

ISBN: 978-84-1091-017-1 (PDF)

DOI: <https://doi.org/10.14201/OAQ0373641649>

# Poetic Machine Translation: An Oxymoron? The Case of Giacomo Leopardi into English

*Traducción poética automática: ¿un oxímoron? El caso de Giacomo Leopardi en inglés*

Letizia LEONARDI

*University of Aberdeen*

[letizialeonardi97@gmail.com](mailto:letizialeonardi97@gmail.com)

**ABSTRACT:** Despite the growing interest in AI, poetic machine translation remains a less explored area. Against this backdrop, this paper offers a quality assessment of automatic poetry translation on the Italian-English language pair. It uses a selection of poems by the 19<sup>th</sup>-century poet Giacomo Leopardi as a case study and compares the outputs produced by three freely accessible MT systems (Google, DeepL, Yandex) to determine which tool better performs with poetry and which are the linguistic areas calling for improvement. This paper does not suggest that AI should be used to translate poetry but rather demonstrates that, due to its stylistic and semantic complexities, poetic language can be used as a tool to improve MT efficiency. It concludes with some considerations on the impact that poetic machine translation may have on authorial representation.

**KEYWORDS:** poetry translation; Leopardi; machine translation.

**RESUMEN:** A pesar del creciente interés por la IA, la traducción automática de poesía sigue siendo un campo poco explorado. En este contexto, este artículo evalúa la calidad de la traducción automática de poesía en la combinación de idiomas italiano-inglés. Utiliza una selección de poemas del poeta del siglo XIX Giacomo Leopardi como caso de estudio y compara los resultados obtenidos por tres sistemas de traducción automática de libre acceso (Google, DeepL, Yandex) para determinar cuál ofrece mejores resultados con la poesía y cuáles son las áreas lingüísticas que requieren mejoras. Este artículo no sugiere que la IA deba utilizarse para traducir poesía, sino que demuestra que, debido a sus complejidades estilísticas y semánticas, el lenguaje poético puede utilizarse como herramienta para mejorar la eficacia de la TA. Concluye con algunas consideraciones sobre el impacto que la traducción automática poética puede tener en la representación autorial.

PALABRAS CLAVE: traducción de poesía; Leopardi; traducción automática.

## 1. INTRODUCTION

The field of MT has attracted increasing scholarly attention in the past few decades. However, the application of AI to poetry translation remains a less explored area. Literary machine translation has often been regarded as a contradictory concept (Voigt and Jurafski 2012). Its complexities, especially in the case of poetry, can be explained by pointing to two main factors. One is the unavailability of «open-sourced multilingual parallel poetic corpora» (Chakrabarty et al. 2021). The other is that MT algorithms and literalness generally prove unsuitable to capture and transfer the emotions, semantic nuances, and stylistic features of literature (Alowedi and Al-Ahdal 2023, 1527).

Against this backdrop, this paper offers a human quality assessment of automatic poetry translation on the Italian-English language pair. It investigates the usability of AI for poetry by comparing the outputs produced by three freely accessible MT translation systems (Google, DeepL, and Yandex). To this aim, it uses a selection of poems by the 19<sup>th</sup>-century Italian poet Giacomo Leopardi as a case study to determine which tool better performs with poetic texts, and which are the linguistic areas calling for improvement.

This paper does not suggest that AI should be used to translate poetry but rather shows that, by offering a wide spectrum of stylistic and linguistic devices, poetic language can be used to improve MT efficacy.

## 2. LITERATURE REVIEW

In the field of poetry, greater academic attention has been devoted to automatic poetry generation than translation (Chakrabarty et al. 2021).

It appears that studies on quality evaluation of poetry machine translation have been confined to a few language pairs, including, for example, Thai to English (Waijanya and Mingkhwan 2012), Portuguese to English (Humbly 2019), Arabic to English (Alowedi and Al-Ahdal 2013), and low-resourced linguistic combinations such as the German-Croatian one (Seljan et al. 2020), among others. Existing literature evidenced that MT especially errs with polysemy, adjectives, and context, and that its efficiency is also dependent on language combinations.

Furthermore, whilst automatic evaluation metrics such as BLEU may be effectively used to assess quality in non-literary machine translation, previous research demonstrated that these may be unsuitable to measure aspects such as style and figurative language, which are typical of poetry (Chakrabarty et al. 2021). This seemingly suggests that human evaluation of poetic machine translation is necessary to better understand how AI performs with poetic texts.

