Machine-Assisted Translation of Literary Text: A Case Study

Antonio Toral
ADAPT Centre
School of Computing
Dublin City University
Dublin 9, Ireland
atoral@computing.dcu.ie

Andy Way
ADAPT Centre
School of Computing
Dublin City University
Dublin 9, Ireland
away@computing.dcu.ie

Abstract

Contrary to perceived wisdom, we explore the role of machine translation (MT) in assisting with the translation of literary texts, considering both its limitations and its potential. Our motivations to explore this subject are twofold: (i) the recent research advances in MT, and (ii) the recent emergence of the ebook, which together allow us for the first time to build literature-specific MT systems by training statistical MT models on novels and their professional translations. A key challenge in literary translation is that one needs to preserve not only the meaning (as in other domains such as technical translation) but also the reading experience, so a literary translator needs to carefully select from the possible translation options. We explore the role of translation options in literary translation, especially in the context of the relatedness of the languages involved. We take Camus' *L'Étranger* in the original French language and provide qualitative and quantitative analyses for its translations into English (a less-related language) and Italian (more closely related). Unsurprisingly, the MT output for Italian seems more straightforward to be post-edited. We also show that the performance of MT has improved over the last two years for this book, and that the applicability of MT does not only depend on the text to be translated but also on the type of translation that we are trying to produce. We then translate a novel from Spanish-to-Catalan with a literature-specific MT system. We assess the potential of this approach by discussing the translation quality of several representative passages.

keywords: machine translation of literary text; literary translation; machine translation; statistical machine translation; human translation.

1. Introduction

The central contention in this paper is that machine translation (MT), in particular statistical MT (SMT), can be useful for the translation of literary works, especially novels. At first glance, this appears to fly in the face of the perceived wisdom in the field:

> *"Taking literary translation as the sole object of translation studies skews all arguments about interlingual communication from the start … That's not what literary translation is about. For works that are truly original — and therefore worth translating — statistical machine translation hasn't got a hope. Google Translate can provide stupendous services in many domains, but it is not set up to interpret or make readable work that is not routine — and it is unfair to ask it to try. After all, when it comes to the real challenges of literary translation, human beings have a hard time of it, too."* (Bellos 2012)

Note also that literary translation is often selected by human translators not overly well-disposed to MT to demonstrate how useless it is for anything; on any randomly selected online translators' forum, you don't have to look too hard to find someone who has selected a section from a book and shown how MT messes up the translation. In reviewing Bellos (2012), Way (2012) suggests reworking the first part of the above quote as follows: "Taking literary translation as the sole object of MT skews all arguments about its potential to help with interlingual communication from the start".

However, despite all this negativity, we contend that MT has the potential to be useful for literary translation; at the very least, the perceived wisdom that it has no hope whatsoever of helping human translators create translations of novels should not be accepted at face value. Accordingly, we propose that a thorough investigation of its utility in this space is warranted, both from the point of qualitative and quantitative evaluations.

At the outset, we agree with Bellos that literary translation is perhaps the hardest task for human translators. Perhaps surprisingly, then, it is worth pointing out that human translators are *extremely* poorly paid for their work on translating literary texts (Kelly and Zetsche 2012, 94). Accordingly, if SMT *could* be shown to be of use in this domain, then MT would be capable of helping these poorly paid literary translators to become more efficient, and make more money.

Despite protestations to the contrary (cf. Penkale and Way (2013) and Way (2013) for a selection of quotes from some particularly forthright scaremongerers regarding MT, as well as many more reasonable views), MT has been used in other professional domains to good effect, especially as a productivity enhancer in the localisation sector. In this regard, Way (2013) notes that there is a large body of evidence pointing to the fact that "the time for MT is now". He observes that:

> *"MT quality is now good enough that millions of people are using it every day to satisfy their requirements. At one end of the spectrum, there are freely available web-based tools such as Google Translate[1] and Bing Translator,[2] which provide strong baseline performance especially given the need to be robust enough to cope with any input."*

[1] http://translate.google.com/

He also provides a number of recent successful use-cases using a range of MT providers for different clients. Accordingly, despite a number of translators resisting the advent of MT, Way avers that "the point of questioning whether MT is useful or not is moot."

That said, it has to be admitted that the domain of application of MT that we investigate in Toral and Way (2015) and in more depth in this paper is quite different in nature from those in which human post-editors normally operate. Way (2013) provides three sets of use-cases where MT has been demonstrated to be effective: for raw MT, for lightly post-edited MT (PEMT), and for full PEMT. Way provides examples of these, including:

- Raw MT: user-generated content, multilingual search, sentiment analysis, real-time translation, forensic investigation, basic product information
- Light PEMT: manuals (with little security or health & safety risks), online help, product support, market research
- Full PEMT: manuals (security/health & safety to be considered), contracts, patents

For all these cases, what is crucially important is rendering the meaning of the source text in the target language. However, when it comes to literary translation, it can be convincingly argued that one critical objective of translation is to preserve the experience of reading the text.

We believe it is timely to explore the applicability of MT to literary text, due to the research maturity and industrial adoption reached by MT and also due to the emergence of the ebook, which allows us to build literary-specific SMT systems trained on novels and their translations. This exploration is relevant both from research and societal points of view:

- Research. MT lags behind state-of-the-art theories in Translation Studies. This field moved decades ago from formal equivalence, where the aim of the translation is to replicate the form of the source text, to dynamic equivalence, where the equivalence is sought at functional and pragmatic levels (Nida and Taber 1969), with more recent theories moving even further away from formal equivalence (Snell-Hornby 1995). Meanwhile, the vast majority of research in MT disregards functional and pragmatic aspects and aims to model – somehow – formal equivalence. It is no wonder then that the biggest success of MT to date is its application to technical documents, where the primary function of the translation is informative, and thus formal equivalence suffices. The challenge with translating literature is that the primary function of its translation is expressive, the aim of the translation being to replicate the source text's effects on the reader. By broadening the application of MT to literary text we ultimately aim to bring the dynamic equivalence theory to the field of MT, thus narrowing the theoretical gap between the fields of MT and Translation Studies.
- Society. Translation of literary texts is a costly task both in terms of time and money, but at the same time it is crucial for literary exchange across different linguistic and cultural communities. Successful application of MT to literary text would then foster such an

[2] http://www.bing.com/translator

exchange, especially for communities of minority languages for which translations have heretofore been rather limited. Furthermore, MT could work as an accelerator allowing novels to be translated in shorter terms (cf. *simship* in localisation workflows).

The remainder of this paper is organised as follows. Section 2 provides an introduction to SMT followed by an overview of previous work on MT for literary texts. Section 3 motivates the current opportunity of applying MT to literary texts. Section 4 explores the challenges that translation options in literary translation present to out-of-the-box MT systems. Section 5 delves into the applicability of literary-adapted MT for related languages. Finally, Section 6 outlines our conclusions and lines of future work.

2. Related Work

In this section, we provide an overview of how SMT models of translation are built, and how they are used in practice. We also present previous efforts at using MT for different areas of literary translation, and compare them to the MT-assisted translation workflow that we envisage.

