

Article

Human versus Neural Machine Translation Creativity: A Study on Manipulated MWEs in Literature

Gloria Corpas Pastor *  and Laura Noriega-Santiañez 

Research Institute of Multilingual Language Technologies (IUITLM), University of Malaga, 29016 Malaga, Spain; laura.noriega@uma.es

* Correspondence: gcorpas@uma.es

Abstract: In the digital era, the (r)evolution of neural machine translation (NMT) has reshaped both the market and translators' workflow. However, the adoption of this technology has not fully reached the creative field of literary translation. Against this background, this study aims to explore to what extent NMT systems can be used to translate the creative challenges posed by idioms, specifically manipulated multiword expressions (MWEs) found in literary texts. To carry out this pilot study, five manipulated MWEs were selected from a fantasy novel and machine-translated (English > Spanish) by four NMT systems (DeepL, Google Translate, Bing Translator, and Reverso). Then, each NMT output as well as a human translation are assessed by six professional literary translators by using a human evaluation sheet. Based on these results, the creativity obtained in each translation method was calculated. Despite the satisfactory performance of both DeepL and Google Translate, HT creativity was highly superior in almost all manipulated MWEs. To the best of our knowledge, this paper not only contributes to the ongoing study of NMT applied to literature, but it is also one of the few studies that delve into the almost unexplored field of assessing creativity in neural machine-translated MWEs.

Keywords: literary translation; neural machine translation; creativity; manipulated multiword expressions; human evaluation



Citation: Corpas Pastor, G.; Noriega-Santiañez, L. Human versus Neural Machine Translation Creativity: A Study on Manipulated MWEs in Literature. *Information* **2024**, *15*, 530. <https://doi.org/10.3390/info15090530>

Academic Editor: Ivan Dunder

Received: 29 July 2024

Revised: 20 August 2024

Accepted: 25 August 2024

Published: 2 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the last decades, both the significant development of artificial intelligent (AI) and the intertwining of different disciplines (such as natural language processing, corpus linguistics, or machine translation) have led to the redefining of technological tools and resources that have shaken the Translation field [1,2]. In fact, translation is considered “a form of human-computer interaction”, as the translation task heavily relies on computer tools [3] (p. 4).

Nevertheless, this technological adoption has not reached or affected all translation genres and text-types in the same way. For instance, literary texts (including novels, comics, or poems, among others) are extensively considered creative, as they revolve around the aesthetics in their production, in contrast to the direct objectives often found in more technical texts [4]. These “creative texts are, to a large extent, defined by their idiosyncrasy, fitting into one and many national, cultural, temporal and even personal styles” [4] (p. 6). Thus, literary translation deals with the dynamic interplay between the ever-renewing form and content of literary texts and the syntactic and semantic limitations of the target language and culture in order to maintain cohesion, coherence, style, and impact [5]. For that reason, Ruffo [6] (p. 18) noticed that “the very nature of creative texts almost implies an inherent degree of resistance to automation” that may constrain the literary translator's skills.

Despite the scepticisms towards technological advances, in recent years, the interaction between machine translation (MT) and literary texts has begun to catch the attention of numerous scholars (cf. [7–9], to name but a few). Specifically, this work focuses on those studies that considered or measured the creative factor presented in MT [10–13].

In the realm of AI, the creative process represents one of the reasons against the absolute adoption of technologies. However, defining creativity is a major challenge. García Álvarez [14] (p. 13) associated creativity with concepts such as “originality of thought, intellectual curiosity, imagination, decision-making capacity, and critical reasoning”, whereas Guerberof-Arenas and Toral [11,12] considered that creativity involved “novelty” and “acceptability”. In general terms, Şahin and Gürses [10] (p. 27) stated that creativity in translation means “solutions that go beyond literal translation and differ from the MT solution”. Regardless of the different definitions, creativity is undoubtedly an essential skill for the literary translator. Precisely, the PETRAE-E framework of reference for the education and training of literary translators [15] highlights, among some of its core competencies, that professional literary translators should achieve an “optimal creative ability” or “find solutions and make choices creatively”.

Against this background, this paper aims at exploring to what extent neural machine translation (NMT) can achieve satisfactory results when translating the creative challenges posed by manipulated multiword expressions (MWEs) found in literary texts, with special reference to idioms. To carry out this pilot study, five manipulated MWEs were selected from the American fantasy novel *Black Sun* (*Black Sun* (2020) by Rebecca Roanhorse, is an epic fantasy novel, the first book in the *Between Earth and Sky* trilogy, inspired by the civilizations and culture of pre-Columbian Americans) by Rebecca Roanhorse [16]. Then, four NMT systems, namely DeepL (Accessed online at <https://www.deepl.com/translator>, accessed on 22 May 2024), Google Translate (Accessed online at <https://translate.google.com>, accessed on 22 May 2024), Bing Translator (Accessed online at <https://www.bing.com/translator>, accessed on 22 May 2024), and Reverso (Accessed online at <https://www.reverso.net/traduccion-texto>, accessed on 22 May 2024) in the English > Spanish language pair were tested, as well as a human translation (HT) made by a professional literary translator. Each NMT output is assessed by six professional literary translators by using a human evaluation sheet. Thus, the study pursues answers to three research questions:

- How do NMT systems and HT perform in translating manipulated MWEs, according to the proposed equivalence parameters?
- To what extent can these NMT systems be compared to a professional HT in terms of creativity?
- What is the opinion of literary translators with regard to integrating NMT systems into their workflow to translate literature, in general, and manipulated MWEs, in particular?

In accordance with these goals, this paper is structured in six sections. Section 2 outlines previous studies related to the use of MT in literature, with special emphasis on creativity and manipulated MWEs. Section 3 describes the protocolised methodology, pinpointing the selected manipulated MWEs in context as well as the human evaluation profiles and sheet. Section 4 presents the results of this study, i.e., the human evaluation of the NMT output in terms of creativity, and the post-evaluation questions. Section 5 discusses the primary findings against previous studies in the field, with a special focus on the three research questions proposed. Finally, Section 6 draws the main conclusions and details the future lines of research.

2. Literary Translation and Creativity in the MT Era

The digital era has introduced multiple technological tools and resources, transforming both the market and the translator’s workflow. For instance, translators have at their disposal online dictionaries, spell checkers, databases, revision tools, lexicons, corpora, MT systems, translation memories (TMs) and termbases in computer-assisted translation (CAT) tools, among many others [2,17,18]. In contrast, as mentioned above, literary translation has partially remained on the sidelines of this technology outburst, being considered “the last bastion of human translation” [19] (p. 174).

Nevertheless, recent studies on translators’ attitudes towards technology pointed in another direction [6,20,21]. Ruffo [6] (p. 34) stated that literary translators are not opposed to technological adoption, but only to tools that compromise “literary translators’ self-

image” or that interfered “with creativity, originality, and freedom”. With the development of technologies such as CAT tools, TMs, and MT systems, literary translators are more likely to embrace their benefits not only to increase their productivity, but also to support the ideation process [22]. In fact, the refinement of MT is very likely to affect the field of literature in the years to come, and so only the most skilled literary translators will be able to easily navigate this new technological scenario [20]. Therefore, literary translators are currently balancing the assistance provided by technologies with the creative nature of their craft [23].

This idea is embodied in the fact that there has been ongoing research of MT applied to literature in recent decades. At the beginning, some authors explored the role of statistical or phrase-based machine translation systems to compare the human translation (HT) against both the raw MT and the machine translation post-edited (MTPE) output [19,24,25]. Later, after the emergence of NMT systems that “can attain better translation quality than the dominant approach to date” [26] (p. 264), handling diverse texts and genres in novel contexts [27], more refined studies appeared on the scene. In particular, some authors employed evaluation metrics (such as BLEU, TER, METEOR, or COMET, among others) to assess the NMT output (provided by NMT systems such as Google Translate, DeepL, Bing Translator, or Phrase TMS) in terms of quality, effort, productivity and/or time compared to HT and MTPE [5,7,13,28–32]. Others specifically focused on the MT quality of literary elements such as the metaphorical language [33,34] and the quality of stylistic and narratological features as well as the accuracy and fluency [35] and the challenges posed by neologisms or manipulated MWEs [13,36]. Finally, some studies introduced customised NMT systems that might improve the performance of general NMT systems to translate [27,29] or even developed literariness algorithms to predict literary quality ratings [35].

