing rates depending on text, user or mode of work.

Further studies of editing could focus on the editing threshold: the effort rates for segments above and below the threshold could be identified in fine-grained studies and the percentage of segments close to upper or lower bounds (with higher or lower quality) could be analysed. Naturally, any threshold will necessarily vary based on different factors, like language pair, TM or MT quality, domain, and project specifications. It would also be interesting to try to identify the threshold using typical translation process research methodologies: is it possible to see an effect on effort (be it technical, measured in keystroke frequency, or cognitive, measured by eye-tracking) when translators move from segments in which they are below the editing threshold to segments where they need to translate? All this research potential is a strong argument for making the editing threshold an object of study.

### *Types of editing actions*

To study editing, we need to have a clear description of the four actions that compose this task: deletion, insertion, movement and replacement. These four actions may be decomposed into two dimensions: the content of the textual units (words and groups of words) and the positions these units occupy in a segment. This helps to distinguish between primary and secondary actions and to develop better models for describing editing (do Carmo 2017).

The actions can be observed while the translator is editing or afterwards as TER comparisons between, for example, an MT suggestion (created by an MT system) and its edited version (created by the post-editor). Deletion and insertion are primary editing actions. This means that the position of the edited unit is occupied for only one of the elements of comparison: deleted units only appear in the MT suggestion; inserted units only appear in the post-edited version. In each case there is an empty position. These are pure cases of missing words, which makes them easy to identify in an alignment check (edit distance measures like TER start by aligning words and identifying deletions and insertions). Replacement and movement are secondary editing actions. This means that there is no empty position on either side: they represent manipulations of units which exist in both versions, either by replacing content in the same position or by maintaining the same content but moving it into a different position. Replacements are easy to identify, especially in sentences with the same numbers of words, but movements are very hard to identify (Snover *et al.* 2006).

Secondary actions can be decomposed into primary actions, as sequences of deletions followed by insertions: replacement is the deletion of a unit followed by the insertion of a different unit in the same position, whereas movement is the deletion of a unit in one position and its insertion in a different position. However, further analyses can be made with the four actions. We can, for example, divide the four actions into those that imply writing content (insertion and replacement) and those that do not affect the content, but just manipulate position (deletion and movement). Secondary actions are also associated with more efficient methods, like overtyping in the case of replacement, or dragging in the case of movement.

Using the editing actions as a guideline or an instrument of analysis can be very helpful. For example, descriptions of the translation process (usually at the character-level, with many

recursions and details that do not survive in the final version) could be more interpretable if they followed this model. In addition, since efficiency is a fundamental feature of translation, this simplified view of editing could become an important consideration in translator training (do Carmo 2017).

### *Complex editing*

We should stress that any description of editing consisting merely of a sequence of edits is a simplification of a complex process. When editing and, most relevantly, when translating, translators manipulate segmentation: for example, they select a word, but replace it with a phrase or a clause; they make non-linear edits when they apply scattered actions, moving in both directions within a sentence; they make recursive edits, coming back to the same word several times; they replace a word which may be embedded in a larger change within a group of words; and they make partial edits, when only part of a word is replaced, or discontinuous ones, in which, for example, an 's' for plural is added to different words in the same sentence.

Many changes may not be visible in the final result because of backtracks, and only the final edits survive in the translated and revised target text. This complicates the study of editing behaviour, and it is another argument in favour of the editing threshold: we need to partition editing data according to varying degrees of complexity. But even the editing threshold is difficult to measure if we use tools like TER, which estimates the minimum number of edits from the starting text to the final edited text, when the actual number of edits and keystrokes is inevitably greater.

### *From description to prediction of editing behaviour*

In "Towards statistical modelling of translators' activity data", Carl & Jakobsen (2009) explain their model for the analysis of typing and reading behaviour, using a method that includes product and process data. On the process side, units of textual data are collected by keystroke logs, essentially composed of deletions and insertions at the character level. They mention that cutting and pasting operations increase the complexity of the analysis, and add that the smaller the units of analysis, the easier it is to capture the actions. Movement is an operation that affects longer sequences, but it was not recorded during their experiments. They also observe that some of the keyboard activities are not linked to a particular word. The main conclusion from their data collection is that the difficulties in capturing a description of actions are related to the lack of alignment with source text words and to recursive operations applied to the same units.

