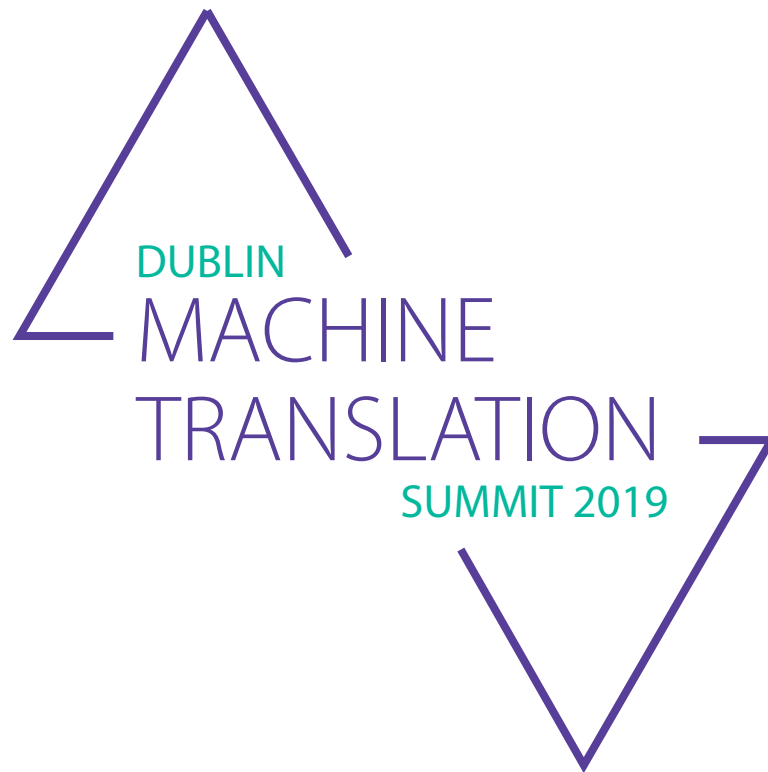# Machine Translation Summit XVII

The Qualities of Literary Machine Translation

19 August, 2019
Dublin, Ireland

The Qualities of Literary Machine Translation

19 August, 2019
Dublin, Ireland

Preface from the co-chairs of the workshop

The question of translation quality and how to define and measure it is one that has occupied a central position in both translation studies (TS) and machine translation (MT) since their respective geneses. TS has largely turned away from questions of absolute quality in recent years, towards a pluralistic notion that any translation produced by a human is a genuine reflection of that human's interpretation of the source text. With this reasoning in mind and lacking any generally agreed standard by which to judge translations, ascribing absolute or relative quality to such translations would be self-contradictory (Drugan 2013: 45). MT, however, cannot adopt the same stance with reference to its own outputs, since they are not the direct products of human interpretations, and so, can simply be inadequate or unacceptable to target readers. Nonetheless, a definable ideal of translation quality remains elusive (Way 2013).

As MT systems have developed, their use by professional translators and by end users by-passing human translators altogether has become more and more an accepted practice. However, this acceptance is only applicable to certain domains of texts. Literature has historically been held up as one domain in which machine translation and computer aided translation (CAT) are both of little or no use (Alcina 2008: 95).

The aim of this workshop is to ask whether literature really is off-limits to technology.

Of the twelve abstract submissions received for this workshop, ten (83.3%) were accepted for presentation after peer review by the workshop's Organizing Committee. Of these ten, six presenters opted also to formulate full articles, which are published in these proceedings. Of those which did not opt for full publication, only abstracts are reproduced here.

The submissions are vary widely in terms of language pairs, with as many as thirteen languages: Catalan, Dutch, English, German, Greek, Irish, Polish, Portuguese, Russian, Scottish Gaelic, Slovene, Spanish, and Turkish, analyzed from a variety of angles and taking in different issues as they pertain to the qualities of literary translations produced wholly or partly by machines.