2.1. Statistical Machine Translation

There are two main processes in statistical models of translation: training and decoding. All training is offline, and involves the computation of three models: a translation model, a reordering model, and a target language model (LM). The first two models are generated from parallel data (such as the contents of a Translation Memory (TM)), while the LM is derived from a large collection of monolingual data.[3]

Hearne and Way (2011) explain that in the original 'noisy channel' version of SMT (Brown et al. 1990, 1993), only the translation model and LM played a role:

> *"The translation model effectively comprises a bilingual dictionary where each possible translation for a given source word or phrase has a probability associated with it. However, the model does not resemble a conventional dictionary where plausible entries only are permitted; many of the entries represent translations that are unlikely but not impossible, and the associated probabilities reflect this. The language model comprises a database of target-language word sequences (usually ranging between 1 and 7 words in length), each of which is also associated with a probability*."

In general terms, the translation model computes a very large number of inferences based on observations in the parallel data and stores these in its phrase-table. At runtime, a set of target-language words and phrases are proposed which are optimal for the translation of the particular source-language sentence at hand. In contrast, the LM computes a very large number of inferences based on observations in the monolingual data. At runtime, it takes the suggested target words and phrases from the translation model and tries to assemble them into the best target-language word order. In cooperating in this way during the decoding (or 'search') phase,

[3] Way and Hearne (2011) demonstrate further that the traditionally understood terms of 'adequacy' and 'fluency' can be applied to the translation and language models,respectively.

the most likely translation of the source sentence is calculated from potentially millions of possible target candidates from a purely *mathematical* point of view. In this regard, Way and Hearne (2011) make the following useful observation:

> *"the methods used ... are not intended to be either linguistically or cognitively plausible (just **probabilistically** plausible), and holding onto the notion that they somehow are or should be simply hinders understanding of SMT."* (original emphasis)

Over the last ten years or so, the noisy channel model of SMT has been supplanted by the log-linear model of SMT (Och and Ney 2002), whereby other components – including a reordering model – can be combined with the language and translation models to improve translation quality. Each of these components (or 'features') is assigned a weight in the 'parameter estimation' (or 'tuning') phase so that the highest score according to a particular automatic evaluation metric (e.g. BLEU, Papineni et al. (2002)) is obtained on a held-out tuning set.

For both models of SMT, and for each of these components, we refer the interested reader to Hearne and Way (2011) in the first instance, and to the primary sources cited herein for the more intrepid.

2.2. MT of Literary Text

There has been recent interest in the Computational Linguistics community regarding the processing of literary text. The best example of this is the establishment of an annual workshop on the topic of Computational Linguistics for Literature since 2012.[4] A popular strand of research in this area has to do with the automatic identification of text snippets that convey figurative devices, such as metaphor (e.g. Shutova et al. 2015), idioms (e.g. Li and Sporleder 2010), humour and irony (e.g. Reyes 2012). All these works apply to monolingual text. To date, there has been only a very limited amount of work on applying MT to literature, as we now survey in detail.

Genzel et al. (2010) explored constraining statistical MT (SMT) systems for poetry to produce translations that obey particular length, meter and rhyming rules. Form is preserved at the cost of producing a lower quality translation, in terms of BLEU, the most widely used automatic evaluation metric in MT, which decreases from 35.3 to 17.3, a drop of around 50% in real terms. It should be noted that their evaluation was not on poetry but on news, i.e. they produced translations of news that obeyed length, meter and rhyming rules. The language pair was French–English.

Greene et al. (2010) also translated poetry, choosing target realisations that conform to the desired rhythmic patterns. Specifically, they translated Dante's *Divine Comedy* from Italian sonnets into English iambic pentameter. Instead of constraining the SMT system, they passed its output lattice through a device that maps words to sequences of stressed and unstressed syllables. These sequences are finally filtered with an iambic pentameter acceptor. Their output translations are evaluated qualitatively only.

---

4 https://sites.google.com/site/clfl2015/previous-workshops

Voigt and Jurafsky (2012) examined how referential cohesion is expressed in literary and non-literary texts, and how this cohesion affects translation. They found that literary texts have more dense reference chains and conclude that incorporating discourse features beyond the level of the sentence (Hardmeier 2014; Meyer 2014) is an important direction for applying MT to literary texts.

Way (2013) presents a list of recent successful case-studies, one of which includes the translation of religious texts from The Church of Jesus Christ of Latter-day Saints using the Microsoft Translator Hub[5] (Richardson 2012). While the underlying technology behind the Microsoft Translator Hub is based very much on SMT, it is a new 'DIY' system where users upload their own translation assets (parallel text, monolingual text, glossaries etc) and the system is built automatically in quite a short period of time.

Jones and Irvine (2013) used existing MT systems to translate samples of French literature (prose and poetry) into English. They then used qualitative analysis grounded in translation theory on the MT output to assess the potential of MT in literary translation and to address what makes literary translation particularly difficult.

Besacier (2014) presented a pilot study where MT followed by post-editing is applied to translate a short story from English into French. In Besacier's work, post-editing is performed by non-professional translators, and the author concludes that such a workflow can be a useful low-cost alternative for translating literary works, albeit at the expense of sacrificing translation quality. According to the opinion of a professional translator, the main errors had to do with using English syntactic structures and expressions instead of their French equivalents and not taking into account certain cultural references.

Finally, our recent work has explored the hypothesis of whether MT can be useful to translate literary texts in a position paper (Toral and Way 2014). A follow-up exploratory experiment provided some preliminary evidence that MT could be useful in this regard (Toral and Way 2015). In this experiment we built a tailored MT system for a contemporary best-selling author (Carlos Ruiz Zafón)1 and then applied it to translate one of his novels between two closely-related languages (Spanish to Catalan). We discovered that for 20% of the sentences, the translations produced by the MT system and the professional translator (i.e. taken from the published novel in the target language) were identical. In addition, a human evaluation revealed that for over 60% of the sentences, native speakers noted the translations produced by MT and by the professional translator to be of the same quality.Our work contributed to the state-of-the-art in two dimensions. On the one hand, we conducted a comparative analysis on the translatability of literary text according to narrowness of the domain and freedom of translation. This can be seen as a more general and complementary analysis to the one conducted by Voigt and Jurafsky (2012). On the other hand, and related to Besacier (2014), we evaluated MT output for literary text. There are two differences though; first, Besacier translated a short story, while we have done so for a longer type of literary text, namely a novel; second, his MT systems were

5 https://hub.microsofttranslator.com/

evaluated against a post-edited reference produced by non-professional translators, while we have evaluated our MT systems against the translation produced by a professional translator.

Our recent results constitute a promising first step since they question the perceived wisdom that MT is of no use for translating literature, at least for closely-related languages. That said, as these results are preliminary for this research topic, they are of course somewhat limited as we dealt only with one novel (it may be that results depend to a large extent on the novel style, genre, etc.), the evaluation was conducted at sentence level, and, most of all, it is evident that MT between related languages leads to better results than between unrelated languages.

From this survey of the state-of-the-art we can conclude that the applicability of MT to literature from an empirical point of view is in its infancy. We argue that the line of research started in our contribution to the state-of-the-art, that we pursue further in this paper, is especially ambitious and relevant since we are the first to (i) build MT systems adapted to the writing and translation styles of novels, and (ii) evaluate their translation outputs against their professionally generated translations.

3. Opportunity is ripe for the use of MT for Literary Translation

We argue that the quest to study the applicability of MT to literary texts is timely. The recent emergence of the ebook is dramatically changing the book publishing industry. The ebook has reduced two of the most important costs of publishing books – distribution and printing – to the extent that they become almost negligible. This cost reduction, of course, applies also to publishing translations of books, for which two main costs remain: (i) publishing rights (the fee paid by the company publishing the translation to the publisher in the original language and/or to the author), and (ii) the translation itself. With distribution and printing costs gone, these two costs become the bottleneck for publishing translations. While publication rights are clearly out of our control, we argue that the use of MT can reduce translation costs, and that this reduction should result in publishers being able to translate more books, thus benefiting (i) readers, who will be able to access a broader selection of translated books in their native language, and (ii) authors, who will reach readers from other linguistic communities.