However, several ethical factors should be carefully considered in terms of “translation as process, product and industry” [20] (p. 692). For instance, the training of MT systems with the output of authors and translators raised concerns about their intellectual property rights. Moreover, the employment landscape for translators is not ideal, plagued by “constant pressure on price, abstract measures of quality, fears of being replaced by AI” [28] (p. 5), and sometimes their own professional ethics are devalued [37]. In addition, there is also an imbalance of linguistic diversity and equitable representation in literature, as MT systems can marginalise or restrict lesser-resourced languages [23]. Furthermore, the literary translator’s creative voice can be constrained in post-edited texts [38], leading to homogenisation and normalisation [39].

Against this background, this paper investigates the quality of NMT systems when it comes to translating literary creative challenges. The studies presented below serve as the methodological and theoretical framework for our study.

First, Şahin and Gürses [10] explored the effects of using MT to retranslate literary texts in terms of creativity. They conducted a study on translating a literary novel in the English–Turkish language pair by undergraduates in translation both with and without the aid of MT systems. After analysing multiple translation units, they concluded that MT is likely to block creativity among novel translators.

To the best of our knowledge, Guerberof-Arenas and Toral’s studies [11,12] were pioneers in exploring the role of creativity in MT based on textual elements in novels. Both studies were carried out within the framework of the CREAMT project (Creativity and narrative engagement of literary texts translated by translators and neural machine translation (<https://cordis.europa.eu/project/id/890697>, accessed on 15 May 2024), which focused on HT, MT, and PE. Their first study used the English > Catalan language combination, and then, the second also added the English > Dutch directionality. The first study tested the impact of different translation modalities in the user’s reading experience, considering the creative factor. The findings revealed that HT and PE showed a similar reading experience, but HT was better with creative shifts. The second study mainly focused on creativity (i.e., novelty and acceptability). The results showed that neither the literary trained NMT system

nor the PE output achieved satisfactory quality for translating creative elements. Finally, both studies highlighted that NMT systems can even constrain translators' creativity, and that better results are achieved when professional literary translators are involved.

In this regard, Webster et al. [9] compared the output of Google Translate and DeepL in translating classic novels from English into Dutch. They concluded that HT fairly surpassed NMT output. In fact, NMT rendered many errors and showed lack of creativity and diversity whereas HT proved a richer style. Finally, they stated that NMT systems can be helpful as an aid during the translation process.

Finally, Noriega-Santiáñez and Corpas Pastor [13] studied the quality of three NMT systems in translating formal neologisms against the HT made from students of the degree in translation and interpreting. Although NMT systems unsurprisingly failed to surpass the creativity of HT, students used a bunch of different technologies, including NMT systems, to tackle the creative challenges of literary translation in the English > Spanish directionality.

Manipulated MWEs

In this technological scenario on creativity and literature, our study addresses manipulated MWEs, specifically idioms, in a fantasy novel. Thus, this section particularly delves into phraseological variability and its connection to creative challenges.

García Campos [40] detailed three linguistic levels in a fantasy novel: (1) the morphosyntactic features of the source language and the style singularities of the author; (2) the specialised language (resulting from the author's documentation to set the novel scenario and plot); and (3) the author's creativity to name and invent certain elements (e.g., creatures, objects, sublanguages, etc.). In addition, there are certainly heterogeneous phraseological challenges involved in translating any fantastic work [36]. For that reason, mastery of phraseology is crucial for literary translators, who face numerous challenges that test their skills [41].

Monti et al. [42] (p. 3) defined MWEs as «meaningful lexical units made of two or more words in which at least one of them is restricted by linguistic conventions in the sense that it is not freely chosen». MWEs entail a series of difficulties due to their pragmatic, idiomatic, metaphorical, phonetic, and/or cultural load [43,44]. Indeed, the backbone features of the phraseological essence encompass fixity, idiomaticity, and plurilexicality [45]. In addition, MWEs have little syntactic and semantic transparency, but have a high degree of lexicalisation and conventionality [46]. Given their highly idiomatic nature, the meaning of these units is sometimes unpredictable without a given context. Furthermore, some other features should be considered. For instance, the psycholinguistic mechanisms, speaker manipulations, and the metaphorical and cultural meaning presented in many of these units [44].

This linguistic dynamism gives rise to phraseological variability, which is also a symptom of creativity [47]. According to Corpas Pastor and Mena Martínez [48], variables can be systematic or occasional. Against this background, this study revolves around manipulated MWEs. Despite different definitions within phraseology studies, discontinuity can be defined as the deliberate manipulation or creative modification of MWEs for semantic, stylistic, and pragmatic purposes [49] and [50] (p. 47). These units defy linguistic norms and yet remain anchored in the language system, as they must be comprehensible to the receiver in order to ensure communication, whether for expressive or stylistic purposes [51,52]. In addition, these units must be novel and unusual [51]. Phraseological manipulation can be classified in two different categories [49]: (1) *internal manipulation*: involving formal structural changes (morphological, lexical, or syntactic) visible in its constituents; or (2) *external manipulation*: without visible alterations in the canonical form. In fantasy novels, these units are manipulated in order to adapt the realities of the imaginary world to popular expressions in the target language. In other words, the author manages to immerse the reader into the novel's world by adding a novel element or a metaphorical twist into a canonical MWE [36].

Thus, translating these units implies a strong linguistic and communicative competence, i.e., not only to achieve semantic and formal equivalence, but also functional equivalence [45]. Hence, a process of encoding and decoding the message takes place [45]. This paper focuses on MWEs (specifically idioms) that have been internally manipulated in literary texts. To translate idioms, Corpas Pastor [53] outlines a comprehensive approach that begins with first identifying the idiom, followed by interpreting it within context, and concluding with conveying its pragmatic and semantic meaning in the target language.

3. Data Selection and Methodology

As stated before, this study intends to shed some light on the evaluation of technologies, specifically MT systems, to translate manipulated idioms.

This section outlines the protocolised methodology, describing the source text as well as the NMT systems and MWEs selected. Finally, it delves into the human evaluation and creative metrics proposed.

3.1. Selection of the Novel and NMT Systems

To select the source text for the study, i.e., *Black Bird* by Rebecca Roanhorse, several criteria have been considered. The source text should be a novel that adheres to the following criteria: (1) is a representative and recently (no more than 5 years) published fantasy novel; (2) is the first book in a saga; (3) is written in English; and (4) has not been translated into Spanish or French to date (as far as we are aware).

Regarding the selection of the NMT systems, this study particularly assesses the output of DeepL, Google Translate, Bing Translator, and Reverso. These general-domain NMT systems are under review for their wide use, not only in the post-editing labour market but also in a number of studies that applied NMT to literature (see [5,9,13,36], to name but a few).

3.2. Selection and Translation of Manipulated MWEs

A total of five manipulated MWEs, particularly idioms, were manually extracted from the novel. These were chosen because of their representativeness and their creative essence. These units are all visibly manipulated, i.e., these experienced formal changes in their canonical form.

With the help of some reference monolingual dictionaries of English, namely the Oxford Dictionary (Accessed online at <https://www.oed.com/>, accessed on 15 May 2024), Cambridge Dictionary (Accessed online at <https://dictionary.cambridge.org/>, accessed on 15 May 2024), and Merriam-Webster Dictionary (Accessed online at <https://www.merriam-webster.com/>, accessed on 15 May 2024), the canonical forms of the manipulated MWEs were first identified and then classified in index cards within a Word document. The table below (Table 1) shows the context, the manipulated MWE, and its canonical form.

All selected manipulated idioms represent imaginary realities from the book, encompassing the author's view of the world created ad hoc. These units have been extracted in context to be coherently translated and evaluated. The context of each was carefully selected by the following criteria: (1) if the manipulated MWE was part of a dialogue, the whole speech has been extracted; and (2) if the manipulated MWE was isolated in the narrative, only the sentence in which it is found has been extracted.

On the one hand, some of the selected units replace a known element of the canonical unit with an ad hoc created reality (as in the case of MWEs 2 and 5). On the other hand, the other selected units replace an element of the canonical unit with another common term (as in MWEs 1, 3, and 4).

Once all manipulated idioms were classified, they were translated by a professional literary translator (i.e., a human translation) and by the four selected NMT systems. Then, all four translations were included in evaluation rubrics (described in Section 3.4, below).