In the discussion section of their paper, the authors comment on how best to assist human translation processes with automated tools: "At what moment during the translation would the mechanical help be most welcome? Would a translator be better supported during the 'linear' translation production or during the translation pauses?" (Carl & Jakobsen 2009: 136). They then discuss the distracting impact of typing suggestions and how to integrate these and other aids into translation tools. They advocate that process analysis may help to identify reading patterns and to develop tools even for the reading task.

By 2016, a few signs seemed to indicate that technology and research might have reached the evolutionary state required to pursue the goal of predicting editing actions:

*We are now at a stage in the development where translation research becomes predictive. The records from keylogging software and eye-trackers make it possible to address Holmes' (1972) second main objective, to "explain and predict" translators' behaviour: at present, we have all the necessary tools to address the challenge of building a model of human translation which makes specific, falsifiable predictions regarding the process and the product of translation.*  
(Carl, Bangalore & Schaeffer 2016:4)

Predicting the editing actions that translators should perform is one of the toughest challenges for the development of tools that specifically support PE. This is the theme of the next section of this chapter.

## **5. The need for interactive editing tools**

As we have seen, PE has become an accepted task in the industry (albeit one that is cognitively demanding and sometimes contentious on the translator side). Feeding translators MT output is now relatively straightforward, but applying the predictive power of MT to support the work of translators is more challenging.

When translators approach sentences that were produced by a high-quality MT system, they hope to just scan through them, not really reading from beginning to end, but merely looking for the errors to correct. In doing this, they are pressured by opposing forces from externally imposed productivity goals and their learned (and required) focus on small quality details.

Having negotiated this conflict and identified the errors, they decide how to correct them. They click on, or navigate to, scattered words in the text, and apply different techniques to correct errors with surgical precision, based on detailed keyboard and mouse actions applied to the words selected. Alternatively, they simply delete the whole sentence and rewrite it when they consider this the most efficient process, even if it means retyping a few words that were in the original sentence. That was the conclusion of a study which compared two modes of editing: traditional PE (the system presents one full MT suggestion and there is no interaction) and Interactive Machine Translation (IMT), where the translator writes and the system presents the next word (Alabau *et al.* 2016). In that study, almost all users preferred not to work with such interactive help. This shows that, although the technology is available to model and predict PE work, the question of how to offer the help in a useful and usable way remains open.

There is one very important difference between writing a translation and editing it: if you are writing a translation from start-to-finish, the whole translation is being generated in your brain even if your tool presents you with suggestions for word or sentence completion. Generating means first creating an abstract notion of the meaning and intention of a sentence, and only then, through syntactic processes, giving it form. When this generation process is confronted with a sequence of ever-changing suggested completions for the sentence, a high cognitive load is created by the dynamics of the process (Alabau *et al.* 2016).

For editing, the generation process should not be triggered, since the user is looking for mistakes. To actually be ‘editable’, the sentence presented by the MT system must be good enough for the translator to only worry about certain points that may be corrected through the application of well-directed actions. When that cannot be done, the generation process is triggered. At that moment, the translator decides to delete everything and this becomes a writing task, which moves us to the level in which editing becomes translating.

Having a good MT suggestion, however, is not enough to support editing. The interface elements – the mechanisms that build the communication between the MT systems and human actions – must also provide the necessary conditions for the editing to proceed in an efficient way. IMT may not be the best model to support editing since it is inspired by translating.