As we saw in Section 2.1, the main resource required to build SMT systems is bilingual parallel text. Given the emergence of the ebook, books such as novels are now available in digital format. Accordingly, we are now able for the first time to build SMT systems trained on novels. In this paper we build MT systems tailored to the styles of literary authors and translators.

Finally, Kelly and Zetzsche (2012:93f.) observe that "literary translation is one of the most challenging types of translation work". That said, they note that "the person who translates the bestselling literary masterpieces would probably earn more working on a factory assembly line … There is very little glamour or money in literary translation, for all but a miniscule percentage of the pool". They also cite Martin de Haan, president of CEATL,[6] Europe's leading association for literary translation, who agrees that "most literary translators are on the verge of poverty … In some countries it is simply impossible to make a living as a professional literary translator".

[6] http://www.ceatl.eu/

To back this up, Kelly & Zetzsche quote from a CEATL study which showed that literary translators earned less than 50% of the per capita GDP. In other words, "the average earning power of a literary translator was inferior to the average wages in manufacturing and services in *every single country* analyzed. Indeed, in the vast majority of countries, translators earned less than 66% of this amount" (op cit, p.94, original emphasis).

In sum, there is clearly both a demand and a resource for MT in the area of literary translation. At the same time, we contend that as in other sectors, the availability of MT as a tool in the translator's armoury has the potential to increase remuneration for currently poorly-paid human translators of literary text.

4. Translation Options in Literary Translation

It is often argued that a key difference between literary translation and other types of translation is that "*how* one says something can be as important, sometimes more important, than what one says" (Landers 2001, 7, original emphasis), in other words, literary translation is not only about preserving meaning but also about preserving the reading experience. This is related to the claim that simple source-language phrases can be rendered in a variety of ways. For instance, Landers (*op cit.*) provides 12 possible translations for a source-language sentence as simple as the Portuguese "Nao vou lá". One source of multiple translation options comes from the open debate among translators on whether one should adapt the source text to the reader in the target language, also known as domesticating the text, or stay as faithful as possible to the original, also referred to as foreignising (Venuti 2008).

Given this myriad of translation options, it is then claimed that if two translators were to translate the same literary text, the translations would be substantially different. This has been claimed at the qualitative level. In this section we measure this at the quantitative level using two translations into English of Camus' *L'Étranger* (Camus 1942). The first was by a British translator (Camus and Gilbert 1946) and was read as the standard English translation for more than thirty years. The second (Camus and Ward 1989), was americanised due to the fact that Camus was influenced by the American literary style. Table 4.1 shows a passage of the novel and its translations by Gilbert and Ward.

French – original
Lui parti, j'ai retrouvé le calme. J'étais épuisé et je me suis jeté sur ma couchette. Je crois que j'ai dormi parce que je me suis réveillé avec des étoiles sur le visage. Des bruits de campagne montaient jusqu'à moi. Des odeurs de nuit, de terre et de sel rafraîchissaient mes tempes. La merveilleuse paix de cet été endormi entrait en moi comme une marée. A ce moment, et à la limite de la nuit, des sirènes ont hurlé. Elles annonçaient des départs pour un monde qui maintenant m'était à jamais indifférent. Pour la première fois depuis bien longtemps, j'ai pensé à maman.
English - translation by Gilbert
Once he'd gone, I felt calm again. But all this excitement had exhausted me and I dropped heavily on to my sleeping plank. I must have had a longish sleep, for, when I woke, the stars were shining down on my face. Sounds of the countryside came faintly in, and the cool night air, veined with smells' of earth and salt, fanned my cheeks. The marvelous peace of the sleepbound summer night flooded through me like a tide. Then, just on the edge of daybreak, I heard a steamer's siren. People were starting on a voyage to a world which had ceased to concern me forever. Almost for the first time in many months I thought of my mother.
English - translation by Ward
With him gone, I was able to calm down again. I was exhausted and threw myself on my bunk. I must have fallen asleep, because I woke up with the stars in my face. Sounds of the countryside were drifting in. Smells of night, earth, and salt air were cooling my temples. The wondrous peace of that sleeping summer flowed through me like a tide. Then, in the dark hour before dawn, sirens blasted. They were announcing departures for a world that now and forever meant nothing to me. For the first time in a long time I thought about Maman.

Table 4.1. A passage from *L'Étranger* by Camus, together with its translations in English by Gilbert and Ward.

Our methodology is as follows. First we sentence-align both translations. We then measure the overlap between those sentences with BLEU. Taking Gilbert's translation as the reference and Ward's as the output and vice versa, perhaps surprisingly the BLEU score (measuring word- and phrase-level overlaps between two texts) is only 18.5. To provide some insight into this, a maximum score of 100 would have meant that the two translations were identical to one another. To system developers, a BLEU score of less than 20 would be indicative of unusable quality in a post-editing workflow. In other words, Gilbert and Ward have translated Camus' work so differently as to render the two outputs incomparable.[7]

We accept there is a possibility that Ward may have deliberately avoided producing a translation close to Gilbert's, but it will become clear in our analysis that the output from an out-of-the-box MT system bears much greater resemblance to Ward's translation than to Gilbert's. Note that Bellos (2012: 266, Ch. 23) observes that "[Translators] behave more like GT [Google

[7] Note that this is not as bizarre an undertaking as might be imagined at first sight. Fancellu et al. (2014) provide a number of use-cases where same-language MT has real-world applications.

Translate]" themselves, which may to some extent explain this finding!

In the remainder of this section we present qualitative and quantitative analyses of MT for *L'Étranger*.

4.1. Qualitative Analysis

We take a passage previously studied by Jones and Irvine (2013), cf. Table 4.1. They selected this passage on the basis that it uses "fairly simple language" and corresponds to a "modern and well-known author".

We first provide an analysis of SMT progress. Contrary to widespread perception – in the translation industry at least – that SMT performance has stagnated, recent research has shown that the performance of MT systems has improved notably in the period 2007–2012.[8] These improvements were measured on newswire, so in what follows we assess the extent to which these improvements carry over to literary texts.

To that end we show in Table 4.2 the translation of the passage into English using Google Translate as it was back in 2013 (i.e. the translation shown in Jones and Irvine (2013)) and in its current status at the time of writing (June 2015). We consider as reference the professional translation by Ward.

[8] For example, Lønning et al. (2004) state that "although statistical approaches can deliver good initial results to MT, they seem to sooner or later suffer from 'ceiling' effects in performance". In contrast, Graham et al. (2014) looked at the best-performing systems of the WMT shared task for seven language pairs during this period, and found the improvement in translation quality during the period to be around 10% absolute, in terms of both adequacy and fluency.

English - Google Translate (2013)
He was gone, I found calm. I was exhausted and I threw myself on my bunk. I think I slept because I woke up with stars on her face. Noises campaign amounted to me. The smell of night, earth and salt refreshed my temples. Heavenly peace this summer sleeping entered me like a tide. At that time, and the limit of the night, sirens screamed . They announced departures for a world that now was never indifferent to me. For the first time in ages I thought mom.
English - Google Translate (June 2015)
He left, I returned to calm. I was exhausted and I threw myself on my bunk. I think I slept because I woke up with stars on the face. Campaign noises were up to me. Night smells of earth and salt were cooling my temples. The wonderful peace this summer asleep entered me like a tide. At that moment, and on the edge of the night, sirens howled. They announced departures for a world that now was never indifferent to me. For the first time in many years, I thought about Mom.