Table 1. Manipulated MWEs.

No.	Context	Manipulated MWEs	Canonical Form
1	“I would not miss it for all the stars in the sky,” xe said, and that time she caught the contempt plain enough.	would not miss it for all the stars in the sky	wouldn’t miss it for the world
2	And what in all the hells is a Kuharan?	what in all the hells	What the hell
3	“Thank the lesser gods for that, at least,” she said, and meant it.	Thank the lesser gods	thank God/goodness/heaven(s)/the Lord
4	“Just a rain cough. I swear to the deep. I wouldn’t take a commission if it was worse. I don’t need the pay.”	swear to the deep	swear to God
5	She must have hit someone, but for all the cacao in Cuecola, she couldn’t remember who.	for all the cacao in Cuecola	not for all the tea in China

3.3. Measuring Creativity

We followed the path of similar studies such as the ones Guerberof-Arenas and Toral [11,12] carried out. They defined creativity as “a combination of novelty (i.e., new, original) and acceptability (i.e., something of value, fit for purpose)” [54] (p. 78). Based on their long-tested classification, we proposed the way to measure creativity in manipulated MWEs, as follows:

- *Acceptability*: If the manipulated MWE complies with some of the three parameters of equivalence proposed by Corpas Pastor [49], as shown in the table below (Table 2): (1) morphosyntactic; (2) semantic; and (3) pragmatic.

Table 2. Corpas Pastor’s parameters of equivalence [49].

Parameters	
Morphosyntactic	Complementation, sentence function, and transformations
Semantic	Phraseological meaning, base image, and lexical composition
Pragmatic	Cultural competence, diasystematic constraints, frequency of use, discursive aspects, and implicatures

Due to the degree of discrepancy in cross-linguistic translation of a MWE from a source language to a target language, the notion of “functional phraseological equivalence” has been developed. This notion comprises a series of parameters of cross-linguistic comparison in order to measure the degree of equivalence [55] (p. 117). Thus, these are the different parameters that will be used to measure the extent to which NMT systems are able to convey the degree of equivalence in MWEs, considering three degrees (full, partial, or zero equivalence) [53].

- *Novelty*: The literary translators will be asked to evaluate the MWE’s degree of originality and innovation generated by the NMT system through a 5-point Likert scale, as follows:
 1. No novelty;
 2. Minor novelty;
 3. Neutral novelty;
 4. Moderate novelty;
 5. High novelty

In addition, a general parameter has been included for the evaluator to give an overall mark on the MWE translation, i.e., to assess to what extent it works in the target context.

Creativity score: Based on the formula used by Guerberof-Arenas and Toral [12], we proposed the following way to measure creativity in the translation of each MWE:

$$\left(\frac{\text{Morphosyntactic} + \text{Semantic} + \text{Pragmatic}}{3} \right) \times \frac{1}{2} + \frac{\text{Novelty}}{2} = \text{creativity score in MWEs}$$

The general parameter has been removed, as it has been used to get an overview of the perception of MWE in context. As acceptability encompasses morphosyntactic, semantic, and pragmatic parameters, the average of these three has been calculated to represent acceptability. However, as Guerberof-Arenas and Toral [12] pointed out, creativity comprises novelty and acceptability. Therefore, we consider that both parameters should be of equal value. To summarise, the percentage obtained from the human evaluation on each parameter is calculated by using this formula to score creativity in manipulated MWEs.

3.4. Human Evaluation: Profiles and Rubric

The human evaluation was carried out by six professional literary translators. To this end, a human evaluation sheet was created, which is anonymised so as not to collect personally identifiable information from the evaluators. This evaluation sheet was divided into four different parts: (1) the informed consent, which included the aim of this study, the journal where it will be published, the data protection law regulating it, and the right to participate voluntarily; (2) Section 1, which encompasses the demographic data; (3) Section 2, which includes the evaluation rubrics detailed below; and (4) Section 3, which contains the post-evaluation questions.

Once agreed to participate in the study, the participants filled in a questionnaire about their profiles and professional experience in Section 1. Their answers are summarised in Table 3.

Table 3. Literary translators' profiles.

Question	Answer
Years old	26–57
Years of experience as a literary translator	2–28
Mother tongue(s)/First language(s)	Spanish
Second language(s)	English/French/German/Portuguese
Number of translated novels	1–149
Language pair(s) in which you translate novels	English > Spanish/French > Spanish/German > Spanish

According to the profile data collected, most of the participants were aged between 26 and 30 years. Translators were selected according to the following criteria: (1) they have at least two years of professional experience; (2) they have translated at least one novel in the English > Spanish language pair; (3) they have Spanish as a first working language (or mother tongue) and English as a second language; and (4) they have a university degree in translation and interpreting and are specialised in literary translation.

Then, in Section 2 they could find the evaluation rubrics (see Table 4). Thus, the human evaluators assessed the translated MWEs in terms of creativity by filling them.

The evaluation rubrics in Section 2 are divided into two colours: blue and grey. The blue section contains the following items: (1) the extracted manipulated MWE; (2) the context where the MWE appears; (3) the extra context, i.e., a manual clarification by the authors so that the evaluator can fully understand why the MWE is manipulated, providing data related to the type of character involved or the socio-cultural or religious context in which the MWE appears in the book; (4) the canonical form of the MWE, including a link to its meaning on the online reference dictionary of English; and (5) the output of both the

HT and the four NMT systems, without specifying which one is which to not influence the results.

Table 4. Example of the evaluation rubric.

MANIPULATED MWE							
Context							
Extra-context							
Canonical form							
Translation method/system	Target language (Spanish)	Parameters	Evaluation				
			Likert				
			1	2	3	4	5
1		Morphosyntactic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Semantic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Pragmatic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		Novelty	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
		General	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The grey section is where the evaluators include their assessment. To simplify the evaluation, we established a 5-point Likert scale. Thus, the literary translators were asked to evaluate both the degree of *acceptability* and *novelty* of the MWE based on 5 points by following the instructions provided within the same evaluation sheet. The *acceptability* parameter was divided into the three parameters of equivalence previously explained.

Finally, once the evaluation had been completed, the evaluators were faced with a series of questions organised within the evaluation sheet. The post-evaluation questions can be found in Table 5. The nature of these questions was to find out their opinions regarding the evaluation and the use of NMT systems.

Table 5. Post-evaluation questions.

Post-Evaluation Questions
1. To what extent do you think NMT systems can help you to translate a manipulated MWE?
2. To what extent do you think NMT can constrain your creativity if you were to translate a MWE?
3. To what extent do you think literary translators should use NMT systems?
4. Please use this space below to add any comments on the evaluation that you consider relevant.

These first three questions were planned to be answered on a 5-point Likert scale as well. The last one was an open-ended question for evaluators to provide further comments or suggestion they may deem relevant for the evaluation. The literary translators' answers served to evaluate their perception about the usefulness of NMT systems to address creative challenges.

4. Results

This section summarises the results from Sections 2 and 3 in the evaluation sheet, presenting the manipulated MWEs output in terms of creativity as well as the post-evaluation questions.

4.1. Manipulated MWEs Evaluation

The evaluation of the manipulated idioms was carried out by six professional literary translators. To this end, each participant assessed the HT and the NMT output by using

a 5-point Likert scale. At the beginning of this subsection, the results obtained are presented. Finally, creativity was averaged per each translation modality by using the formula previously explained. This subsection covers the first and second research questions of our study.

4.1.1. NMT Systems against HT

The figures below (Figures 1–5) analyse the degree of acceptability of each parameter (morphosyntactic, semantic, pragmatic, novelty, and general) obtained by each of the manipulated MWEs. The average was calculated as a percentage of the rating, with 100% being the maximum score. In other words, each participant’s answers were processed, then the total score obtained for each parameter using each translation method was added up and then averaged. As the Likert scale ranked from 1 to 5, the results are adjusted to obtain a fair percentage.