### *Open challenges for editing tools*

Several research projects have delved into the challenges of the interactivity required to reuse MT content in a way that supports translators through consideration of recommendations from fields such as Human-Computer Interaction (O’Brien 2012). Some of these projects explore mobile interfaces (O’Brien, Moorkens & Vreeke 2014), while others consider how to learn from post-edits (Simianer, Karimova & Riezler 2016). Interactive forms of PE have been studied within the CasMaCat project (Koehn, Saint-Amand & Alabau 2015). Techniques like Online Learning (to learn from user actions and incrementally adapt translation suggestions) and Active Learning (to determine which sentences need to be edited first so as to adapt the translation models most efficiently) have been applied to the task (Ortiz-Martínez *et al.* 2016). A thorough analysis of the development of interactive systems that apply both MT and TM in complex CAT environments has also been made (Forcada & Sánchez-Martínez 2015). However, this wealth of resources is often associated with an increased cognitive load (Christensen 2011, Pym 2011b). A more recent study compares statistical MT and NMT interactive systems (Daems & Macken 2019), and another (Coppers *et al.* 2018) tries to combine several resources in an interface that is intended to be intelligible and practical for users.

Translation throughput is nowadays often only limited by the capacity of the interfaces between humans and technology, as is recognised by analyses of typing speed, mobile input interfaces and voice recognition (Moorkens, O’Brien & Vreeke 2016). Furthermore, technological development must be based on correct models of processes, at the risk of not being useful (in the sense that they solve problems) or usable (in the sense that they improve processes) (Rabadán 2010).

The main conceptual model of how translators produce translations is ‘translating’: translators write full sentences, typing characters in a linear way. Therefore, auto-completion features have been the basis for the features that support writing in interactive translation systems – as translators start typing, words or full sentences are suggested to them (Green *et al.* 2014, Hokamp & Liu 2015). However, while auto-completion supports linear writing, editing is scattered and implies other actions, some of which involve only the position of words. These actions cannot be adequately supported by auto-completion systems and would benefit from

systems that were able to predict which words may require deletion, or in which positions they should be placed (do Carmo 2017).

An interactive tool that supports editing should present suggestions, depending not only on the words selected, but also on a contextual choice from the editing actions available: if there is a learnt model that estimates a high probability that a certain word is to be deleted, the tool may suggest that action beforehand. The interactive tool may use predictive writing functionalities to support the insertion and replacement actions, but it should also incorporate adaptive features to support frequent deletions and movements.

The challenge posed by editing is, as we have seen, not just a technical matter. Having tools that support decisions about which words to delete, or tools that present alternatives to replace words, affects not just the daily life of translators but our own theories and conceptions of the processes involved, as well as our overall perception of the value of tasks performed by professional translators.

## **6. Conclusions**

Throughout this chapter, we have reviewed descriptions and analyses of the editing task (when translators make small changes to text) that enable a clear consideration of the PE process and support our assertion of PE as a translation task. We have discussed how the view of PE as a form of revision may have contributed to devaluing this process, mostly in the professional world, but also affecting pedagogical approaches.

As an argument to counter that view, we propose a threshold that, even if it is only based on a measure of post-editors' technical work, serves to highlight the level of writing and translating that is required during PE. As we have seen, the four editing actions, which are at the centre of our view of editing, appear in very early studies of the translation process such as those by Nida and Toury, but they also appear very early in Computer Science in the form of edit distance metrics. Although apparently simple, these four actions may be fundamental to supporting descriptions of the complexity of translating, revising, and post-editing, and to fostering new methods of studying the details of production behaviours.

Throughout the chapter, there are several indications that PE is a specialised process which should be carried out by specialised translators. We have shown how editing is part of a creation process which has to be efficient and achieve different levels of quality. During PE, translators need to read more content than in normal revision, and they need to write in a more varied way, most frequently editing, but with the ability to quickly decide to delete an MT suggestion and translate the sentence from scratch. Post-editors need to have more strategies than mere transference; they need to know how to avoid replicating the source content in a structure that is inappropriate in the target language. Knowledge and practice of the four editing actions may help translators become more efficient at editing, but this must be done with the awareness that they can move above the threshold, into translating. Finally, the evolution of translation tools, offering more interactivity and support, requires users who are proficient at using these features. That way, the tools and resources at their disposal become instruments that sustain their professional development.