Table 4.2. Machine translations produced by Google Translate in 2013 and 2015 for the passage of *L'Étranger* shown in Table 4.1.

Jones and Irvine (2013) analysed this passage in terms of lexical variation and time as an aspect of translation. We now re-explore their criticisms of the output produced by the MT system in 2013, and analyse whether the translation produced by the MT system in 2015 yields any improvement:

- Line 1. Ward translates the "re" of "retrouvé" as "again". This part of the translation allows him to express the fact that the speaker is "calming himself after the departure of the warden". In the translation produced by the MT system, this nuance was lost. The newer MT system improves upon this by translating "retrouvé" into "returned to" instead of as "found".
- Line 3. The MT system introduced a wrong pronoun ("her"). The more recent MT system improves on this by rendering a correct translation ("the"), despite not be as appropriate as Ward's ("my").
- Line 4. "Des bruits" was wrongly translated as "noises" instead of "sounds". The newer MT system still translates this incorrectly, but at least improves on word order ("Campaign noises" instead of "Noises campaign").

Aside from the aspects analysed in Irvine and Jones (2013), there are other clear examples in these passages that indicate that SMT has improved in the last two years, namely:

- Regarding tense, "rafraîchissaient" was wrongly translated by the 2013 MT system in the simple past ("refreshed"), while the 2015 system matches the translation produced by Ward ("were cooling").
- As for lexical choice, let us comment on two cases where MT has improved. First, "cet été endormi" was translated by the 2013 system as "summer sleeping", which

improves to “summer asleep” with the 2015 system. Second, for the phrase “des sirènes ont hurlé”, the 2015 system produces a more literary-sounding “sirens howled” compared to the 2013 system’s “sirens screamed”.

- Finally, regarding the translation of particles, there are again two clear cases where MT shows improvement over the last two years. The first (“à la limite de la nuit”) was translated by the 2013 MT system as “the limit of the night”, whereas the 2015 system gives “on the edge of the night”. While the 2013 system dropped “à”, the 2015 system translates it as “on”. The second (“j’ai pensé à maman”) yields an ungrammatical translation with the 2013 system (“I thought mom”) in dropping “à”, while the translation is grammatical and fluent with the 2015 system (“I thought about Mom”); note too the correct casing here.

Hitherto, most literature on literary translation has focused solely on English as the target language. Conversely, in this work we contemplate other target languages in order to study the effect of language relatedness as regards the potential usefulness of MT. We hypothesise that MT will be more applicable to the translation of literary texts between languages that belong to the same family, as the number of potential translation options ought to be lower. Accordingly, we now look at the translation into Italian of the passage from *L’Étranger* shown above (Table 4.1). We give two translations in Table 4.3, a machine translation produced by Google Translate at the time of writing and the professional translation by Zevi (Camus and Zevi 1987).

Italian - Google Translate (June 2015)
Ha lasciato, sono tornato alla calma.
Ero esausto e mi sono buttato sulla mia cuccetta.
Credo di aver dormito perché mi sono svegliato con stelle sul viso.
Rumori della campagna sono stati fino a me.
Odori notturni di terra e di sale sono stati raffreddando le tempie.
La pace meraviglioso questa estate addormentato mi è entrato come una marea.
In quel momento, e sul bordo della notte, sirene ululavano.
Hanno annunciato partenze per un mondo che ormai era mai indifferente.
Per la prima volta in molti anni, ho pensato a mamma.
Italian - Translation by Zevi
Partito lui, ho ritrovato la calma.
Ero esausto e mi sono gettato sulla branda.
Devo aver dormito perché mi sono svegliato con delle stelle sul viso.
Rumori di campagna giungevano fino a me.
Odori di notte, di terra e di sale rinfrescavano le mie tempie.
La pace meravigliosa di quell’estate assopita entrava in me come una marea.
In quel momento e al limite della notte, si è udito un sibilo di sirene.
Annunciavano partenze per un mondo che mi era ormai indifferente per sempre.
Per la prima volta da molto tempo, ho pensato alla mamma.

Table 4.3. Translation into Italian of the passage of *L’Étranger* shown in Table 4.1.

There are several issues with the Italian translation produced by the MT system, some of the more obvious ones being the following:

- Lexical choice, mainly affecting verbs, e.g. in the first line, “lasciare” does not express the exact meaning of the French “parti”, rendered correctly by “partire”. A similar case is presented with “sono stati” versus “giungevano”.

- Verbal tense. There are two clear cases where the MT produces a wrong verbal form in terms of tense, i.e. “sono stati raffreddando” (instead of “rinfrescavano”) and “hanno annunciato” (instead of “annunciavano”).
- Particles. The professional translator uses particles that make the translation more fluent in Italian compared to the output produced by the MT system, e.g. “con delle stelle” vs “con stelle”, “ho pensato alla mamma” vs “ho pensato a mamma”.
- Agreement, mainly gender. This occurs, for example, between noun (“pace”, feminine) and adjective (“meraviglioso”, masculine), between noun (“estate”, feminine) and participle (“addormentato”, masculine).

While some of these types of issues are similar to those occurring for English (e.g. lexical choice, verbal time and particles), others are different (e.g. agreement).

In particular, we note that the MT output for Italian seems to provide a better basis than that for English to reach the reference (i.e. the translation by Zevi or Ward, respectively) by means of post-editing. Table 4.4 gives the MT outputs for Italian and English, indicating the portions that match the respective language references and those that would need to be post-edited. Note that sequences in bold match the reference, sequences in regular font do not (so would need to be post-edited) and sequences between brackets indicate required insertions. As shown in the table, Italian MT output results in longer sequences in bold. Measured in number of character edits to reach the reference, 185 are needed for Italian and 212 for English. Note that 226 edits would be needed for the English output produced with the 2013 system, which demonstrates a 6% improvement in real terms in the intervening two years.

Italian - Google Translate (June 2015)
Ha lasciato, sono tornato al**la calma**. **Ero esausto e mi sono** buttato **sulla** mia cuccetta. Credo di **aver dormito perché mi sono svegliato con** [delle] **stelle sul viso**. **Rumori d**ella **campagna** sono stati **fino a me**. **Odori** notturni **di terra e di sale** sono stati raffreddando **le** [mie] **tempie**. **La pace meraviglios**o [di] **que**sta **estate** addormentato mi è **entra**to **come una marea**. **In quel momento**, **e** sul bordo **della notte**, [si è udito un sibilo di] **sirene** ululavano. Hanno **annuncia**to **partenze per un mondo che** [mi era] **ormai** era mai indifferente [per sempre]. **Per la prima volta** in molti anni, **ho pensato a**[lla] **mamma**.
English - Google Translate (June 2015)
He left**, I** returned **to calm** [down again]. **I was exhausted and** I **threw myself on my bunk**. **I** think I slept **because I woke up with** [the] **stars on** the **face**. Campaign noises **were** up to me. Night **smells of** [night,] **earth and salt** [air] **were cooling my temples**. **The** wonderful **peace** [of] **this** [sleeping] **summer** asleep entered **me like a tide**. At that moment, and on the edge of the night, **sirens** howled. **They** [were] **announc**ed **departures for a world that now** was never indifferent **to me**. **For the first time in** many years, **I thought about** Mom.