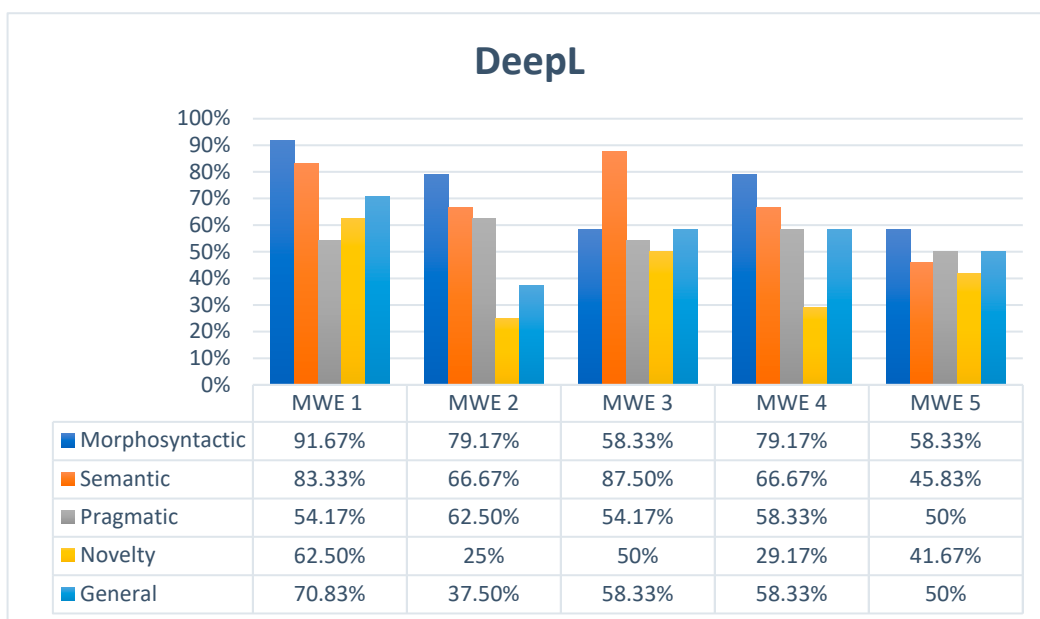


Figure 1. DeepL performance.

DeepL is the first NMT system to be under review by participants. As can be seen, it renders prominent results in terms of morphosyntactic and semantic parameters, but it notably falls short in novelty, especially in MWE 2 and MWE 4. However, MWE 1 scores the best result in novelty among the other NMT systems’ outputs. But, despite not having a high degree of novelty, MWE 4 is still acceptable to the professional literary translators in general terms. Conversely, since DeepL performs poorly for MWE 2 in terms of novelty, the evaluators consider that the output is generally unacceptable. Finally, on a pragmatic level, this NMT system is fairly average.

Then, the next NMT system to be evaluated was Google Translate. It performed similarly to DeepL in terms of morphosyntactic and semantic parameters, scoring between 70% and 80% on average. However, it is slightly better on a pragmatic level in some MWEs, namely MWE 1 and MWE 3. In contrast, Google Translate was less novel than DeepL for MWE 1 and MWE 2, but rendered better results for MWE 4 and MWE 5. Despite its similarities in performance, Google Translate was generally less acceptable than DeepL among the evaluators.

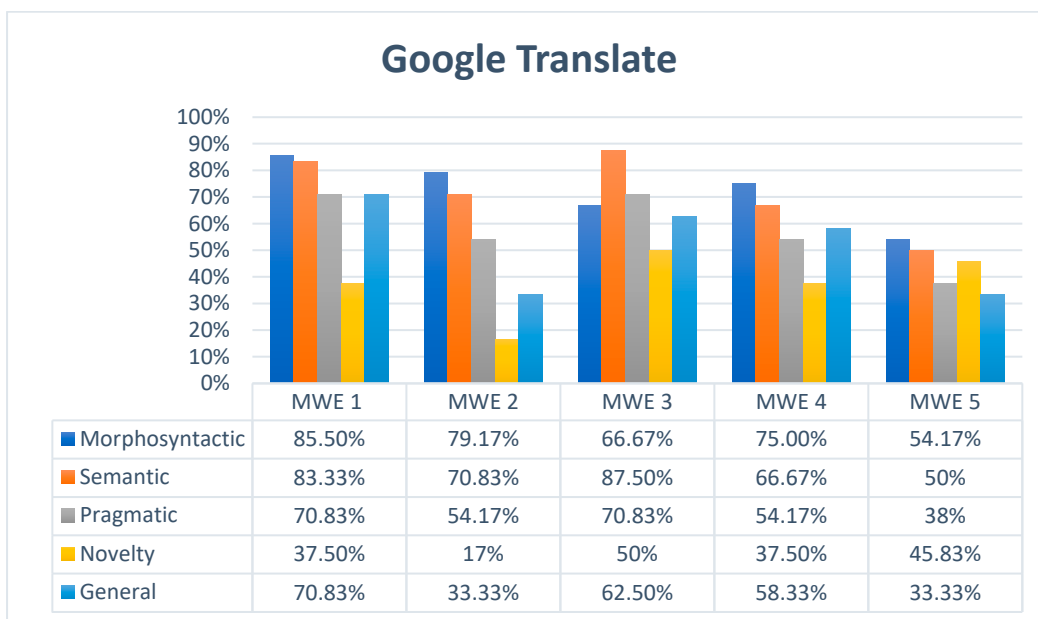


Figure 2. Google Translate performance.

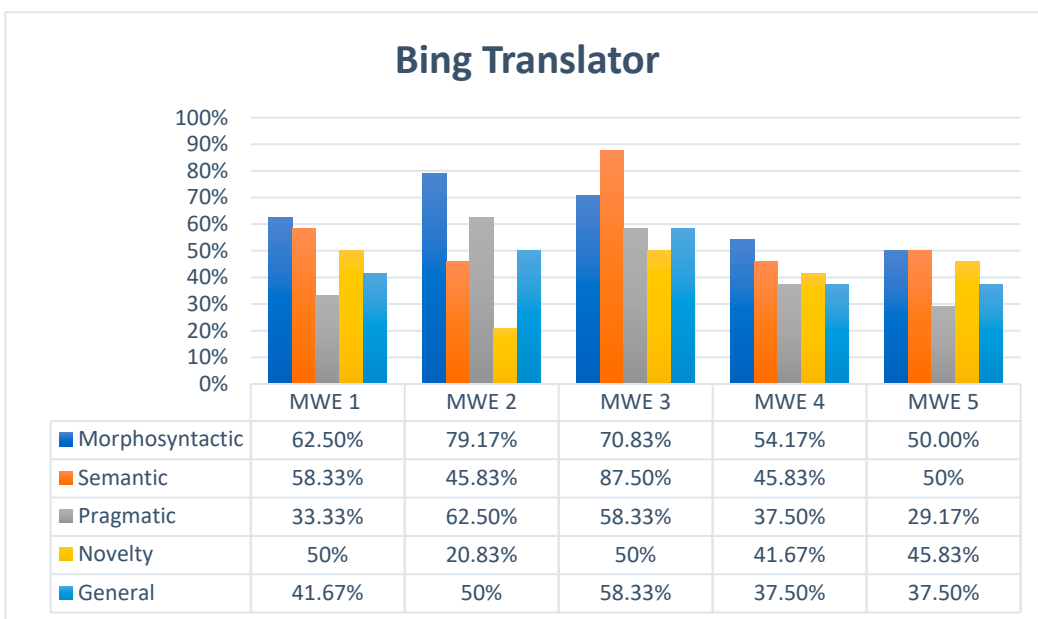


Figure 3. Bing Translator performance.

Bing Translator was actually the fourth translation method to be under review in the evaluation rubric, right after the HT. Both morphosyntactically and semantically, this NMT system performed worse than DeepL and Google Translate. Indeed, it did not achieve satisfactory results in terms of semantic parameters for MWE 2 and MWE 3. However, our results reveal that Bing Translator output meets the novelty parameter better than Google Translate, although it still falls short of DeepL’s performance for some MWEs, notably MWE 1 and MWE 2.

Finally, Reverso was the last NMT system to be evaluated among literary translators. This NMT system showed the greatest variability in performance. Specifically, it demonstrated a more than satisfactory level of several parameters, as in the case of MWE 4. However, in other manipulated idioms, as in MWE 1, it was still far below the morphosyntactic or pragmatic levels achieved by previous NMT systems. Regarding the novelty

parameter, it usually performed better than Google Translate, but behind DeepL and Bing Translator in some MWEs. Indeed, three out of five MWEs did not reach 50% of acceptance.