Through the study of editing, we have offered a view in which PE may be a more rewarding process for translators, a process in which optimisers and satisficers alike can identify their role in a demanding professional environment.

## 7. References

- Alabau, Vicent *et al.* (2016) 'Learning Advanced Post-Editing' in *New Directions in Empirical Translation Process Research* (Michael Carl, Srinivas Bangalore and Moritz Schaeffer eds.), Heidelberg: Springer International Publishing Switzerland, 95–110.
- Allen, Jeffrey (2003) 'Post-Editing' in *Computers and Translation* (Harold Somers ed.), Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Alves, Fábio *et al.* (2010) 'Translation Units and Grammatical Shifts: Towards an Integration of Product and Process-Based Translation Research' in *Translation and Cognition* (Gregory Shreve and Erik Angelone, eds.), Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Angelone, Erik (2010) 'Uncertainty, Uncertainty Management, and Metacognitive Problem Solving in the Translation Task' in *Translation and Cognition* (Gregory Shreve and Erik Angelone eds.), Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Austermühl, Frank (2013) 'Future (and Not-so-Future) Trends in the Teaching of Translation Technology'. *Tradumàtica: tecnologies de la traduciò* 11: 326–37.
- Cadwell, Patrick, Sharon O'Brien and Carlos S. C. Teixeira (2018) 'Resistance and Accommodation: Factors for the (Non-) Adoption of Machine Translation among Professional Translators' *Perspectives: Studies in Translatology* 26:3, 301-321.
- Carl, Michael, Srinivas Bangalore and Moritz Schaeffer (eds.) (2016) *New Directions in Empirical Translation Process Research*. Heidelberg: Springer International Publishing Switzerland.
- Carl, Michael, Barbara Dragsted and Arnt Lykke Jakobsen (2011) 'On the Systematicity of Human Translation Processes' *Tralogy 2011. Translation Careers and Technologies: Convergence Points for the Future*. Paris, France.
- Carl, Michael and Arnt Lykke Jakobsen (2009) 'Towards Statistical Modelling of Translators' Activity Data' *International Journal of Speech Technology* 12:4, 125–38.
- do Carmo, Félix (2017) 'Post-Editing: A Theoretical and Practical Challenge for Translation Studies and Machine Learning'. PhD dissertation, Universidade do Porto. <https://repositorio-aberto.up.pt/handle/10216/107518> [consulted 31 August 2019].
- Castilho, Sheila *et al.* (2017) 'Is Neural Machine Translation the New State of the Art?' *The Prague Bulletin of Mathematical Linguistics* 108:1.
- Catford, J. C. (1965) *A Linguistic Theory of Translation*. London: Oxford University Press.
- Christensen, Tina Paulsen (2011) 'Studies on the Mental Processes in Translation Memory-Assisted Translation - the State of the Art' *Trans-kom* 4: 137–60.
- Coppers, Sven *et al.* (2018) 'Intellingo : An Intelligible Translation Environment' in *Proceedings of CHI*, Montreal, Canada, 1–13.
- Daems, Joke and Lieve Macken (2019) 'Interactive Adaptive SMT versus Interactive Adaptive NMT: A User Experience Evaluation' *Machine Translation* 33:135