Table 4.4. Edits required in the MT outputs of the passage of *L’Étranger* into Italian and English to reach the respective reference translations (cf. Tables 4.1 and 4.3).

4.2. Quantitative Analysis

In order to carry out a quantitative analysis, we consider not just a passage, as in the last section, but rather the whole novel. We preprocess the datasets (*L'Etranger* in French, its translation by Zevi in Italian, and the two translations in English by Ward and Gilbert) as follows. The books are sentence split with language-specific splitters included in the NLTK toolkit.[9] Then we use Hunalign (Varga et al., 2005) to align the sentences of the following book pairs: French–Italian, French–English (Ward) and French–English (Gilbert). We keep only the subsets of 1-to-1 sentence alignments.[10] Finally, we build a multilingual dataset by keeping the 1-to-1 sentence alignments that are common in our three aligned datasets. Our final dataset thus comprises equivalent translations in each of our initial datasets: French, Italian, English (Ward) and English (Gilbert). Our initial datasets contained 2,289 (French), 2,176 (Italian), 2,315 (English/Ward) and 2,288 (English/Gilbert) sentences, while our final multilingual dataset is made of 1,572 groups of four sentences.

Table 4.5 shows the results obtained when using Google Translate to translate *L'Étranger* into Italian (using Zevi as the reference) and English (using both Ward and Gilbert as the references). We report results using two widely used automatic metrics: BLEU and TER (Snover et al., 2006). BLEU is the *de facto* standard metric in the MT field. We use also TER as it is an error-rate metric whose score is based on the number of operations (insertions, deletions and edits) that are required to bring the MT output to match the reference, which makes it more applicable to the machine-assisted translation scenario we envisage. Furthermore, TER has been shown to correlate well with PE time (O'Brien, 2011). In order to interpret the results, it should be borne in mind that both metrics operate in the scale 0 to 100.[11] For BLEU, the higher the scores the better (100 indicating that the MT output and the reference are identical), while for TER the lower the score the better (0 indicating that the MT output and the reference are identical).

Translation direction	BLEU	TER
French to Italian	24.73	59.78
French to English (Ward)	28.45	56.10
French to English (Gilbert)	11.22	80.08
French to English (Ward and Gilbert)	32.46	52.12

Table 4.5. Scores by automatic metrics on MT output of *L'Étranger* into Italian and English.

There are a number of observations to be made on these results. Firstly, while one might expect the best scores to be obtained for Italian, due to its closer relatedness to French, this is not the case; the results into English using Ward's translation as reference are slightly better (around 3.7 points both for BLEU and TER). This has to do with a number of factors, as we provide

9 http://nltk.org/

10 Hunalign can produce 1-to-1, 1-to-many and many-to-1 sentence alignments.

11 Note that TER can provide results higher than 100.

indications for in the following: (i) Italian has a relatively complex morphology compared to English, which leads, as shown in the previous section, to agreement errors; (ii) Ward's translation seems to use plainer language compared to the translation into Italian and (iii) Google Translate's LM for English is probably much better than that for Italian as the amount of monolingual data available on-line, and thus probably used by Google Translate, in the first language is considerably bigger than that for the second.

In the previous section we noted the differences between the two English translations. This is underlined still further by the huge difference in scores between using either Ward's or Gilbert's as the reference (17 absolute BLEU points and 24 absolute TER points, respectively). To reiterate, a BLEU score of just 11 indicates a system whose output is far too poor to be of use in a machine-assisted translation scenario; in contrast, using Ward's translated sentences as reference translations against which to compare the MT output, a much higher – and satisfactory – score of 28 BLEU points is obtained. When both sets of translations are used as reference, unsurprisingly the score rises again to 32 BLEU points.

From this we can conclude that the type of literary translation that one aims to produce (at first sight Gilbert's appears to be a considerably freer translation compared to Ward's) is a major factor in whether MT can be of assistance or not, even more so than the level of relatedness between the source and target languages. In the future, we would like to analyse different types of literary translations to identify what makes them different and what those differences imply in terms of challenges for MT.

5. Literary-Adapted Machine Translation between Related Languages

In the experiments described in the previous section, we used freely available generic web-based MT systems. We now consider literary-adapted MT systems. In our previous work (Toral and Way, 2015), we proposed a methodology to build SMT systems adapted to novels and we conducted an experiment comparing generic and adapted systems to translate *El Prisionero del Cielo* (Ruiz Zafón 2011) between two closely-related languages, Spanish to Catalan. We measure the translation quality of the MT outputs with automatic metrics using the professional translation as reference on the whole book (4,846 sentences)[12] and observe that the adapted system leads to considerably better scores (47.2 versus 42.9 BLEU and 39.7 versus 42.1 TER).

In MT evaluation, sentences are extracted at random for testing. However, this sentence-level evaluation does not take context into account, so to try to mitigate this somewhat, we now provide an analysis of three short passages, each containing 5 to 10 sentences. These passages are selected to be representative of:

[12] The original book in Spanish contains 5,044 sentences (according to the sentence splitter we used). When this is sentence-aligned to the professional translation into Catalan, it leads to 4,846 1-to-1 sentence pairs.

1. The average MT quality. We select a passage for which its BLEU score is similar to the BLEU score obtained on the whole novel.
2. Low MT quality.We select a passage whose BLEU score is similar to the average BLEU score of the 20% lowest-scoring passages.
3. High MT quality. We select a passage whose BLEU score is similar to the average BLEU score of the 20% highest-scoring passages.

For each passage, we consider the text source text, the MT output (with indication of the edits required to reach the reference, as in Table 4.4) and the reference translation, i.e. the published translation in Catalan (Ruiz Zafón and Pelfort Gregori 2012). In addition, as English gloss, we consider its published translation into this language (Ruiz Zafón and Graves 2012). In the remainder of this section we analyse the MT output produced for each of the three passages.

Spanish - original

La abracé y permanecimos en silencio unos minutos.
— He estado pensando — dijo Bea.
Tiembla, Daniel, pensé.
Bea se incorporó y se sentó en cuclillas sobre el lecho frente a mí.
— Cuando Julián sea algo mayor y mi madre pueda cuidarlo unas horas durante el día, creo que voy a trabajar.
— ¿Dónde?
— En la librería.
La prudencia me aconsejó callar.
— Creo que os vendría bien — añadió —.
Tu padre ya no está para echarle tantas horas y, no te ofendas, pero creo que yo tengo más mano con los clientes que tú y que Fermín, que últimamente me parece que asusta a la gente.

Catalan - adapted SMT

La vaig abraçar i ens **vam** quedar **en silenci** [, durant] **uns minuts**.
— He estat rumiant — va dir la Bea.
[Ja pots] T**remola**[r], **Daniel, vaig pensar** [jo].
La Bea es va incorporar i es **va** as**seure** a la gatzoneta sobre **el llit** davant meu.
— Quan el Julià sigui una mica més gran i la meva mare pugui cuidar-lo **unes** quantes **hores** durant e**l dia**, **em sembla que** vaig a **treballar**[é].
— ¿On?
— A la llibreria.
La prudència em va aconsellar [mantenir-me] **calla**r.
— Em sembla que us aniria bé — [hi] **va afegir —**.
El teu pare ja no és per fer **tantes hores i, no t'**ofenguis, però crec que jo **tinc més** mà amb **els clients que tu i que el Fermín, que últimament** em **sembla que** [fins i tot] **espanta la gent**.