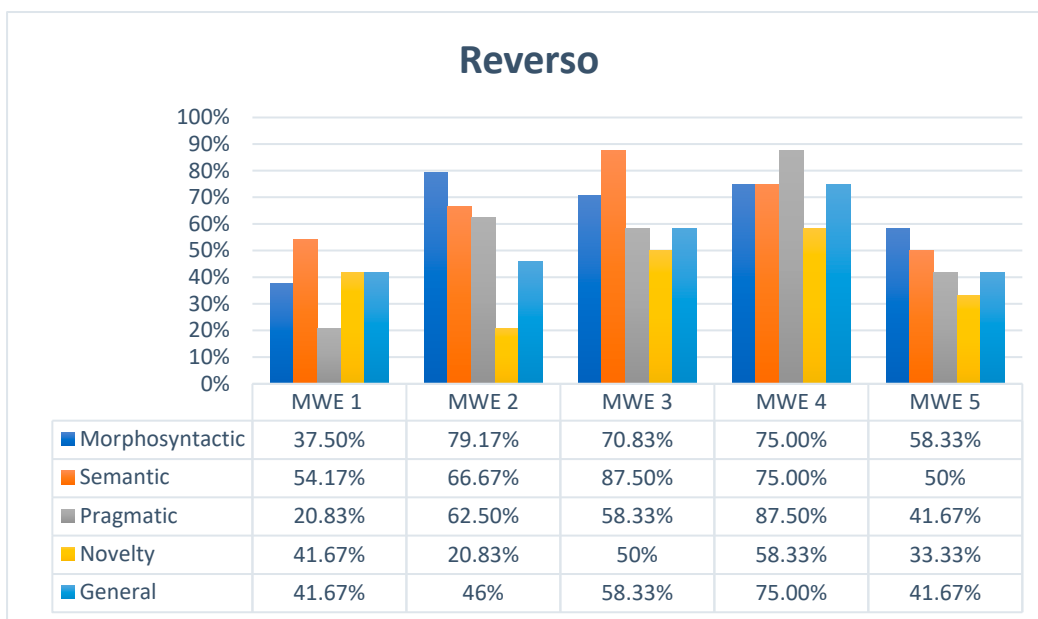


Figure 4. Reverso performance.

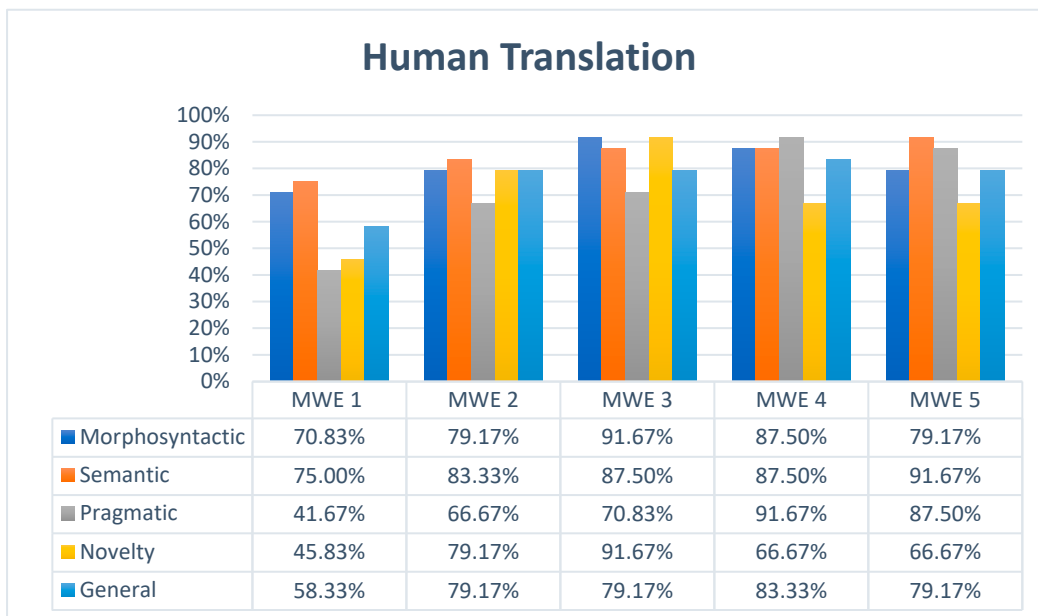


Figure 5. Human translation performance.

In this study, human translation was used as a gold standard to compare the level of creativity that these NMT systems can achieve. In general, HT was much better in all parameters than NMT systems, with overall average scores between 70% and 80%. However, there was a surprisingly lower score in MWE 1, which may be due to an error of the human translator, as it did not achieve the quality standards at a pragmatic and novelty level. In fact, it failed to reach the performance of several NMT systems, namely DeepL and Google Translate. However, the degree of novelty was very high compared to NMT systems, especially in MWE 2 and 3.

4.1.2. NMT and HT Creativity

After analysing the performance of each translation method, creativity is calculated. Thus, we can evaluate to what extent NMT systems can achieve human creativity. The results can be found in Figure 6.

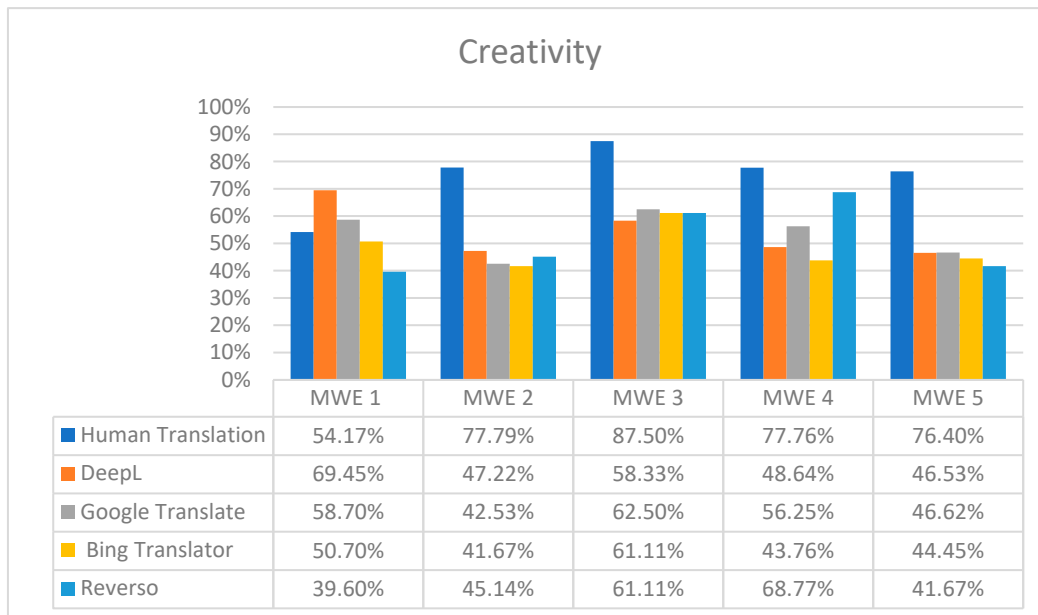


Figure 6. Creativity in HT and NMT systems.

As can be seen in the figure above, HT (74.72% on average) highly surpassed all NMT systems in almost all MWEs, thus unsurprisingly being the most creative translation method. Indeed, HT output was not comparable in most of the instances with NMT, as it generally showed a high level of creativity, specifically novelty.

Regarding NMT systems, the most creative in this pilot study was DeepL (54.03%), closely followed by Google Translate (53.32%). Then, Reverso slightly outperformed (51.26%) Bing Translator (48.34%). However, the scores of all four translation systems were fairly similar.

4.2. Post-Evaluation

This subsection shows the results obtained in the post-evaluation questions answered by the six participants of this study (cf. the third research question of our study). The figures below (Figures 7–9) collect their answers. Each colour represents the Likert results (from 1 to 5), as shown in the legend.

According to the data collected, most respondents were reluctant to use NMT systems to help them translate manipulated MWEs. However, we can see diverse opinions among them, but no one believed that NMT systems should always be used. In fact, most of their answers fell between 1 and 3 on a 5-point Likert scale.

Regarding the second question, half of the participants (three evaluators) believed that NMT systems can constrain literary translators’ creativity when it comes to translating manipulated MWEs. The others are less inclined to use these systems, but no one marked less than a 3 on the 5-point Likert scale.

Finally, there was significant discrepancy among the evaluators’ responses to the final question. However, they unanimously agreed that NMT systems are not to be always used in literature. The majority (two evaluators) indicated that NMT systems could be used occasionally, while the remaining two stated that these should never be used.