- Damerau, Fred J. (1964) 'A Technique for Computer Detection and Correction of Spelling Errors' *Communications of the ACM* 7:3, 171–76.
- Dorr, Bonnie (editor) (2010) 'Part 5: Machine Translation Evaluation' in *Handbook of Natural Language Processing and Machine Translation*, (Joseph Olive, Caitlin Christianson and John McCary eds.). Springer Verlag, 801–94.
- European Master's in Translation Expert group (2017) 'European Master's in Translation - Competence Framework 2017'  
[https://ec.europa.eu/info/sites/info/files/emt\\_competence\\_fwk\\_2017\\_en\\_web.pdf](https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_en_web.pdf) [consulted 31 August 2019].
- Forcada, Mikel L. and Felipe Sánchez-Martínez (2015) 'A General Framework for Minimizing Translation Effort: Towards a Principled Combination of Translation Technologies in Computer-Aided Translation' *Proceedings of the 18th Annual Conference of the European Association for Machine Translation*: 27–34.
- Green, Spence, Jason Chuang, Jeffrey Heer and Christopher D Manning (2014) 'Predictive Translation Memory: A Mixed-Initiative System for Human Language Translation' in *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*, New York: ACM.
- Guerberof Arenas, Ana and Joss Moorkens (2019) 'Machine Translation and Post-Editing Training as Part of a Master's Programme' *The Journal of Specialised Translation* 31: 217–38.
- Halliday, M.A.K. and J.R. Martin (1993) *Writing Science: Literacy and Discursive Power*. London: The Falmer Press.
- Hokamp, Chris and Qun Liu (2015) 'Handycat: The Flexible CAT Tool for Translation Research' in *EAMT 2015*, Istanbul.
- House, Julianne (2003) 'English as Lingua Franca and Its Influence on Discourse Norms in Other Languages' in *Translation Today: Trends and Perspectives* (Gunilla Anderman and Margaret Rogers eds.) Clevedon: Multilingual Matters, 168.179.
- ISO. 2017. 'ISO 18587:2017 - Translation Services - Post-Editing of Machine Translation Output - Requirements' 1–15.
- Kenny, Dorothy (2018) 'Sustaining Disruption? On the Transition from Statistical to Neural Machine Translation' *Tradumàtica: tecnologies de la traducció* 16: 59.
- Kenny, Dorothy and Stephen Doherty (2014) 'Statistical Machine Translation in the Translation Curriculum: Overcoming Obstacles and Empowering Translators' *Interpreter and Translator Trainer* 8:2, 276–94.
- Koehn, Philipp (2009) 'A Process Study of Computer-Aided Translation' *Machine Translation* 23:4, 241–63.
- Koehn, Philipp, Herve Saint-Amand and Vicent Alabau (2015) 'Final Report on Interactive Editing'  
<https://cordis.europa.eu/docs/projects/cnect/6/287576/080/deliverables/001-deliverable33Ares20152495853.pdf> [consulted 31 August 2019].
- Krings, Hans P. (2001) *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes*, (Geoffrey S. Koby ed.). Kent, Ohio & London: The Kent State University Press.

