LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 27, No. 2, October 2024, pp. 1049-1061

LLT Journal: A Journal on Language and Language Teaching
http://e-journal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

POETIC TRANSLATION AND STUDENTS' APPRECIATION THROUGH HUMAN TRANSLATOR AND MACHINE SYSTEM

Danti Pudjiati[1], Maria Vincentia Eka Mulatsih[2*], Ilham Ilham[3], Deta Maria Sri Darta[4], and Febrio Rahim Suhel[5]
[1,5]STKIP Kusuma Negara, Indonesia
[2] Sanata Dharma University, Indonesia
[3]Muhammadiyah Palangkaraya University, Indonesia
[4]Satya Wacana Christian University, Indonesia
dantipudjiati@stkipkusumanegara.ac.id[1], mv_ika@usd.ac.id[2],
ilhamroy88@gmail.com[3], deta.darta@uksw.edu[4],
riosukiandi@stkipkusumanegara.ac.id[5]
*correspondence: mv_ika@usd.ac.id
https://doi.org/10.24071/llt.v27i2.9299
received 25 July 2024; accepted 28 October 2024

Abstract
The engagement with poetry is a personalized journey that transcends standardized methodologies. Key to this process is complete immersion in the poetic experience, alongside evaluators' openness to both human and machine-generated translations from Indonesian to English. The overarching goal is to enhance students' discernment and appreciation of these translated works. The study's specific objectives involved comparing poetic translation assessments by evaluators for both human translators and machine systems. It was to assess students' appreciation of poetry through the lens of both translation methods across three different institutions. The research employed a mixed-methods approach, combining quantitative and qualitative descriptive analyses, including descriptive statistics. The dataset comprised two Indonesian poems by Taufik Ismail, rated using a score as proposed by Nababan. Findings indicate that human translators outperformed machine systems in terms of accuracy, acceptance, and readability. While students from Institutions 1 and 2 preferred human-translated poetry, students from Institution 3 favoured machine-translated versions. This suggests that human translation quality remains superior.

Keywords: human translation, machine English, poetry translation

Introduction
Poetry appreciation is the scholarly process of assessing or valuing poetry. In a more comprehensive sense, poetry appreciation entails the systematic observation, evaluation, and admiration of the crucial components embedded in poetry (Faulkner, 2019; Finnegan, 2018; Stockwell, 2019). This encompasses the constituents of poetry such as meter, theme, moral, ambiance, poetic imagery including visual, auditory, gustatory, olfactory, tactile, and kinetic aspects, and the values conveyed in poetry comprising cultural, social, ethical, and moral values.

Indeed, the process of appreciating a poem is a complex effort that unfolds in several distinct stages (Gönen, 2018; Hidayat et al., 2022). The commencement of this journey involves reading the poem for personal enjoyment. This stage is pivotal as it establishes the foundation for the entire process. It is noteworthy that a single reading may not be sufficient to fully comprehend the essence of the poem. Multiple readings create many possibilities to arise excitement and provide a deeper understanding of sensing poetic expression that is depicted by a poem. Thus, each reading may reveal a new point of view for enriching kinds of experience.

The following steps should be underlined to examine the poems comprehensively. Such examination focuses on words or expressions that are notable and fascinating to explore (Mulatsih, 2020; Pudjiati & Zuriyati, 2022). These elements function as the centre of interpreting the insightful meaning and themes that are described by the poet. Thus, another important step is to identify the poem's structure, for example, free-verse, sonnet, haiku, limerick, and others in relation to the choice of the poet's stylistic which tends to cause a further deep meaning of a poem. The knowledge of the genre, rhyme scheme, figures of speech, linguistic style, speaker's tone, and references to another literary work should be familiarized (Houache & Zedek, 2023; Khan et al., 2023; Mahbub et al., 2023). This makes someone understand both the context and the literary traditions that are written in the poem. The final step comprises some understanding toward the speaker, the title, the connotation and denotation, and the objective of the poem. This step confirms that all steps are taken to comprehend the poem thoroughly, including the nuances and the complexities (Nuryadi, 2021; Rustandi, 2020).

Human translation is the process of transferring text from one language into another language that is conducted by humans. The human translation is often measured to be more precise and pragmatic than machine translation regarding the complicated language aspects, for example, grammar, syntax, and context (Heilmann, 2020; Qassem & Aldaheri, 2023). Furthermore, professional human translators produce their translation with the touch of culture and natural empathy which is difficult to imitate by a machine. Moreover, humans are certainly more sensible to communicate with the audience, especially in the domain of creative content or persuading. On the contrary human translation need cost much and more duration in completing the work.