### 3. METHODOLOGY

This paper investigates the usability of MT for poetry using a corpus of five poems («L'infinito», «Il passero solitario», «A Silvia», «La quiete dopo la tempesta», «Il sabato del villaggio») by the poet Giacomo Leopardi. Leopardi's poems lend themselves well to the purposes of this paper for two main reasons: their peculiar stylistic features have often gained them the reputation of being untranslatable into another language (Stewart 2017, 44); the repeated use of specific lexemes. The latter aspect also enables us to explore how MT deals with lexical consistency in poetic texts, especially in the case of polysemous words.

The main criterion guiding text selection was the presence of items which are generally considered challenging in poetry translation. These include figurative language, disused vocabulary, syntactic inversions, and passages with ambiguous interpretations. Machine-generated translations were acquired using three freely accessible AI translation tools: Google and DeepL, which are both neural MT systems, and Yandex, combining neural and statistical methods. Once all translations were automatically generated, a double contrastive analysis was conducted: each machine output was first juxtaposed to the ST to determine accuracy; the translations produced by each AI system were then compared between each other to determine which tool better performs with poetic texts, and which are the linguistic areas calling for improvement. Results shall be divided into different categories (lexical, semantic, syntactic, grammatical, rhetorical, and stylistic). In a few instances, MT outputs shall be compared to human translation to check how AI translation systems rank with respect to professional translators.

### 4. RESULTS

#### 4.1. Lexical Level

As shown by the table below, lexical issues were particularly caused by polysemy, apocope, homographs, outdated diction, and orthographic similarity (see table 1).

SL-word	Meaning in the ST	Google	DeepL	Yandex	Type of mistranslation
<i>Mirare</i> (6)	To gaze	Aiming (1)	Mirrored (1)	Aim (1); target (1)	Polysemy
<i>German</i> (1)	Brother	German (1)	German (1)	German (1)	Apocope/orthographic similarity
<i>Mi fingo</i> (1)	I conceive in my mind (1)	I pretend (1)	I pretend (1)	I pretend (1)	Polysemy/outdated diction
<i>Onde</i> (4)	Whereof; whence	Waves (2); why (1)	-	Waves (3); how (1)	Homograph/outdated diction
<i>Odi</i> (5)	Hear	Hate (2)	-	Hate (5)	Homograph

<i>Romore/romorio</i> (2)	Noise	-	Roar (1); Romance (1)	Romance (1); Romorio (1)	Orthographic similarity
------------------------------	-------	---	-----------------------------	-----------------------------------	----------------------------

*Table 1. Examples of lexical mistranslations and number of occurrences across the corpus*

*Mirare* is an example of polysemous word as in Italian it can have the double meaning of «to gaze» and «to aim at» (Garzanti Linguistica 2024). Six occurrences were registered across the corpus. Although Leopardi always employs the word in the sense of «to gaze», Google and Yandex often mistranslated the term and replaced it with verbs such as «to aim» or «to target».

A feature common to all the outputs analysed was the presence of untranslated words. These generally resulted from machines' inability to recognise apocope and disused vocabulary. Apocope is often used in poetry for metrical reasons. An example is *german*, elided form of *germano* [brother]<sup>1</sup> in *german di giovinezza* [brother of youth], line 20 of «Il passero solitario». The term was probably left untranslated due to its similarity to the English word «German». As shown by the table, indeed, orthographic likeness was another common pitfall for MT. However, since «German» is also an adjective of nationality in English, its presence in the TT causes ambiguity.

As for outdated diction, an example is the verb *mi fingo* in *io nel pensier mi fingo*<sup>2</sup> [I conceive in my mind] (line 7, «L'infinito»). In contemporary Italian, the verb  *fingere* is generally used as a synonym of «to pretend». In Leopardi's figurative lexicon, however, the verb has the Latin meaning of «to imagine» or «to create in one's mind» (Treccani n.d.). Leopardi uses *mi fingo* in relation to the imagination of spatial infinity described in the text. This was mistranslated as «I pretend» by all three MT systems. This suggests that, apart from lacking contextual knowledge, AI tools do not take into account diachronic linguistic variations. This lexical mistranslation has a negative impact on authorial representation since it fails to convey the idea of imagination, this being a key concept in Leopardi's poetics.