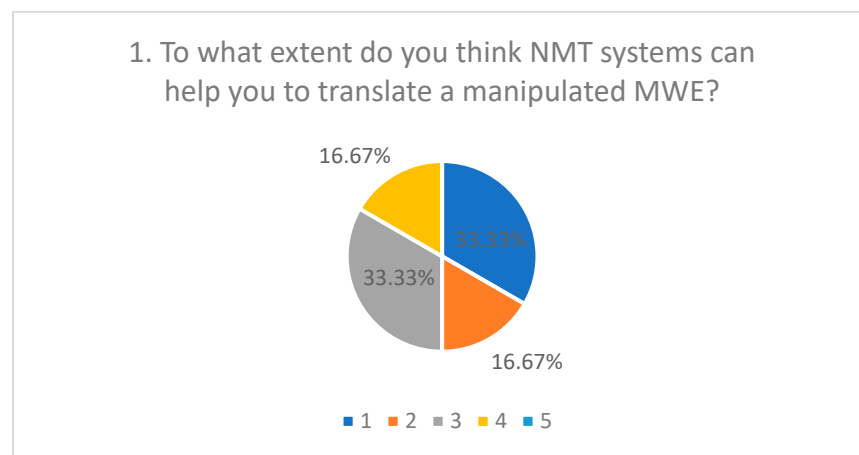


Figure 7. Answers to question 1.

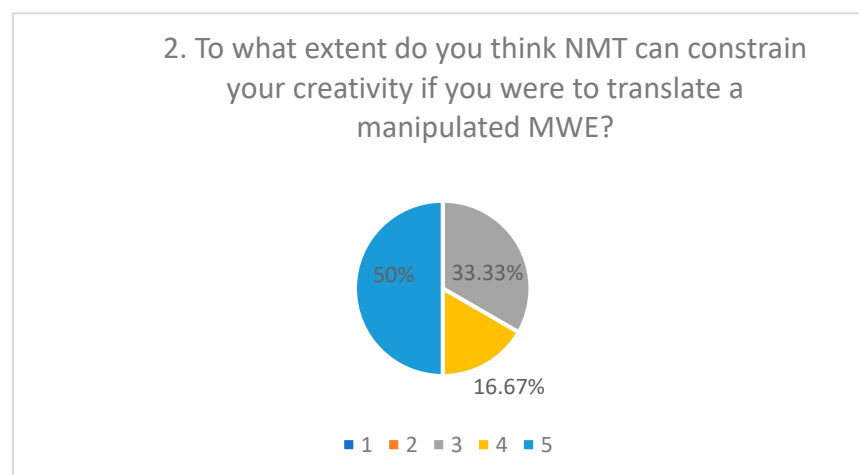


Figure 8. Answers to question 2.

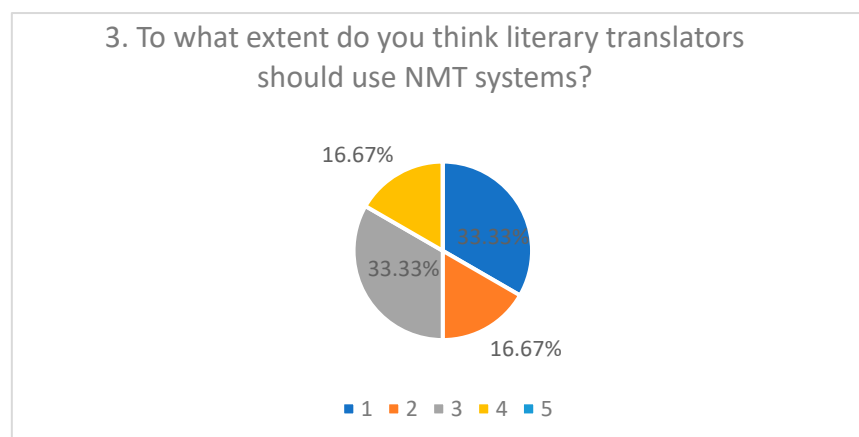


Figure 9. Answers to question 3.

5. Discussion

The findings of this pilot study are focused on the three main questions outlined earlier, which will be discussed in relation to previous studies:

- How do NMT systems and HT perform in translating manipulated MWEs, according to the proposed equivalence parameters?

- To what extent can these NMT systems be compared to a professional HT in terms of creativity?
- What is the opinion of literary translators with regard to integrating NMT systems into their workflow to translate literature, in general, and manipulated MWEs, in particular?

Regarding the first research question, our results show that Google Translate is slightly more accurate (in terms of morphosyntactic, semantic, and pragmatic parameters) for translating manipulated idioms in literature compared to DeepL. This result contradicts some of the findings of Webster et al. [9] regarding the NMT performance of some creative challenges in literature, in which Google Translate fell short of accuracy. However, DeepL general parameter scores were better in our study, which is in line with the results obtained by Noriega-Santiáñez and Corpas Pastor [13] when it comes to translating creative MWEs. Concerning the other NMT systems, Bing Translator performed significantly worse than Google Translate on the pragmatic parameter, and also failed to reach DeepL's results. This finding corroborates Brusasco [5]'s, as the author found that DeepL, Google Translate, and Bing Translator needed more editing regarding pragmatic adequacy. Furthermore, although Reverso excelled in novelty, it still performed worse compared to the other NMT systems at a morphosyntactic or pragmatic level. These findings partly support Ibrahim and Alkhawaja's [32]'s study, as they stated that Google Translate slightly outperformed Reverso in terms of quality.

In contrast, human translation mainly outperformed NMT systems in almost all parameters, followed closely by DeepL and Google Translate when it comes to morphosyntactic and pragmatic parameters. HT also stood out in the novelty parameter. This finding corroborates some conclusions reached by Guerberof and Toral [11,12], as they pointed out that NMT could not achieve a satisfactory level of novelty in creative shifts, such as idioms, among others. Furthermore, HT demonstrated a markedly higher degree of acceptance in the general parameter among literary translators than any other NMT system, which is in line with some results that involved HT against NMT [10,11,13,26].

Regarding the second research question, HT unsurprisingly outperformed NMT output in virtually all MWEs in terms of creativity. These findings are in line with the study by Guerberof-Arenas and Toral [11,12], as they noticed that to translate creatively, human involvement is needed (whether in the form of postediting or translating from scratch). In addition, some other authors such as Castilho and Resende [30] or Şahin and Gürses [10] agreed that HT is much more creative than MT or even PE. Concerning NMT systems' performance, both Google Translate and DeepL reached similar results, which was also noticed in Van Egdome et al.'s [35] study, and both performed far better than Reverso or Bing Translator. Thus, our findings are in line with previous studies that concluded that MT renders far worse results than HT [13,25,28]. In fact, Brusasco [5] pointed out that current systems cannot achieve human experience.

However, there is a certain discrepancy in our findings, as it seems that some NMT systems are not far ahead of HT in terms of acceptability parameters. For example, MWE 1 was the only instance that HT output fell short of DeepL or Google Translate's performance results. Indeed, our study demonstrates that NMT systems show better results when the manipulated idiom is formed by common vocabulary, hence the NMT systems might process it better. In fact, Zajdel [33] reached a similar conclusion when comparing MT against HT in metaphors, as the study showed that HT specially rendered better results with multi-word metaphors.

Finally, regarding the last research question, the post-evaluation results generally show that most literary translators in our study were reluctant to integrate NMT systems into their workflow, which is in line with the larger-scale findings reached by Ruffo [6,21]. Indeed, they hold the same view as the participants of these studies, i.e., they truly believe that NMT systems are still not able to convey all the literary challenges, but not all of them were completely against them. This partial acceptance might occur due not only to the current challenges that NMT systems face in dealing with literary texts [19], but also to the difficulty of translating MWEs [41,49], especially discontinuous or manipulated MWEs [36].

In fact, half of these professional literary translators agreed that NMT might constrain creativity, which corroborates what Kenny and Winters [38] and Şahin and Gürses [10] stated in their research. In addition, some of the literary translators that participated in this study pointed out that NMT systems recycle data, thus these cannot help to create anything new. Hence, contrary to other studies that showed translation students sometimes use NMT systems for creative challenges [13], the findings of this study suggest that professional literary translators are inclined to reject this practice.