- Läubli, Samuel *et al.* (2013) 'Assessing Post-Editing Efficiency in a Realistic Translation Environment' in *MT Summit XIV Workshop on Post-Editing Technology and Practice*, 83–91.
- Levenshtein, Vladimir I. (1966) 'Binary Codes Capable of Correcting Deletions, Insertions and Reversals.' *Soviet Physics Doklady* 10:8, 707–710.
- Lommel, Arle (2018) 'Where's My Translation Jet Pack?' in *Proceedings of Translating and the Computer 40*, London: Asling.
- Lommel, Arle and Donald Depalma (2016) 'Europe's Leading Role in Machine Translation - How Europe Is Driving the Shift to MT'. [http://cracker-project.eu/wp-content/uploads/Europes\\_Leading\\_Role\\_in\\_MT.pdf](http://cracker-project.eu/wp-content/uploads/Europes_Leading_Role_in_MT.pdf) [consulted 31 August 2019].
- Moorkens, Joss and Sharon O'Brien (2017) 'Assessing User Interface Needs of Post-Editors of Machine Translation Interfaces for Editing Human Translation and Post-Editing Machine' in *Human Issues in Translation Technology* (Dorothy Kenny and Jenny Williams eds.) London: Routledge.
- Moorkens, Joss, Sharon O'Brien and Joris Vreeke (2016) 'Developing and Testing Kanjingo: A Mobile App for Post-Editing' *Tradumàtica: Tecnologies de la traducció* 14: 58–66.
- Moorkens, Joss and Andy Way (2016) 'Comparing Translator Acceptability of TM and SMT Outputs' *Baltic Journal of Modern Computing Modern Computing* 4:2, 141–51.
- Mossop, Brian (2020) *Revising and Editing for Translators*. (4th ed.) London & New York: Routledge.
- Nida, Eugene Albert (1964) *Toward a Science of Translating*. Leiden: E. J. Brill.
- O'Brien, Sharon (2012) 'Translation as Human Computer Interaction' *Translation Spaces* 1:1, 101–22.
- O'Brien, Sharon, Joss Moorkens and Joris Vreeke (2014) 'Kanjingo: A Mobile App for Post-Editing.' *Third Workshop on Post-Editing Technology and Practice*: 125.
- Ortiz-Martínez, Daniel *et al.* (2016) 'Integrating Online and Active Learning in a Computer-Assisted Translation Workbench' in *New Directions in Empirical Translation Process Research* (Michael Carl, Srinivas Bangalore and Moritz Schaeffer eds.), Heidelberg: Springer International Publishing Switzerland 57-76.
- Plitt, Mirko and François Masselot (2010) 'A Productivity Test of Statistical Machine Translation Post-Editing in a Typical Localisation Context' *The Prague Bulletin of Mathematical Linguistics* January 20:93, 70–99.
- Pym, Anthony (2011a) 'Translation Research Terms: A Tentative Glossary for Moments of Perplexity and Dispute' in *Translation Research Projects* 3:3, 75–99.
- (2011b) 'What Technology Does to Translating' in *Translation and Interpreting* 3:1, 1–9.
- Rabadán, Rosa (2010) 'Applied Translations Studies' in *Handbook of Translation Studies - Volume I*, (Yves Gambier and L. van Doorslaer eds.). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sager, Juan (1993) *Language Engineering and Translation: Consequences of Automation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Sánchez-Gijón, Pilar, Joss Moorkens and Andy Way (2019) 'Post-Editing Neural Machine Translation versus Translation Memory Segments' *Machine Translation* 33:1–2, 31–59.

- Sanchis-Trilles, Germán et al (2014) 'Interactive Translation Prediction versus Conventional Post-Editing in Practice: A Study with the CasMaCat Workbench' *Machine Translation* 28:3–4, 217–35.
- Schwartz, Lane (2014) 'Monolingual Post-Editing by a Domain Expert Is Highly Effective for Translation Triage.' *Proceedings of the Third workshop on post-editing technology and practice (WPTP-3)* 34–44.
- Simianer, Patrick, Sariya Karimova and Stefan Riezler (2016) 'A Post-Editing Interface for Immediate Adaptation in Statistical Machine Translation.' *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: System Demonstrations*: 16–20.
- Snover, Matthew et al (2006) 'A Study of Translation Edit Rate with Targeted Human Annotation.' *Proceedings of AMTA 2006* August: 223–31.
- Specia, Lúcia, Carolina Scarton and Gustavo Henrique Paetzold (2018) *Quality Estimation for Machine Translation*. Morgan & Claypool.
- TAUS. 2010. 'Post-Editing Guidelines' <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines> [consulted 31 August 2019].
- Toury, Gideon (1995) *Descriptive Translation Studies – and Beyond*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Vasconcellos, Muriel (1987) 'Post-Editing on-Screen: Machine Translation from Spanish into English' in *Proceedings of the Conference Translating and the Computer 8*, (Catriona Picken ed.). London: Aslib, 133–46.
- Wagner, Emma (1985) 'Post-Editing Systran - A Challenge for Commission Translators' *Terminologie et Traduction* 3.
- Zaretskaya, Anna (2019) 'Raising the TM Threshold in Neural MT Post-Editing: A Case-Study on Two Datasets' in *Proceedings of MT Summit XVII Vol. 2*, 213–218.
- Zhechev, Ventsislav (2014) 'Analysing the Post-Editing of Machine Translation at Autodesk' in *Post-Editing of Machine Translation: Processes and Applications* (Sharon O'Brien et al. eds.) Newcastle upon Tyne: Cambridge Scholars Publishing.