The presenters who opted to provide full articles are:

- Kuzman, Vintar and Arčan, who examine productivity and output quality in the case of the poorly resourced and under-studied language pair of English and Slovene;

- Matusov, who looks at stories translated from English to German and Russian by NMT systems, proposing a new form of error evaluation specifically tailored to literary prose;

- Ó Murchú, who examines the issues related to post-editing literary translations produced by an ad-hoc hybrid machine translation system, with a focus on the time and effort required to bring the output to the standard required for publication;

- Şahin and Gürses, who consider the pertinent question of retranslation as it relates to NMT, asking whether and how NMT systems might be brought to bear on practicing literary translation professionals' work in order to improve productivity;

- Taivalkoski-Shilov, who points out an important feature of literary texts that has thus far been over-looked in our research into literary machine translation and which future models may need to cater for, that of free indirect discourse;

- Tezcan, Daems, and Macken, whose work is a case study of NMT used to translate Agatha Christie's The Mysterious Affair at Styles into Dutch, with a focus on error rates and stylistic differences between this NMT and the published human translation.

The presenters who did not opt to produce full articles are:

- Oliver, Toral, and Guerberof, who focus on bilingual ebooks as they relate to the training of NMT systems with the aim of increasing the number of ebooks available.

- Sklaviadis, Gong, and Crane, who bring NMT models and a wide range of lexical resources to bear on the translation of the Homeric Classics.

- Toral, Oliver, and Pau Ribas, who compare the outputs of literary translations produced with generalized NMT systems and those specifically tailored to literature.

- Zajdel, who compares the decision-making processes of human translators and NMT as they relate to the translation of metaphor in literary texts.

James Hadley, Maja Popović, Haithem Afli and Andy Way

Organizers

Workshop Chairs

James Hadley	ADAPT Centre, Dublin City University
Maja Popović	ADAPT Centre, Dublin City University
Haithem Afli	ADAPT Center, Cork Institute of Technology
Andy Way	ADAPT Centre, Dublin City University

Contents

InLéctor: Neural Machine Translation for the creation of bilingual ebooks

Antoni Oliver González
Universitat Oberta de Catalunya
`aoliverg@uoc.edu`

Antonio Toral
University of Groningen
`A.Toral.Ruiz@rug.nl`

Ana Guerberof
Dublin City University
`ana.guerberof@dcu.ie`

InLéctor is a collection of bilingual ebooks intended for helping people willing to read the original version of a novel. The reader can move from a sentence in the original to the corresponding sentence in the translated version with a click. This can be of great help to readers facing problems in difficult passages. To date we have published several books in English, French and Russian with translation into Spanish or Catalan. These bilingual ebooks are freely available (https://inlector.wordpress.com) in epub, mobi and html, so they can be read in almost any device. Until now, we have published books in the public domain with translation also in the public domain or, in some cases, with the translation rights donated to our collection. It is difficult to find novels in the public domain with translations also in the public domain and for this reason we have been able to publish a limited number of books.

In this paper we present the process of training such a literary-adapted neural machine translation (NMT) system from English to Catalan and its use to derive parallel ebooks. We also present the results of a survey conducted by a user group who have read a short history in this format, namely Arthur Conan Doyle's *The yellow face* in the bilingual English-Catalan version. Our hypothesis is that bilingual ebooks save time consulting dictionaries and make the whole reading experience more fluent.

The use of NMT systems can boost our InLéctor collection as we can now publish a large number of novels in the public domain. This also means that we can offer readers machine translated versions of books that have not been translated to date into their native language. We also plan to train NMT systems for other language pairs in order to increase the number of source and target languages in the InLéctor collection.