Machine translation supports the procedure of transferring text from one language into another language using computer software. Statistical machine translation and neural translation are considered to be the most significant types of machine translation systems (Banik et al., 2019; Stahlberg, 2020). The machine translation software serves greater speed and efficiency as well as human translation which leads to preferred usage by plentiful businesses and organizations. On the contrary, such a machine may blunder when handling documents with specialized or nuanced content. Even though machine translation serves the fascination of turnaround and produces more lucrative, human translator promises high accuracy, linguistic appropriateness, and cultural thoughtfulness, particularly for sophisticated or industry-specific content (Macken et al., 2020; Melgar Hernández, 2022). Accordingly, the decision between a human translator and a machine system is basically based on the specific needs and requirements of the assignment at hand.

The previous research to probe poetic translation by human and machine systems was completed (Dunder et al., 2021; Ni & Wang, 2022; Seljan et al., 2020).

The results discovered that human translators showed greater flexibility and dynamism, while machine methods fell short in recognizing the errors of both grammar and rationality because of their strict programming model. Regardless of this issue, the researchers approved that the machine system can overcome a huge volume of translation assignments, however, it cannot replace the human translator completely.

Another research investigated the translation of ancient Arabic poetry stanzas by human and machine systems (Adiel et al., 2023). The result exposed that the machine method presented particular advantages but it could not succeed in the role of humans to translate Arabic poetry into English. In spite of its abilities, the machine system still requires considerable post-editing to achieve acceptable outcomes.

There are three findings dealing with poetic translation research (Seljan et al., 2020). First, poetic translations are extremely detailed and exposed to reach the individual interpretation. Secondly, the machine system integrates novel algorithms than traditional which had not been examined intensively for reflective languages, such as Croatian. Thirdly, to evaluate the machine system in terms of the small amount of dataset and the involvement of human evaluators.

In recent comparative literature research, the primary focus is on the analysis of semantic meaning within individual lines of poetry. This study leverages the Siamese MaLSTM algorithm to assess originality by discerning subtle semantic nuances and identifying similar wording in pairs of poetic lines. The model, trained on an Indonesian poetry collection, undergoes essential stages including the processing of data, the embedding word, training, and testing as well. Such a translation evaluation model explored the accuracy of a praiseworthy level to identify sentence pairs in terms of semantic similarity. Furthermore, the study introduces a system application designed for the management of Indonesian poetry, incorporating originality checks facilitated by the machine learning model developed by researchers (Anggriawan et al., 2023).

Previous research has primarily focused on poetry translation using both human and machine methods. However, the aspect of poetry appreciation resulting from these distinct translation approaches remains unexplored. Consequently, to address this gap, a research study is undertaken to assess poetic translation and explore its appreciation across different translation methods by students of higher education.

Method

The design of this research was a mixed-methods to scrutinize poetic translations. In the beginning, it used a quantitative approach to evaluate the quality of poetic translations. Then, this research also managed a comparative study through descriptive statistics for comparing students' appreciation of poetic translations conducted by human translators and machine systems. The participants of his study came from three different campuses in Java. Such analysis offers a depiction comprehensively of the phenomenon which is observed to expose essential patterns (Glesne, 2016; Hennink et al., 2020).

Two principal instruments in this research were a set of translation assessment questions suggested by Nababan (Nababan et al., 2012) in terms of score and a survey with online mode. The assessment and the survey were held through an

online platform. The objective of the assessment was to investigate the translation quality by human translators and machine systems. The survey aimed at gathering the relevant data. Implementing various methods, namely the assessments and the survey questionnaires for triangulation in collecting data led the researchers to improve the credibility of their findings through gathering perceptions from a larger sample (Kim et al., 2017; Rahman, 2020).

The participants in this research were 129 students around 20 years old and studying in an English education program for an undergraduate degree. An information-oriented sampling method was employed as a guide to select the participants. Explicitly, the participants were those who already completed the subject of translation and they agreed to be involved in this research voluntarily and anonymously. This study was managed in a natural site of three private higher education institutions in three provinces.