Catalan - translation by Pelfort Gregori

La vaig abraçar i vam estar així, en silenci, durant uns minuts.
— He estat rumiant — va fer ella.
Ja pots tremolar, Daniel, vaig pensar jo.
La Bea es va alçar i va seure al meu costat del llit.
— Quan el Julià sigui una mica més gran i la meva mare se 'n pugui fer càrrec unes hores al dia, em sembla que treballaré.
— ¿On?
— A la llibreria.
La prudència em va aconsellar de mantenir-me callat.
— Em sembla que us aniria bé — hi va afegir —.
El teu pare ja no està en condicions de dedicar-hi tantes hores i, no t'ho prenguis malament, em fa l'efecte que tinc

més traça jo, a l'hora de tractar els clients, que tu i que el Fermín, que últimament sembla que fins i tot espanta la gent.
English - translation by Graves ‘I've been thinking,' said Bea . Tremble, Daniel, I thought . Bea sat up and then crouched down on the bed facing me . ‘When Julián is a bit older and my mother is able to look after him for a few hours a day , I think I'm going to work.' I nodded. ‘Where?' ‘In the bookshop.' I thought it best to keep quiet. ‘I think it would do you all good,' she added. ‘Your father is getting too old to put in all those hours and, don't be offended, but I think I'm better at dealing with customers than you , not to mention Fermín, who recently seems to scare business away.'

Table 5.1. Source, MT output, reference translation and English gloss for a passage of average MT quality of *The Prisoner of Heaven*

5.1. Average Quality Passage

As can be observed by the long sequences shown in bold, the MT output seems a reasonable starting point for post-editing. Some of the differences between the MT output and the professional translation could be considered of equivalent quality, with the MT outputs being more literal with respect to the source, e.g. “va dir la Bea” (“Bea said”) vs “va fer Ella” (“she said”). In some other cases though it is clear that the MT output is of lower quality. From these errors we can identify two types according to how serious they are:

- *Disfluencies*, e.g. lack of pronoun *hi* in “hi va afegir”. *Hi* is a weak pronoun. While other Romance languages such as Italian contemplate this grammatical element, Spanish does not. Hence, when translating from Spanish into Catalan, these pronouns are problematic as it is a challenge for the system to produce them out of the blue.
- *Errors*, e.g. “vaig a treballar” vs. “treballaré” (I’m going to work). The construction chosen by the MT system (to go + infinitive) is a calque of the grammatical construction used in the original (“voy a trabajar”). While this is a correct construction to express future tense in Spanish, it is not in Catalan.

Spanish – original — A ver si tenemos suerte, porque el muñón le está empezando a supurar y eso va a oler que no le cuento … — Joder — dijo el carcelero alejándose a toda prisa. Tan pronto como lo oyó llegar al extremo del corredor, Fermín procedió a desnudar a Salgado y luego se desprendió de sus ropas. Se vistió con los harapos pestilentes del ladrón y le puso los suyos. Colocó a Salgado de lado en el camastro, de cara al muro, y lo tapó con la manta hasta cubrirle medio rostro. Entonces agarró el saco de lona y se introdujo dentro. Iba a cerrar la saca cuando recordó algo. Volvió a salir a toda prisa y se acercó al muro. Rascó con las uñas entre las dos piedras donde había visto a Salgado esconder la llave hasta que asomó la punta. Intentó asirla con los dedos, pero la llave resbalaba y quedaba apresada entre la piedra
Catalan - adapted SMT **— A veure si** tenim **sort, perquè el monyó** [ja] **li** està **començ**ant **a supurar** y eso va a ensumar que no li explico

…
— Joder — **va dir el carceller** allunyant-se a corre-cuita.
Tan aviat com ho **va sentir** arribar a l'extrem del corredor, **el Fermín va** procedir **a despullar** [el] **Salgado i**[,] **després**[,] **es va** desprendre **de la** [seva] **roba**.
[Aleshores,] **Es va vestir amb els parracs pestilents del lladre i** [a ell] **li va posar els seus**.
Va col·locar [el] **Salgado de costat** al **llit** tronat, **de cara al mur, i el va tapar amb la** manta **fins** cubrirle mig rostre.
Llavors va agafar el sac de lona i es va introduir **a dins**.
Anava **a tancar** la treu **quan va recordar** alg**una cosa**.
[En] **Va tornar a sortir** a corre-cuita **i es va acostar** al mur.
Rascó **amb les ungles** [el forat que quedava] **entre** les **dues pedres,** [que era el lloc] **on havia vist** [que el] **Salgado amaga**r **la clau**[,] **fins que** [en] **va** treure **la punta**.
Va intentar asirla **amb els dits, però la clau** [li re]**lliscava i** [es] **quedava** capturada **entre l**a **pedr**a.

Catalan - translation by Pelfort Gregori

— A veure si estem de sort, perquè el monyó ja li comença a supurar i això farà una ferum que Déu n'hi do …
— Cagondéna — va dir el carceller, que va tocar el dos ben de pressa.
Tan aviat com va sentir que ja era a l'altre extrem del passadís, el Fermín va començar a despullar el Salgado i, després, es va desfer de la seva roba.
Aleshores, es va vestir amb els parracs pestilents del lladre i a ell li va posar els seus.
Va col·locar el Salgado de costat, damunt del llit i de cara al mur, i el va tapar amb la màrfega fins a cobrir-li la meitat de la cara.
Llavors, va agafar el sac de lona i s'hi va ficar a dins.
Ja es disposava a tancar el sac quan va recordar una cosa.
En va tornar a sortir de pressa i es va acostar a la paret.
Va gratar amb les ungles el forat que quedava entre dues pedres, que era el lloc on havia vist que el Salgado hi amagava la clau, fins que en va sortir la punta.
La va voler agafar amb els dits, però la clau li relliscava i es quedava entaforada entre les pedres.

English - translation by Graves

‘ Let's hope we're in luck and it works out , because his stump is starting to ooze and I can't begin to tell you what that's going to smell like … '
‘ Shit , ' said the jailer , scuttling off .
As soon as he heard him reach the end of the corridor , Fermín began to undress Salgado .
Then he removed his own clothes and got into the thief 's stinking rags .
Finally , Fermín put his own clothes on Salgado and placed him on the bed , lying on his side with his face to the wall , and pulled the blanket over him , so that it half-covered his face .
Then he grabbed the canvas sack and got inside it .
He was about to close it when he remembered something .
Hurriedly , he got out again and went over to the wall .
With his nails , he scratched the space between two stones where he'd seen Salgado hide the key , until the tip began to show .
He tried to pull it out with his fingers , but the key kept slipping and remained stuck between the stones .

Table 5.2. Source, MT output, reference translation and English gloss for a passage of low MT quality of *The Prisoner of Heaven*

5.2. Low Quality Passage

The main errors in this section have to do with out-of-vocabulary words (i.e. source words that are not known by the MT translation model and thus the system just outputs them as they are). Examples include mainly verbs, e.g. “Rascó” (he scratched) and “cubrirle” (cover his), but also other linguistic elements, e.g. “Joder” (shit), “y” (and) and “eso” (that). In addition, we again see disfluencies in the MT output, where weak pronouns, such as *li* and *en*, are missing.