Embeddings for Literary NMT

Sophia Sklaviadis
Tufts University
Sophia.Sklaviadis@tufts.edu

Bowen Gong
Tufts University
Bowen.Gong@tufts.edu

Gregory R. Crane
Tufts University
gregory.crane@tufts.edu

With c. 100 million surviving words produced over more than 2,000 years —conventionally c. 750 BCE through 1453 CE— classical Greek offers a significant literary corpus. Homer's Iliad and Odyssey are two of the oldest Greek texts (c. 750 BCE) with linguistic-literary connections to the preceding Sanskrit oral poetry (Nagy 1974), as well as to the later European literary traditions. Homer has been consistently translated from antiquity to the present. The enthusiasm with which scholars have translated Homer has resulted in a complex accumulation of parallel texts, ranging from Chinese to Persian and Hindi. In French, there are more than 20 different modern translations of the Odyssey. In Modern Greek, translations at different time periods reflect changes in a language continuous with Homer's, yet inaccessibly distant without training. This paper presents a preliminary application of state-of-the-art neural machine translation (NMT) to the texts of Homer. We focus on modeling a standard edition of the source texts and English translations. We compare the effect of static, pre-trained embeddings on a seq2seq NMT model. First, we report on fitting the NMT model itself without a static embedding layer. We then discuss a qualitative evaluation of embedding spaces based on the mood-tense morphological variation of Ancient Greek verbs. Finally, we summarize the effect on the seq2seq model of pre-trained static embeddings trained (i) only on the texts of Homer (c. 200,000 words), and (ii) on the canonical-GreekLit corpus (c. 10,000,000 words, https://github.com/PerseusDL).

References

Nagy, G., 1974. Comparative studies in Greek and Indic meter (Vol. 33). https://chs.harvard.edu/CHS/article/display/6448

Automatic and Human Evaluations of Neural Machine Translation on Novels

Antonio Toral
University of Groningen
`A.Toral.Ruiz@rug.nl`

Antoni Oliver González
Universitat Oberta de Catalunya

`aoliverg@uoc.edu`

Pau Ribas
Universitat Oberta de Catalunya

`pribasba@uoc.edu`

Recently, neural machine translation (NMT) has emerged as a new paradigm in MT, and has been shown to considerably improve the translation quality achieved, regardless of the language pair (Toral and Sánchez-Cartagena, 2017). In addition, compared to the translations produced by previous paradigms to MT, those by NMT are much more fluent (Bentivogli et al., 2016) and also less literal.

Due to the above, we deem it appropriate to evaluate NMT on a content type that has historically been considered particularly challenging for MT: literary texts. Specifically, we target novels for the English-to-Catalan language direction and consider different NMT systems: commercial offerings as well as in-house systems tailored to novels trained under the recurrent with attention architecture (Bahdanau et al., 2014) and with an attention-only approach, commonly referred to as Transformer (Vaswani et al., 2017). We conduct two evaluations:

- An automatic evaluation with BLEU (Papineni et al., 2002), the most widely-used automatic evaluation metric in MT, on a set of twelve widely-known novels (Toral and Way, 2018), including for example J. Joyce's Ulysses and J. K. Rowling's Harry Potter The Deathly Hallows . The results show that NMT systems, particularly Transformer, bring notable improvements in translation performance.
- A human evaluation, on a fragment of Arthur Conan Doyle's The yellow face. In this evaluation a human post-edition of the text has been performed, making the minimum changes for the target segments to be acceptable. After this postedition, the errors have been manually classified in several categories.

Both automatic and human evaluations show that specifically tailored systems using a literary corpus perform much better than general-purpose commercial systems. The quality levels obtained with the tailored systems are good enough to use the MT system in certain situations, as for example where a human translation of the work is not available or for the creation of reading aids.

Machine versus human: Comparing human and machine translations of metaphors in *The Picture of Dorian Gray*

Alicja Zajdel

Trinity Centre for Literary and Cultural Translation, Trinity College Dublin

`zajdela@tcd.ie`

Although the recent shift from statistical to neural machine translation (MT) systems has made MT a frequently used tool in the translation industry, specialists in literary translation remain sceptical of the usefulness of the technology for literature. This study puts MT to the test, by exploring its possibilities and limitations when translating literary texts rich in metaphorical language. It does this by comparing solutions used by Google Translate to translate metaphors in *The Picture of Dorian Gray*, with those used by human translators across three languages: Spanish, Portuguese and Polish. Using a parallel corpus, this study identifies patterns in the decision-making processes of both MT and human translators and evaluates how and to what extent they differ. Through analysis and visualisation of the collected data, the results of this study provide an opportunity to assess the current suitability of Google Translate for literary texts and may be useful in the programming of improved MT systems in the future.