#### 4.2. *Semantic Distortions*

In the corpus analysed, semantic distortions were generally caused by homographs. This is due, again, to machines' inability to detect the correct grammatical functions of words based on contextual information.

This was, for instance, the case of the adjective *muti* [silent] used in the expression *quando muti questi occhi all'altrui core* [when these eyes shall be silent to others' heart], line 53, «Il passero solitario». Apart from being an adjective, *muti* can also be the second person singular of the verb *mutare* [to change] in the simple present. The fact that it is used in connection to a noun, *occhi* [eyes], may have suggested machines that the term is employed with adjectival function. However, Google disregarded this aspect and mistranslated the sentence as «when I change these eyes». This could be also considered

<sup>1</sup> All translations in square brackets are my own.

<sup>2</sup> All passages from Leopardi's *Canti* are quoted from Blasucci (2019).

as an example of «hallucination», a TL output which is «fluent but unrelated to the source» (Müller et al. 2020, 151).

In the dataset analysed, it was Yandex, however, which generated the highest number of hallucinations, sometimes producing ludicrous results. For example, in «L'infinito», the passage *e mi sovvien l'eterno* [and I call to mind eternity] (line 11) is misleadingly changed into «LORD helps me». Apart from being erroneous, such a rendering is completely out of tune with the poet's agnostic disposition. Knowledge of the source poet is essential for accurate rendition. Nevertheless, this is completely out of the reach of MT, thus evidencing that human intervention is essential in poetic rendition.

#### 4.3. Interpretation

The issue of semantic distortions is connected to that of interpretation. Previous research on poetic machine translation showed that Google was able to correct wrong interpretations by human translators (Humblé 2019). A misleading passage in Leopardi's poetry is the expression *ove per poco il cor non si spaura* in «L'infinito» (lines 7-8). The passage should not be read literally as «where my heart is not afraid» but should be rather understood in the sense of «where my heart is almost afraid». This is because the adverb *non* does not introduce a negation, as it generally does in Italian, but is employed here in a pleonastic way (Perella 2000, 367). As remarked by Perella (2000, 367), human translators have often been misled by the grammatical structure of the passage, producing erroneous renderings. As shown by the below table, the same remark is valid for machine translation.

ST (lines 7-8, «L'infinito»)	TTs
<i>Ove per poco il cor non si spaura</i>	«And for a moment I am calm» (Townsend 1887)
	«The heart is not afraid» (Google)
	«The heart is not dismayed» (DeepL)

Table 2. Interpretative issues

The above translations convey a meaning which is opposite to that expressed by the ST, thus confuting the argument that MT can correct erroneous interpretations by human translators.

#### 4.4. Syntactic Level

In the STs, syntactic inversions were frequently used to foreground key themes and achieve particular rhythmic effects. Machines employed dissimilar approaches: whilst DeepL and Yandex generally maintained the inversions displayed in the STs, in several cases Google repristinated the standard word order of the English sentence. Let us compare, for example, the subject-verb inversion in line 31 of «Il sabato del villaggio»: *poi quando intorno è spenta ogni altra face*, which literally translates as «then when around is extinguished every other light» (see Table 3).

ST (line 31, «Il sabato del villaggio»)	TTs
<i>Poi quando intorno è spenta ogni altra face</i>	«Then when around every other face is extinguished» (Google)
	«Then when all around is extinguished every other face» (DeepL)
	«Then when around is off each other face» (Yandex)

Table 3. Syntax

All three MT systems were unable to recognise the correct meaning of *face* [light], leaving it untranslated in the TT. However, unlike Google, DeepL and Yandex maintained the ST's syntactic inversion. As is known, the English language is marked by lesser syntactic flexibility than Italian, making inversions less easily replicable in translation. Apart from resulting in increased fluency in the TL, Google's sentence restructuring seems to be an indication that the machine may be able to discern between the different syntactic functioning of the two languages in question.