The participants in this research were 129 students around 20 years old and studied in an English education program for undergraduate degrees. An information-oriented sampling method was employed as a guide to select the participants. Explicitly, the participants were those who already completed the subject of translation and they agreed to be involved in this research voluntarily. This study was managed in a natural site of three private higher education institutions in three provinces. These institutions include Sanata Dharma University in Yogyakarta, Satya Wacana University in Salatiga, Central Java, and Kusuma Negara Teachers College in East Jakarta. To provide further context, 88 students were from Yogyakarta, 10 from Salatiga, and 31 from East Jakarta. Given these contextual details, these research sites were deemed suitable for addressing the investigative objectives and collecting relevant data.

In the examination of the research data, the investigators utilized a hybrid methodology that incorporated both quantitative and qualitative approaches. Given the comparative nature of the study, the data was bifurcated into two distinct categories: (1) an evaluation of translation quality between human translator and machine system, and (2) an assessment of the appreciation of poetic translation using human and machine methods across three different campuses. In the first phase, three expert raters evaluated poetic translations produced by both human translators and machine systems. Two poems entitled “Pantun Terang Bulan di Midwest” and “Adakah Suara Cemara” written by Taufiq Ismail consisted of 34 lines to become the data. These evaluators possess doctoral degrees in the English language, exhibit a keen interest in translation, boast over decades of English teaching experience, and notably, hold senior lecturer positions. Three anonymous evaluators check the data analysis. The assessment protocol was based on the translation quality assessment (TQA) model initiated by Nababan et al. This model comprises three critical dimensions: accuracy, acceptability, and readability, each of which is further delineated by three specific indicators in terms of score as presented by the table as follows.

The qualitative data collection process involved administering an online survey in the form of closed-ended questionnaires. Participants were specifically instructed to complete these surveys, which aimed to gather insights regarding their appreciation of poetic translation—both from human translator and machine systems. The questionnaire comprised a series of structured questions, and participants responded by selecting predefined options. Through this method, we

tried to find nuanced viewpoints on the quality and reception of translated poems. Therefore, it was to enrich the understanding of the phenomenon entirely.

In total, 34 samples were analyzed qualitatively. Then, the participants were assigned to evaluate and select the translation based on their appreciation toward human translators and machine systems. Consequently, the quantitative assessment was measured to quantify the degree of preference conveyed by those participants. Such assessment was stated in the percentage to signify the comparative favourability of the translations from two modalities. The coincidence of human translator and machine system permitted a nuanced examination of both aesthetic and linguistic nuances innate in a poem, shedding light on the relationship between human creativity and computational algorithms in the domain of literary translation.

Findings and Discussion

The translation assessment

Each evaluator assessed the English translations of Indonesian poetry, which constituted a dataset of 34 samples. The evaluation focused on three key aspects: accuracy, acceptance, and readability. Ratings were assigned on a score from 1 to 3, representing the overall quality. This kind of poetry translation evaluation or called as translation quality assessment is advocated to be used for an accountability result (Ma & Wang, 2020; Pallavi & Mojibur, 2018). After evaluating the translation result in detail for three evaluators, namely E1, E2, and E3, then it was presented in the mean score of evaluation results to be considered as follows.

Table 1. The mean score of the translation evaluation

Total Data	Evaluator	Human Translator			Machine System		
		Accuracy	Acceptance	Readability	Accuracy	Acceptance	Readability
34	E1 E2 E3	2.27	2.25	2.49	1.89	1.56	1.76

In Table 1, descriptive statistics illustrated a dataset based on scores provided by three evaluators. Specifically, the mean scores for accuracy, acceptance, and readability were higher for human translators compared to machine systems. Such results proved that the poetic translation quality of a human translator was better than the machine system (Seljan et al., 2020). Furthermore, this created a suggestion that human translators outperformed machine systems in terms of accuracy, acceptance, and readability. It was stated that the mean scores for human translators were 2.27, 2.25, and 2.49, while those for machine systems were 1.89, 1.56, and 1.76.

Table 2. Descriptive statistic

System	Criterion	N	Std.Dev.
Human	Accuracy	34	0.71
	Acceptance	34	0.59
	Readability	34	0.71
Machine	Accuracy	34	0.86
	Acceptance	34	0.59
	Readability	34	0.90

In Table 2, the standard deviation values for both the accuracy and readability of human translators were 0.71. Apart from this, the machine system showed higher standard deviations, namely 0.86 for accuracy and 0.90 for readability. Remarkably, it was found that the acceptance of standard deviation between human translator and machine system was in the same amount, namely 0.59. This implied that the smaller the standard deviation, the more similar the values in the items (Casquilho & Buescu, 2022). In other words, it can be inferred that the more accurate they were with the mean.