Finally, in the open-ended question, the literary translators also pointed out that there are some ethical issues that should be considered, some of which were also examined by Taivalkoski-Shilov [20] and Kenny and Winters [38]. For instance, NMT systems might use data that can infringe authors' and translators' copyrights or can even constrain literary translators' creativity.

6. Conclusions

To the best of our knowledge, this article presents one of the few studies that deals with creativity in NMT applied to manipulated MWEs present in literary texts. In addition, it tentatively proposes a formula to score the degree of creativity based on a human evaluation.

This pilot study reinforces the idea that NMT systems, especially DeepL, Google Translate, Bing Translator, and Reverso, are still unable to handle all creative phraseological challenges found in literary texts, specifically manipulated idioms. In fact, this study presents some evidence of how creativity is such an intrinsic human characteristic that it is highly difficult to reproduce in novel contexts by a NMT system. Despite some instances that NMT systems perform satisfactorily, NMT are not near as good as HT in terms of creativity, especially if the MWE incorporates any term created ad hoc.

However, DeepL or Google Translate showed potential, notably when dealing with morphosyntactic or semantic parameters, and even received acceptable overall scores in the human evaluation. Nevertheless, these did not generally achieve a satisfactory degree of creativity, as the complexity of manipulated MWEs requires human intervention, either to revise the NMT of a MWE, to post-edit the unit, or to translate it from scratch.

In addition, this study gives voice to a small group of professional literary translators to express their opinions. Regardless of age or experience, their answers point to not always using NMT systems in literature. However, there is still a level of disagreement as to whether these systems could be helpful or, instead, could lead to constraining the creative genius of literary translators.

However, our findings cannot be easily extrapolated, as this is a pilot study with several limitations. For instance, both the number of creative challenges extracted and the number of participants were limited. In addition, there was an imbalance of literary translators' profiles that could lead to bias in the study. For that reason, following our promising preliminary results, we intend to expand our pilot study in various ways. For instance, we intend to add more language pairs (e.g., English > French or English > Italian), more creative phraseological challenges (such as metaphors, puns, etc.), and more participants, including both student and professional translators. In addition, we would like to compare MTPE output evaluation with the HT and the NMT output evaluation.

Finally, our study contributes to expanding the on-going discourse of human-centred machine translation. In particular, studying the link between technologies and creativity in literature could help professionals to integrate them into their workflow and raise awareness about misuses or unethical practices in the field.

Author Contributions: Conceptualization, G.C.P. and L.N.-S.; methodology, G.C.P. and L.N.-S.; validation, G.C.P. and L.N.-S.; formal analysis, G.C.P. and L.N.-S.; investigation, L.N.-S.; resources, G.C.P.; data curation, L.N.-S.; writing—original draft preparation, L.N.-S.; writing—review and editing, G.C.P.; visualisation, L.N.-S.; supervision, G.C.P.; project administration, G.C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by a predoctoral contract granted by the University of Malaga and it has been carried out in the framework of several research projects: “Multi-lingual and Multi-domain Adaptation for the Optimisation of the VIP system” (VIP II, ref. no. PID2020-112818GB-I00/AEI/10.13039/501100011033, 2021–2025, Spanish Ministry of Science and Innovation), and “Multilingual machine interpretation for COVID-19 cases in emergency departments” (RECOVER, ref. ProyExcel_00540, 2022–2025, Andalusian Regional Government).

Institutional Review Board Statement: The study was approved by the Ethics Committee, Faculty of Arts and Philosophy, University of Malaga (25 July 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Dataset available on request from the authors.

Acknowledgments: The authors would like to thank the professional literary translators who have participated in this research.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Corpas Pastor, G.; Bautista Zambrana, M.R.; Hidalgo-Ternerero, C.M. *Sistemas Fraseológicos en Contraste: Enfoques Computacionales y de Corpus*; Comares: Granada, Spain, 2021.
2. Rothwell, A.; Way, A.; Youdale, R. *Computer-Assisted Literary Translation*; Routledge: New York, NY, USA, 2024.
3. O’Brien, S. Translation as human–computer interaction. *Transl. Spaces* **2012**, *1*, 101–122. [CrossRef]
4. Hadley, J.L.; Taivalkoski-Shilov, K.; Teixeira, C.S.C.; Toral, A. *Using Technologies for Creative-Text Translation*; Routledge: New York, NY, USA, 2022.
5. Brusasco, P. Pragmatic and cognitive elements in literary machine translation: An assessment of an excerpt from J. Polzin’s Brood translated with Google, DeepL, and Bing Translator. In *Using Technologies for Creative-Text Translation*, 1st ed.; Hadley, J.L., Taivalkoski-Shilov, K., Teixeira, C.S.C., Toral, A., Eds.; Routledge: New York, NY, USA, 2022; pp. 139–160.
6. Ruffo, P. Collecting literary translators and narratives: Towards a new paradigm for technological innovation in literary translation. In *Using Technologies for Creative-Text Translation*; Hadley, J.L., Taivalkoski-Shilov, K., Teixeira, C.S.C., Toral, A., Eds.; Routledge: New York, NY, USA, 2022; pp. 18–39.
7. Besacier, L.; Schwartz, L. Automated translation of a literary work: A pilot study. In Proceedings of the NAACL-HLT Fourth Workshop on Computational Linguistics for Literature, Denver, CO, USA, 4 June 2015; Kazantseva, A., Szpakowicz, S., Koolen, C., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2015; pp. 114–122.
8. Toral, A.; Wieling, M.; Way, A. Post-editing Effort of a Novel With Statistical and Neural Machine Translation. *Front. Digit. Humanit.* **2018**, *5*, 9. [CrossRef]
9. Webster, R.; Fonteyne, M.; Tezcan, A.; Macken, L.; Daems, J. Gutenberg Goes Neural: Comparing Features of Dutch Human Translations with Raw Neural Machine Translation Outputs in a Corpus of English Literary Classics. *Informatics* **2020**, *7*, 32. [CrossRef]
10. Şahin, M.; Gürses, S. Would MT kill creativity in literary retranslation? In Proceedings of the Qualities of Literary Machine Translation, Dublin, Ireland, 19 August 2019; pp. 26–34.
11. Guerberof-Arenas, A.; Toral, A. The Impact of Post-editing and Machine Translation on Creativity and Reading Experience. *Transl. Spaces* **2020**, *9*, 255–282. [CrossRef]
12. Guerberof-Arenas, A.; Toral, A. Creativity in translation: Machine translation as a constraint for literary texts. *Transl. Spaces* **2022**, *11*, 184–212. [CrossRef]
13. Noriega-Santiáñez, L.; Corpas Pastor, G. Machine vs Human Translation of Formal Neologisms in Literature: Exploring E-tools and Creativity in Students. *Rev. Tradumàtica. Tecnol. De La Traducció* **2023**, *21*, 233–264. [CrossRef]
14. García Álvarez, A.M. Reflexiones sobre la creatividad en la enseñanza de la traducción literaria. In *El Viaje de la Literatura: Aportaciones a Una Didáctica de la Traducción Literaria*; Fortea, C., Ed.; Cátedra: Madrid, Spain, 2018; pp. 13–32.
15. PETRA-E Framework of Reference for the Education and Training of Literary Translators. Available online: <https://petra-educationframework.eu/> (accessed on 8 May 2024).
16. Roanhorse, R. *Black Sun*; Rebellion Publishing Ltd: Oxford, UK, 2021.
17. Carl, M.; Braun, S. Translation, interpreting and new technologies. In *Routledge Handbook of Translation Studies and Linguistics*; Malmkjaer, K., Ed.; Routledge: Brixham, UK, 2018; pp. 374–390.
18. Bowker, L.; Corpas Pastor, G. Translation Technology. In *The Oxford Handbook of Computational Linguistics*, 2nd ed.; Mitkov, R., Ed.; Oxford University Press: Oxford, UK, 2022; pp. 871–905.
19. Toral, A.; Way, A. Is Machine Translation Ready for Literature? Proceedings of Translating and the Computer 36, London, UK, 27–28 November 2014; pp. 174–176.
20. Taivalkoski-Shilov, K. Ethical Issues Regarding Machine(-Assisted) Translation of Literary Texts. *Perspectives* **2018**, *27*, 689–703. [CrossRef]