#### 4.5. Grammatical Issues

Regarding grammar, most issues occurred at the verbal and pronominal level. In «A Silvia», for instance, problems were posed by the presence of verbal forms such as *apparìa* (line 30) [it appeared], *perìa* [it died] (line 49), and *splendea* [it shone] (line 3). These are imperfect indicative third person inflections and are nowadays obsolete in contemporary Italian. Machines were able to detect the correct meanings of such verbs, but mistakenly replaced them with different tenses (simple present and future simple): «it appears» (DeepL, Google), «it looks like» (Yandex), «it will perish» (Google), and «it shines» (Yandex). Apart from being grammatically inaccurate, the missed reproduction of imperfect verbal forms also has a negative impact on meaning. This is because in the original text the alternation between past and present tenses on the formal level serves to mark the antithesis between the time of memory and illusions, and that of truth, which are central concepts in Leopardi's production.

Previous studies demonstrated that Google was unable to capture cohesion (Voigt and Jurafski 2012). This piece of research confirms these results, adding that the same consideration is also valid for DeepL and Yandex. Apart from problems of cohesion and anaphoric reference, the misuse of pronouns seems symptomatic of the often-discussed gender bias in machine translation (Prates et al. 2020). In the corpus here analysed, whilst Google alternated between masculine and feminine pronouns in a seemingly random fashion, Yandex showed a neat masculine skew. For example, let us consider the opening of «Il sabato del Villaggio» depicting a *donzella* [young lady] coming from the countryside (see Table 4).

Il sabato del villaggio (lines 1-4)	TT generated by Yandex
<i>La donzella vien dalla campagna, in sul calar del sole, col suo fascio dell'erba; e reca in mano</i>	«The lady comes from the countryside On the setting of the sun,

<i>un mazzolin di rose e viole, [...]</i>	With <b>his</b> bundle of grass; and <b>he</b> carries in <b>his</b> hand A bouquet of roses and violets [...]
---	---

Table 4. Gender bias

Although the term *donzelletta* has been appropriately rendered as «lady», the predominance of masculine adjectives and pronouns («his» and «he») in the TT suggests that the machine completely overlooked the feminine gender of the character in question.

#### 4.6. Rhetorical Level

Despite its literality, Google was surprisingly able to infer the figurative meaning of some of the STs' passages. For instance, *il fior degli anni tuoi* [the flower of your years] (line 43, «A Silvia») is a metaphor where «flower» stands for youth. Instead of translating it literally, Google detected the metaphorical sense of the expression and appropriately rendered it as «the prime of your years». The term «prime» has also often been used by human translators of the *Canti* to refer to youth, for example by Joseph Tusiani in his 1998 anthology.

Another example is *Porgea gli orecchi al suon della tua voce* (line 20, «A Silvia»), which roughly translates as «I gave my ears to the sound of your voice». The expression *porgea gli orecchi* is figuratively used in the sense of «to listen to» (De Robertis, 1969, 202). Unexpectedly, Google inferred the implied sense of the word *orecchi*, meaning «ears» (which stands for the act of listening), and appropriately substituted it for «I listened to the sound of your voice».

#### 4.7. Stylistic Considerations

Some differences were also registered at the stylistic level. Unlike Google and Yandex, DeepL made a conspicuous use of archaic words including «thee» and «thou», for example. Their presence in the TTs signals that, unlike the other two automatic translation tools, DeepL is able to discern between different types of registers and reproduce them in the TL.

## 5. DISCUSSION AND CONCLUSION

Results evidenced that machine-generated translations did not meet acceptable publishing standards. It can be, therefore, concluded that AI is unsuited for Italian to English poetry translation without human intervention. The analysis highlighted a series of distinct features for each of the tool utilised. For instance, DeepL can produce better poetic translations than Google and Yandex, especially regarding style and word choice. Yandex generated the most elevated number of untranslated words and semantic distortions. Google stood in an intermediate position: it generally achieved increased fluency on the syntactic level and was able to understand figurative language more often than expected. On the negative side, the machine's inability to infer contextual

information caused a number of grammatical inaccuracies. In general, lexicon was the most problematic area for all MT systems since word choice was often arbitrary and overlooked context and grammatical relations. By illustrating some of the common weaknesses of automatic poetry translation, this paper showed that poetry can turn into a useful linguistic tool to detect and improve MT limitations. With respect to previous research, it highlighted that MT systematically errs with homographs, apocope, and outdated diction.