Table 3. Score per evaluator

Evaluator	1			2			3		
Criterion	A	Ac	R	A	Ac	R	A	Ac	R
Human Translation	1.91	1.92	1.68	2.62	2.71	2.79	2.59	2.12	3.00
Machine Translation	1.59	1.38	1.32	1.26	1.29	1.24	2.82	2.00	2.74
Human Translation	A (Average for all evaluators): 2.37 Ac (Average for all evaluators):2.25 R (Average for all evaluators): 2.49								
Machine Translation	A (Average for all evaluators):1.89 Ac (Average for all evaluators): 1.56 R (Average for all evaluators): 1.76								
Note: A = Accuracy, Ac = Acceptance, R = Readability									

Table 3 presented the mean scores assigned by three evaluators respectively. These evaluators compared human translators and machine systems of original source text. Evaluator one gave scores of the poetic translation by a human translator below 2 in three aspects: accuracy, acceptance, and readability. In contrast, both evaluators two and three rated the poetic translation by a human translator above a score of 2 in those aspects.

However, evaluators one and two scored poetic translation by machine system below 2 of those aspects, namely accuracy, acceptance, and readability Unlike evaluators one and two, evaluator three scored poetic translation by machine system above 2 in those aspects. Overall, the mean scores across all three evaluators favoured human translators over machine systems, indicating that human-translated poems were superior. This statement was in line with the research on semantic distinction in literary translation (Sedlanić, 2022).

Cronbach's alpha coefficient is used to assess the internal constancy or reliability of the survey item set (Amirrudin et al., 2020). It measures the degree of agreement on a consistent scale ranging from 0 to 1. Specifically, a score of 0.00 indicates a complete absence of agreement, while a score of 1.00 represents perfect agreement. Researchers generally believe a Cronbach's alpha value of 0.70 or higher is reliable. Such computation assists in determining the consistency of measurement of similar characteristics toward the collection of data. Below is the detailed measurement of Cronbach's alpha values.

$\alpha < 0.5$ = No agreement/considered unacceptable
$0.5 \leq \alpha < 0.6$ = Poor agreement
$0.6 \leq \alpha < 0.7$ = Acceptable agreement
$0.7 \leq \alpha < 0.9$ = Good agreement
$\alpha \geq 0.9$ = Excellent agreement

Table 4. Cronbach's Alpha score

System/Criterion	Accuracy	Acceptance	Readability
Human Translation	0.92	0.95	0.99
Machine Translation	0.99	0.95	0.98

The specific alpha values mentioned were likely reported in Table 4 of the research. This table provided details about the statistical tests performed and their associated significance levels. The research findings indicated that there was an excellent agreement between human quality evaluation and machine quality evaluation. Such agreement supported the process of evaluation to be more truthful (Vaske et al., 2017). Therefore, this suggested that both methods in the research yield similar results.

The translation appreciation

Students are assigned to state their view in poetic translation toward the two methods of translators. The main consideration is translating is not only transferring one language to another but involves cultural knowledge to obtain a good understanding of the source text. Moreover, the translator can translate such poetic translation to readers who have different beliefs, behaviors, values, morals, and rules.

In Figure I, students across different institutions appreciated poetry translation between human translator and machine system, expressing their preferences as percentages. Specifically, Institution 1 achieved 65.17%, Institution 2 gained 60%, and Institution 3 received 49.23% in appreciating poetic translation through a human translator. In contrast, Institution 1 obtained 34.83%, Institution 2 reached 40% and Institution 3 attained 50.77% in appreciating poetic translation by machine system. These findings indicated that students from Institutions 1 and 2 preferred human translators for poetic translation as supported by the study of comparing the poetic translation between human and machine (Dawah, 2024). Surprisingly, students from Institution 3 favoured the machine system for poetic translation.