21. Ruffo, P. Human-Computer Interaction in Translation: Literary Translators on Technology and Their Roles. In Proceedings of the 40th Conference Translating and the Computer, London, UK, 15–16 November 2018; pp. 127–131.
22. Way, A.; Rothwell, A.; Youdale, R. Why more Literary Translators should embrace Translation Technology. *Rev. Tradumàtica. Tecnol. Traducció* **2023**, *21*, 87–102. [\[CrossRef\]](#)
23. Declercq, C.; Van Egdom, G. No more buying cats in a bag? Literary Translation in the age of language automation. *Rev. Tradumàtica. Tecnol. Traducció* **2023**, *21*, 49–62. [\[CrossRef\]](#)
24. Voigt, R.; Jurafsky, D. Towards a Literary Machine Translation: The Role of Referential Cohesion. In Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature, Montréal, Canada, 8 June 2012; pp. 18–25.
25. Toral, A.; Way, A. Machine-Assisted Translation of Literary Text: A Case Study. *Transl. Spaces* **2015**, *4*, 240–267. [\[CrossRef\]](#)
26. Toral, A.; Way, A. What Level of Quality can Neural Machine Translation Attain on Literary Text? In *Translation Quality Assessment: From Principles to Practice*; Moorkens, J., Castilho, S., Gaspari, F., Doherty, S., Eds.; Springer International Publishing AG: Dublin, Ireland, 2018; pp. 263–287.
27. Brglez, M.; Vintar, Š. Lexical Diversity in Statistical and Neural Machine Translation. *Information* **2022**, *13*, 93. [\[CrossRef\]](#)
28. Moorkens, J.; Toral, A.; Castilho, S.; Way, A. Translators' Perceptions of Literary Post Editing using Statistical and Neural Machine Translation. *Transl. Spaces* **2018**, *7*, 240–262. [\[CrossRef\]](#)
29. Matusov, E. The Challenges of Using Neural Machine Translation for Literature. In Proceedings of the Qualities of Literary Machine Translation, Dublin, Ireland, 19 August 2019; pp. 10–19.
30. Castilho, S.; Resende, N. Post-Editese in Literary Translations. *Information* **2022**, *13*, 66. [\[CrossRef\]](#)
31. Toral, A.; Van Cranenburgh, A.; Nutters, T. Literary-Adapted Machine Translation in a Well-Resourced Language Pair: Explorations with More Data and Wider Contexts. In *Computer-Assisted Literary Translation*, 1st ed.; Rothwell, A., Way, A., Youdale, R., Eds.; Routledge: London, UK, 2023; pp. 27–52.
32. Ibrahim, H.; Alkhawaja, L. Comparative Evaluation of Neural Machine Translation of fiction literature: A case study. *J. Namib. Stud.* **2023**, *34*, 2806–2822.
33. Zajdel, A. Catching the meaning of words: Can Google Translate convey metaphor? In *Using Technologies for Creative-Text Translation*, 1st ed.; Hadley, J.L., Taivalkoski-Shilov, K., Teixeira, C.S.C., Toral, A., Eds.; Routledge: Abingdon, UK, 2022; pp. 116–138.
34. Dorst, A.G. Metaphor in Literary Machine Translation. In *Computer-Assisted Literary Translation*, 1st ed.; Rothwell, A., Way, A., Youdale, R., Eds.; Routledge: London, UK, 2023; pp. 173–186.
35. Van Egdom, G.; Kosters, O.; Declercq, C. The Riddle of (Literary) Machine Translation Quality: Assessing Automated Quality Evaluation Metrics in a Literary Context. *Rev. Tradumàtica. Tecnol. Traducció* **2023**, *21*, 129–159. [\[CrossRef\]](#)
36. Noriega-Santiáñez, L.; Corpas Pastor, G. La traducción del género fantástico mediante corpus y otros recursos tecnológicos: A propósito de 'The City of Brass'. *Moenia Rev. Lucence Lingüística Lit.* **2023**, *29*, 1–30. [\[CrossRef\]](#)
37. Kruger, A.; Wallmach, K.; Munday, J. *Corpus-Based Translation Studies: Research and Applications*; Continuum International Publishing Group: London, UK, 2011.
38. Kenny, D.; Winters, M. Machine translation, ethics and the literary translator's voice. *Transl. Spaces* **2020**, *9*, 123–149. [\[CrossRef\]](#)
39. Farrel, M. Machine Translation Markers in Post-edited Machine Translation Output. In Proceedings of the 40th Conference Translating and the Computer, London, UK, 15–16 November 2018; pp. 50–59.
40. García Campos, M. Niveles lingüísticos en la traducción de literatura fantástica: El ladrón cuántico, de Hannu Rajaniemi. *La Linterna Trad.* **2013**, *8*, 45–50.
41. Corpas Pastor, G. Detección, descripción y contraste de las unidades fraseológicas mediante tecnologías lingüísticas. In *Fraseo-pragmática*; Olza Moreno, I., Manero Richard, E., Eds.; Frank & Timme: Berlin, Germany, 2013; pp. 335–374.
42. Monti, J.; Seretan, V.; Corpas Pastor, G.; Mitkov, R. Multiword units in machine translation and translation technology. In *Multiword Units in Machine Translation and Translation Technology*; Mitkov, R., Monti, J., Corpas Pastor, G., Seretan, V., Eds.; John Benjamins Publishing Company: Amsterdam, The Netherlands, 2018; pp. 1–38.
43. Sinclair, J. Language and computing, past and present. In *Evidence-Based LSP: Translation, Text and Terminology*; Ahmad, K., Rogers, M., Eds.; Peter Lang: Bern, Switzerland, 2007; pp. 21–52.
44. Leal Riol, M.J. Contraste fraseológico: Similitudes y diferencias existentes entre las unidades fraseológicas del inglés y del español. *ES Rev. Filol. Inglesa* **2008**, *29*, 103–116.
45. Mena Martínez, F.; Sánchez Manzanares, C. Los usos creativos de las UF: Implicaciones para su traducción (inglés-español). In *Enfoques Actuales Para la Traducción Fraseológica y Paremiológica: Ámbitos, Recursos y Modalidades*; Conde Tarrío, G., Mogorrón Huerta, P., Martí Sánchez, M., Prieto García Seco, D., Eds.; Instituto Cervantes: Madrid, Spain, 2015; pp. 59–76.
46. Calzolari, N.; Fillmore, C.J.; Grishman, R.; Ide, N.; Lenci, A.; MacLeod, C.; Zampolli, A. Towards Best Practice for Multiword Expressions in Computational Lexicons. In Proceedings of the Third International Conference on Language Resources and Evaluation, Las Palmas, Spain, 29–31 May 2002; pp. 1934–1940.
47. Mena Martínez, F. En torno al concepto de desautomatización fraseológica: Aspectos básicos. *Tonos Digit.* **2003**, *5*.
48. Corpas Pastor, C.; Mena Martínez, F. Aproximación a la variabilidad fraseológica de las lenguas alemana, inglesa y española. *ELUA* **2003**, *17*, 181–201. [\[CrossRef\]](#)
49. Corpas Pastor, C. *Manual de Fraseología*, 1st ed.; Gredos: Madrid, Spain, 1996.
50. Zuluaga, A. Traductología y Fraseología. *Paremia* **1999**, *8*, 537–549.

51. Mellado Blanco, C. Parámetros específicos de equivalencia en las unidades fraseológicas (con ejemplos del español y el alemán). *Rev. Filol.* **2015**, *33*, 153–174.
52. Llopart-Saumell, E. Desautomatización fraseológica: De la norma a la creatividad. *CLINA Rev. Interdiscip. Traducción Interpret. Comun. Intercult.* **2020**, *6*, 119–136. [[CrossRef](#)]
53. Corpas Pastor, G. *Diez Años de Investigación en Fraseología: Análisis Sintáctico-Semánticos, Contrastivos y Traductológicos*; Vervuert: Madrid, Spain, 2003.
54. Guerberof-Arenas, A.; Valdez, S.; Dorst, A.G. Does training in post-editing affect creativity? *J. Spec. Transl.* **2024**, *41*, 74–97. [[CrossRef](#)]
55. Corpas Pastor, G. Fraseología y traducción. In *El Discurs Prefabricat: Estudis de Fraseologia Teòrica i Aplicada*; Liern, S., Piquer Vidal, A., Eds.; Universitat Jaume I: Valencia, Spain, 2000; pp. 107–138.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.