Previous studies had claimed that some automatic translation systems were able to correct wrong human interpretations. The results of this paper confute this conclusion, showing that semantic alterations and lexical mistranslations were a frequent issue in the dataset analysed. MT also blunders when interpretation is dependent on mere grammatical issues. Distortions resulted in either the introduction of concepts alien to Leopardi's philosophy or the omission of key ideas in his poetics, thus producing negative effects on authorial representation.

This piece of research presents a series of limitations. Firstly, the results here discussed may soon become obsolete due to MT systems being constantly updated and improved. Secondly, the texts produced by the same poet are generally marked by recurring features and idiosyncrasies. Therefore, working on a single-author corpus may prevent a fuller representation of MT problems in poetic texts. Lastly, this study focused on a single language pair. Considering that the effectiveness of MT is also dependent on linguistic combinations, future research on automatic poetry translation may have a multilanguage focus, simultaneously comparing not only different language pairs but also wider corpora of texts from different authors, periods, and literary traditions.

## REFERENCES

- Alowedi, Noha Abdullah, and Arif Ahmed Mohammed Hassan Al-Ahdal. 2023. «Artificial intelligence based Arabic-to-English machine versus human translation of poetry: an analytical study of outcomes». *Journal of Namibian Studies* 33 (S2): 1523-38.
- Chakrabarty, Tuhin, Arkadly Saakyan, and Smaranda Muresan. 2021. «Don't go far off: an Empirical Study on Neural Poetry Translation». *Proceedings of the 2021 Conference on Empirical methods in Natural Language Processing*, 7253-65.
- Garzanti Linguistica. 2024. «Mirare». Accessed April 15, 2024. <https://www.garzantilinguistica.it/ricerca/?q=mirare>.
- Humblé, Philippe. 2019. «Machine Translation and Poetry. The case of English and Portuguese». *Ilha do Desterro* 72 (2): 41-56.
- Leopardi, Giacomo, and Blasucci, Luigi (2019). *Canti* (Volumes I and II). Milan: Fondazione Pietro Bembo/ Ugo Guanda Editore.
- Leopardi, Giacomo, and Giuseppe De Robertis. 1969. *Canti*. Nuova ristampa. Firenze: Monnier.
- Müller, Mathias, Annette Rios, and Rico Sennrich. 2020. «Domain robustness in neural machine translation». *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas*, 151-64.



- Perella, Nicholas. 2000. «Translating Leopardi?». *Italica* 77 (3): 357-85.
- Prates, Marcelo, O.R., Pedro H. Avelar, and Lamb Luís C. 2020. «Assessing gender bias in machine translation: a case study with Google Translate». *Neural Computing and Applications*, 32: 6363–81. Accessed May 7, 2024. [https://link.springer.com/article/10.1007/s00521-019-04144-6?utm\\_source=getftr&utm\\_medium=getftr&utm\\_campaign=getftr\\_pilot](https://link.springer.com/article/10.1007/s00521-019-04144-6?utm_source=getftr&utm_medium=getftr&utm_campaign=getftr_pilot).
- Seljan, Sanja, Ivan Dunder, and Marko Pavloski. 2020. «Human Quality Evaluation of Machine-Translated Poetry». *43<sup>rd</sup> International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1040-45.
- Stewart, Susan. 2017. «Translating Leopardi's "L'Infinito": an Infinite Task». In *Into English: Poems, Translations, Commentaries*, edited by Pruffer Kevin and Collins Martha, 42-47. Minneapolis: Graywolf Press.
- Townsend, Frederick. 1887. *The Poems of Giacomo Leopardi*. New York: Putnam's son.
- Treccani (n.d.). «Fingere». Accessed April 15, 2024. <https://www.treccani.it/vocabolario/fingere/>.
- Tusiani, Joseph, Pietro Magno, and Franco Foschi. 1998. *Leopardi's Canti*. Fasano: Schena Editore.
- Voigt, Rob, and Dan Jurafsky. 2012. «Towards a Literary Machine Translation: The Role of Referential Cohesion». *Proceedings of the NAACL-HLT 2012 Workshop on Computational Linguistics for Literature*, 18-25.
- Waijanya, Sajjaporn, and A. Mingkhwan. 2012. «Thai Poetry in Machine Translation: An Analysis of Poetry Translation using Statistical Machine Translation». *AIJSTPME* 5 (3): 49-56.