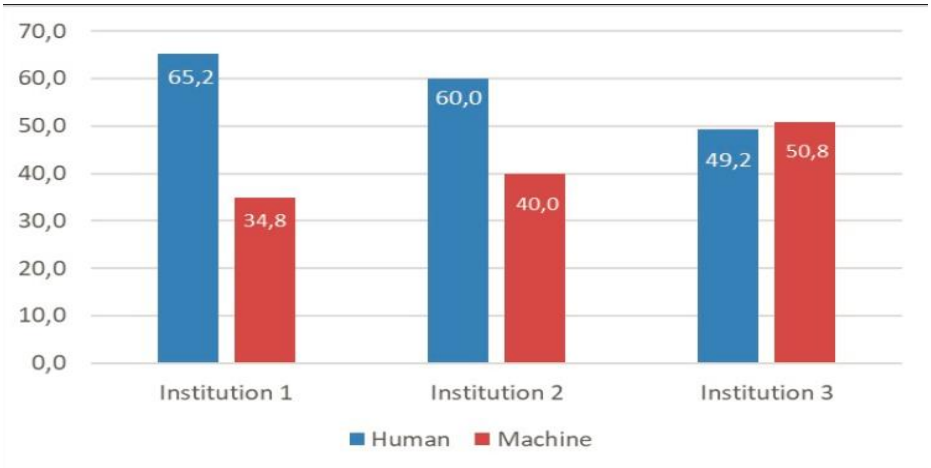

Figure 1. Students' evaluation

In Table 5 according to student feedback from three campuses, human translators were more appreciated for poetic translation across 34 data points. However, there were exceptions. Institutions 1 and 2 favoured human translators for poetic translation for almost all of the data, except the poem title (data point 1). However, Institution 3 preferred a machine system for poetic translation. Interestingly, both Institutions 1 and 3 preferred machine systems for poetic translation specifically for the poem title (data point 22). Institution 2 expressed a balanced preference for either a human translator or a machine system when it came to the title of poem translation.

Table 5. Students appreciation of poetry translation in percentage

No of	Institution 1		Institution 2		Institution 3	
Data	Human	Machine	Human	Machine	Human	Machine
1	67	33	70	30	45.2	54.8
2	72.7	27.3	70	30	61.3	38.7
3	72.7	27.3	70	30	61.3	38.7
4	72.7	27.3	70	30	61.3	38.7
5	72.7	27.3	70	30	61.3	38.7
6	78.4	21.6	50	50	61.3	38.7
7	78.4	21.6	50	50	61.3	38.7
8	78.4	21.6	50	50	61.3	38.7
9	78.4	21.6	50	50	61.3	38.7
10	56.8	43.2	60	40	54.8	45.2
11	56.8	43.2	60	40	54.8	45.2
12	56.8	43.2	60	40	54.8	45.2
13	56.8	43.2	60	40	54.8	45.2
14	67	33	80	20	41.9	58.1
15	67	33	80	20	41.9	58.1
16	67	33	80	20	41.9	58.1
17	67	33	80	20	41.9	58.1
18	71.6	28.4	50	50	25.8	74.2
19	71.6	28.4	50	50	25.8	74.2
20	71.6	28.4	50	50	25.8	74.2
21	71.6	28.4	50	50	25.8	74.2
22	44.3	55.7	50	50	41.9	58.1
23	56.8	43.2	60	40	51.6	48.4
24	56.8	43.2	60	40	51.6	48.4
25	56.8	43.2	60	40	51.6	48.4
26	56.8	43.2	60	40	51.6	48.4
27	61.4	38.6	50	50	51.6	48.4
28	61.4	38.6	50	50	51.6	48.4
29	61.4	38.6	50	50	51.6	48.4
30	61.4	38.6	50	50	51.6	48.4
31	61.4	38.6	60	40	48.4	51.6
32	61.4	38.6	60	40	48.4	51.6
33	61.4	38.6	60	40	48.4	51.6
34	61.4	38.6	60	40	48.4	51.6
Average	65.17	34.83	60	40	49.23	50.77

Generally, it can be inferred that Institution 1 holds a greater appreciation for human translators, namely 33 data out of 34. In contrast, Institution 2 has 21 data out of 34 to poetic translation by human translator and this institution maintained a balanced appreciation for both human translator and machine system, especially across data points 6, 7, 8, 9, 18, 19, 20, 21, 22, 27, 28, 29, and 30. Conversely,

Institution 3 has 20 data out of 34 to poetic translation by human translator and it has 14 data points 14, 15, 16, 17, 18, 19, 20, 21, 30, 31, 32, 33, and 34 to poetic translation by machine system.

Conclusion

The empirical evidence has suggested that human-produced poetic translations outperform machine-generated ones. This was proved by evaluating three critical dimensions of poetic translation, namely accuracy, acceptance, and readability. In a comparative evaluation, there were three independent evaluators from three different institutions awarded consistently higher scores to human translators.

Such evaluation indicated that people can depend on humans to translate poetic expressions. This led to the machine translation result of poetic expressions still being examined by humans to confirm satisfaction of quality requirements, for example in the point of the clarity of the message being described and the style of language being conveyed. Therefore, it can be considered that post-editing translation is needed in translating poetic documents by machine to generate a qualified translation to reach similar quality compared to human-produced poetic translations.