Spanish – original

Eran casos de poca monta, pero todos los clientes habían abonado un retente y firmado un contrato.
— Fermín, le voy a poner un sueldo fijo.
— Ni hablar.
Fermín se negó a aceptar emolumento alguno por sus buenos oficios excepto pequeños préstamos ocasionales con los que los domingos por la tarde se llevaba a la Rociíto al cine, a bailar a La Paloma o al parque del Tibidabo, donde en la casa de los espejos la joven le dejó un chupetón en el cuello que le escoció una semana y donde, aprovechando un día en que eran los dos únicos pasajeros en el avión de falsete que sobrevolaba en círculos el cielo en miniatura de Barcelona, Fermín recuperó el pleno ejercicio y goce de su hombría tras una larga temporada alejado de los escenarios del amor apresurado.
Un día, magreando las beldades de la Rociíto en lo alto de la noria del parque, Fermín se dijo que casi parecía que aquéllos, contra todo pronóstico, estaban resultando ser buenos tiempos.
Y le entró el miedo, porque sabía que no podían durar y que aquellas gotas de paz y felicidad robadas se evaporarían antes que la juventud de la carne y los ojos de la Rociíto.

Catalan - adapted MT

Eren casos de pa sucat amb oli, **però tots els clients havien abonat** un retente **i firmat un contracte**.
— Fermín, li haig de posar **un sou fix**.
— Ni parlar-ne.
El Fermín es va negar **a acceptar** emolumento d'interès **pels seus bons oficis excepte petits préstecs ocasionals amb** els que **el**s **diumenge**s **a la tarda** s'enduia a **la Rociíto al cine, a ballar a La Paloma o al parc del Tibidabo, on a la casa dels miralls la noia li va deixar un** chupetón **al coll que li** escoció **una setmana i on, aprofitant un dia en què eren els dos únics passatgers a l'avió de** falsete **que sobrevolava en cercles el cel en miniatura de Barcelona, el Fermín va recuperar el ple exercici i** el **gaudi de la seva** homenia **després d'una llarga temporada allunyat dels escenaris de l'amor** apressat.
Un dia, magreando els atributs **de la Rociíto a dalt de la** nòria **del parc, el Fermín es va dir que** gairebé semblava **que aquells**, **contra tot pronòstic**, estaven resultant ser **bons temps**.
I li va entrar la por, **perquè sabia que no podien durar i que aquelles gotes de pau i felicitat robades** es evaporarían **abans que la joventut de la carn i els ulls de la Rociíto**.

Catalan - translation by Pelfort Gregori

Eren casos de poca volada, però tots els clients havien abonat paga i senyal i firmat un contracte.
— Fermín, li posaré un sou fix.
— Ni parlar-ne.
El Fermín es negava a acceptar emoluments pels seus bons oficis excepte petits préstecs ocasionals amb què el diumenge a la tarda portava la Rociíto al cine, a ballar a La Paloma o al parc del Tibidabo, on a la casa dels miralls la noia li va deixar un xuclada al coll que li va fer picor durant una setmana i on, aprofitant un dia en què eren els dos únics passatgers de l'avió de fireta que sobrevolava en cercles el cel en miniatura de Barcelona, el Fermín va recuperar el ple exercici i gaudi de la virilitat després d'una llarga temporada allunyat dels escenaris de l'amor a corre-cuita.
Un dia, grapejant les belleses de la Rociíto a dalt de tot de la roda del parc, el Fermín es va dir que,

contra tot pronòstic, resultava que aquells eren bons temps. I va tenir por, perquè sabia que no podien durar i que aquelles gotes de pau i felicitat robades s'evaporarien abans que la joventut de les carns i els ulls de la Rociíto.
English - translation by Graves They were small cases , but all the clients had paid a deposit and signed a contract . ‘Fermín, I'm going to put you on the payroll.' ‘I won't hear of it. Consider my services strictly pro bono .' Fermín refused to accept any emolument for his good offices, except occasional small loans with which on Sunday afternoons he took Rociíto to the cinema, to dance at La Paloma or to the funfair at the top of the Tibidabo mountain. Romance was in the air, and Fermín was slowly reclaiming his old self. Once, in the funfair's hall of mirrors, Rociíto gave him a love bite on the neck that smarted for a whole week. On another occasion, taking advantage of the fact that they were the only passengers on the full-sized aeroplane replica that gyrated, suspended from a crane, between Barcelona and the blue heavens, Fermín recovered full command of his manhood after a long absence from the scenarios of rushed love. Not long after that, one lazy afternoon when Fermín was savouring Rociíto's splendid attributes on the top of the big wheel, it occurred to him that those times, against all expectations, were turning out to be good times. Then he felt afraid, because he knew they couldn't last long and those stolen drops of happiness and peace would evaporate sooner than the youthful bloom of Rociíto's flesh and eyes.

Table 5.3. Source, MT output, reference translation and English gloss for a passage of high MT quality of *The Prisoner of Heaven*

5.3. High Quality Passage

The main errors in this passage regard again out-of-vocabulary words. Most of them are nouns, e.g. “emolumento” (emolument), “retente” (deposit), “chupetón” (love bite), and verbs, e.g. “escoció” (smarted), “evaporarían” (would evaporate), “magreando” (savouring).

We would like to emphasise two positive achievements of the literary-adapted MT system in this passage:

- MT results in a rather fairly accurate translation for a very long sentence of over 100 words, for which just a few edits would cause it to match the reference.
- MT produces an appropriate and fluent Catalan expression, “de pa sucat amb oli” (literally “of bread dipped in oil” but meaning “petty”) from the Spanish “de poca monta” (literally “of little importance”). The unadapted system translates this literally as “de poca volada”.

6. Conclusion and Future Work

In this paper, we motivated the opportunity to explore the applicability of MT to literary texts and explored the role of MT in assisting with the translation of this type of texts.

First, we studied the role of translation options in literary translations and its relation to language relatedness. Taking Camus' L'Étranger as our case study, we discovered that (i) different professional translations can be very divergent, which poses a challenge to MT as there is no unique 'gold standard' reference translation to aim for; (ii) general-domain SMT has progressed in the last couple of years, to the extent that for the passage considered, 6% fewer character edits are required with the latest available system; and (iii) translation between related languages would seem to be easier to be post-edited.

Second, we analysed the quality attained by literary-adapted SMT to translate a novel between closely-related languages. We provided three passages representative of (i) the average quality attained by MT, (ii) low-performing subsets and (iii) high-performing subsets. For each of these we have analysed the main errors committed by MT and shown how suitable it would be for them to be post-edited to match the reference.

Finally, we outline our future research plans to build on the preliminary results presented in this paper to fully MT-assisted workflows for the translation of novels. We believe there are two lines of work in going forward.

1. Improvement of MT for literary texts. We have so far explored adapting MT systems to the writing style of an author and to the translation style of a translator. Related to this we propose to adapt MT systems to the different aspects of prose fiction (descriptions, dialogue, action, etc.). Another important aspect on improving MT for literature regards currently weak aspects of MT such as its treatment of cohesion and figurative language.

2. In order for MT to be used to assist with the translation of literary text, we not only need to improve its performance but also find out suitable literary MT-assisted translation workflows. Literary text is not translated in the same way as other domains on which MT is successfully applied commercially (e.g. technical documentation), so it might be the case that MT-assisted workflows used on these domains (post-editing) are not suitable and other alternatives, such as interactive MT, in which the translator is provided with MT suggestions as he/she types the translation, might suit better.

Acknowledgments

This research is supported by the European Union Seventh Framework Programme FP7/2007-2013 under grant agreement PIAP-GA-2012-324414 (Abu-MaTran) and by Science Foundation Ireland through the CNGL Programme (Grant 12/CE/I2267) in the ADAPT Centre (www.adaptcentre.ie) at Dublin City University.

References

Bellos, David. 2012. *Is That a Fish in Your Ear?: Translation and the Meaning of Everything*. London: Particular Books.

Besacier, Laurent. 2014. "Traduction automatisée d-une oeuvre littéraire: une étude pilote". In *Traitement Automatique du Langage Naturel (TALN)*, Marseille, France. pp. 389–394.

Brown, Peter, John Cocke, Stephen Della Pietra, Vincent Della Pietra, Fred Jelinek, John Lafferty, Robert Mercer, and Paul Roosin. 1990. "A statistical approach to machine translation". *Computational Linguistics* 16: 79–85.

Brown, Peter, Stephen Della Pietra, Vincent Della Pietra, and Robert Mercer. 1993. "The mathematics of statistical machine translation: parameter estimation". *Computational Linguistics* 19: 263–311.

Camus, Albert. 1942. "L'Étranger". Paris: Librairie Gallimard.

Camus, Albert, and Stuart Gilbert. 1946. "The Stranger". New York: Alfred A. Knopf, Inc.

Camus, Albert, and Matthew Ward. 1989. "The Stranger". New York: Knopf Doubleday Publishing Group.

Camus, Albert, and Alberto Zevi. 1987. "Lo straniero". Milan: Bompiani.

Fancellu, Federico, Morgan O'Brien, and Andy Way. 2014. "Standard language variety conversion using SMT". In *EAMT-2014: Proceedings of the Seventeenth Annual Conference of the European Association for Machine Translation*, Dubrovnik, Croatia, pp.143–149.

Genzel, Dmitriy, Jakob Uszkoreit, and Franz Och. 2010. "'Poetic' Statistical Machine Translation: Rhyme and Meter". In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Cambridge, Mass., USA, pp.158–166.

Graham, Yvette, Timothy Baldwin, Alistair Moffat, and Justin Zobel. 2014. "Is machine translation getting better over time?" In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Gothenburg, Sweden, pp.443–451.

Greene, Erica, Tugba Bodrumlu, and Kevin Knight. 2010. "Automatic analysis of rhythmic poetry with applications to generation and translation". In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, Cambridge, MA, USA, pages 524–533.

Hardmeier, Christian. 2014. *Discourse in Statistical Machine Translation*. PhD Thesis, University of Uppsala, Uppsala, Sweden.

Hearne, Mary, and Andy Way. 2011. "Statistical Machine Translation: A Guide for Linguists and Translators". *Language and Linguistics Compass* **5**:205—226.

Jones, Ruth, and Ann Irvine. 2013. "The (Un)faithful Machine Translator" In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, Sofia, Bulgaria, pp.96–101.

Kelly, Nataly, and Jost Zetzsche. 2012. *Found in Translation: How Language Shapes Our Lives and Transforms the World*. New York: Perigee Trade.

Landers, Clifford E. 2001. *Literary Translation: A Practical Guide*. Bristol: Multilingual Matters Ltd.

Li, Linlin, and Caroline Sporleder. 2010. "Using gaussian mixture models to detect figurative language in context". In *Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Los Angeles, CA., USA. pp.297–300

Lønning, Jan T, Stephan Oepen, Dorothee Beermann, Lars Hellan, John Carroll, Helge Dyvik, Dan Flickinger, Janne Bondi Johannessen, Paul Meurer, Torbjørn Nordgård, Victoria Rosén, and Erik Velldal. 2004. "LOGON. A Norwegian MT effort". In *Proceedings of the Workshop in Recent Advances in Scandinavian Machine Translation*, Uppsala, Sweden, 6pp.

Meyer, Thomas. 2014. *Discourse-level features for statistical machine translation*. PhD Thesis, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Nida, Eugene, and Charles Taber. 1969. *The Theory and Practice of Translation, With Special Reference to Bible Translating*. Leiden: Brill.

O'Brien, Sharon. 2011. "Towards predicting post-editing productivity". *Machine Translation*, 25(3):197–215.

Och, Franz, and Hermann Ney. 2002. "Discriminative training and maximum entropy models for statistical machine translation". In *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics*, Philadelphia, PA, USA, pp.295–302.

Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. "BLEU: a method for automatic evaluation of machine translation". In *ACL-2002: 40th Annual meeting of the Association for Computational Linguistics*, Philadelphia, PA, USA, pp.311–318.

Penkale, Sergio, and Andy Way. 2013. Tailor-made Quality-controlled Translation. In *Proceedings of Translating and the Computer 35*, London, 7pp.

Reyes, Antonio. 2012. Linguistic-based Patterns for Figurative Language Processing: The Case of Humor Recognition and Irony Detection. *PhD Dissertation*, Universitat Politècnica de València, Valencia, Spain.

Richardson, Stephen. 2012. "Using the Microsoft Translator Hub at The Church of Jesus Christ of Latter-day Saints". In *AMTA 2012, Proceedings of the Tenth Conference of the Association for Machine Translation in the Americas*, San Diego, California, USA, 8pp.

Ruiz Zafón, Carlos. 2011. "El Prisionero del Cielo". Barcelona: Planeta.

Ruiz Zafón, Carlos, and Lucia Graves. 2012. "The Prisoner of Heaven". London: Weidenfeld & Nicholson.

Ruiz Zafón, Carlos, and Josep Pelfort Gregori. 2012. "El Presoner del Cel". Barcelona: Planeta.

Snell-Hornby, Mary. 1995. *Translation Studies, an Integrated Approach*. Amsterdam: John Benjamins.

Shutova, Ekaterina, Tony Veale, and Beata Klebanov. 2015. *Computational Modelling of Metaphor*. San Rafael, CA: Morgan & Claypool.

Snover, Matthew, Bonnie Dorr, Richard Schwartz, Linnea Micciulla, and Ralph Weischedel. 2006. "A Study of Translation Error Rate with Targeted Human Annotation". In *AMTA 2006: Proceedings of the 7th Conference of the Association for Machine Translation in the Americas, "Visions for the Future of Machine Translation"*, Cambridge, MA, USA, pp.223—231.

Toral, Antonio, and Andy Way. 2014. "Is Machine Translation Ready for Literature?" In *Proceedings of Translating and the Computer 36*. London, pp. 174-176.

Toral, Antonio, and Andy Way. 2015. "Translating Literary Text between Related Languages using SMT". In *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, NAACL, Denver, Colorado, USA, pp. 123—132.

Varga, Dániel, László Németh, Péter Halácsy, Andrés Kornai, Viktor Trón, and Viktor Nagy. 2005. "Parallel corpora for medium density languages". In *Recent Advances in Natural Language Processing*, Borovets, Bulgaria, pp.590–596.

Venuti, Lawrence. 2008. *The Translator's Invisibility: A History of Translation*. New York: Routledge.

Voigt, Rob, and Dan Jurafsky. 2012. "Towards a Literary Machine Translation: The Role of Referential Cohesion". In *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, Montreal, Quebec, Canada, pp.18–25.

Way, Andy. 2012. "Is That a Fish in Your Ear: Translation and the Meaning of Everything – David Bellos, Book Review". *Machine Translation* 26(3): 255—269.

Way, Andy. 2013. "Traditional and Emerging Use-Cases for Machine Translation". In *Proceedings of Translating and the Computer 35*, London, 12pp.

Way, Andy, and Mary Hearne. 2011. On the Role of Translations in State-of-the-Art Statistical Machine Translation. *Language and Linguistics Compass* 5:227—248.