This research revealed the students' assessment of appreciation toward poetic expression using human and machine methods of three campuses with distinguished results. Based on their assessment students from Institution 1 and Institution 2 stated a preference for human-translated poems nonetheless students from Institution 3 favoured machine translations for poetic translation. This meant that further research pertinent to human versus machine poetic translations could be investigated from different insights with more respondents involved who come from different campuses.

References

Adiel, M. A. E., Elsadig, M. A., Altigani, A., Mohamed, Y. A., Ahmed, B. E. S., & Elhassan, S. M. O. (2023). Accuracy and problems of machine-based translation in contrast to human-based translation when rendering health awareness texts versus poetry texts. *Academic Journal of Interdisciplinary Studies*, *12*(4), 223–231. https://doi.org/10.36941/ajis-2023-0108

Amirrudin, M., Nasution, K., & Supahar, S. (2020). Effect of variability on Cronbach alpha reliability in research practice. *Jurnal Matematika, Statistika dan Komputasi*, *17*(2), 223–230. https://doi.org/10.20956/jmsk.v17i2.11655

Anggriawan, E., Farid, F., & Sari, R. F. (2023). Sentences similarity detection in Indonesian poetry comparison using Siamese malstm. *ICIC Express Letters*, *17*(4), 389–396. https://doi.org/10.24507/icicel.17.04.389

Banik, D., Ekbal, A., Bhattacharyya, P., Bhattacharyya, S., & Platos, J. (2019). Statistical-based system combination approach to gain advantages over different machine translation systems. *Heliyon*, *5*(9), 1-9. https://doi.org/10.1016/j.heliyon.2019.e02504

Casquilho, M., & Buescu, J. (2022). Standard deviation estimation from sums of unequal-size samples. *Monte Carlo Methods and Applications*, *28*(3), 235–253. https://doi.org/10.1515/mcma-2022-2118

Dawah, W. A. A.-S. (2024). The human translation and the electronic translation, an applied comparative study of selected Hebrew poetic texts. *Journal of Language Studies*, *8*(5), 58–83. https://doi.org/10.25130/Lang.8.5.4

Dunder, I., Seljan, S., & Pavlovski, M. (2021). What makes machine-translated poetry look bad? A human error classification analysis. *Central European Conference on Information and Intelligent Systems*, 183–191.

Faulkner, S. L. (2019). *Poetic inquiry: Craft, method, and practice*. Oxfordshire: Routledge.

Finnegan, R. (2018). *Oral poetry: Its nature, significance, and social context*. Oregon: Wipf and Stock Publishers.

Glesne, C. (2016). *Becoming qualitative researchers: An introduction*. Washington: Eric.

Gönen, S. İ. K. (2018). Implementing poetry in the language class: A poetry-teaching framework for prospective English language teachers. *Advances in Language and Literary Studies*, *9*(5), 28–42. https://doi.org/10.7575/aiac.alls.v.9n.5p.28

Heilmann, A. (2020). *Profiling effects of syntactic complexity in translation: A multi-method approach* (Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen).

Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative research methods*. Melbourne: Sage.

Hidayat, A., Muliastuti, L., & Dewanti, R. (2022). Condition of poetry appreciation teaching materials at Kuningan University. *UNISET 2021: Proceedings of the 2nd Universitas Kuningan International Conference on System, Engineering, and Technology*,1-9. http://dx.doi.org/10.4108/eai.2-12-2021.2320228

Houache, A., & Zedek, H. (2023). *A stylistics study of the linguistic deviations in Emmett Williams' selected poems: She loves me, do you remember'and' the moon is green'* (Dissertation Université IBN Khaldoun-Tiaret).

Khan, H., Ali, W., & Naeem, R. (2023). Stylistic analysis of the poem "Ode to a nightingale" by John Keats. *Journal of Namibian Studies: History Politics Culture*, *33*(3), 5535–5560. https://doi.org/10.59670/jns.v33i.4967

Kim, H., Sefcik, J. S., & Bradway, C. (2017). Characteristics of qualitative descriptive studies: A systematic review. *Research in Nursing & Health*, *40*(1), 23–42. https://doi.org/10.1002/nur.21768

Ma, Y., & Wang, B. (2020). Description and quality assessment of poetry translation: Application of a linguistic model. *Contrastive Pragmatics*, *3*(1), 89–111.

Macken, L., Prou, D., & Tezcan, A. (2020). Quantifying the effect of machine translation in a high-quality human translation production process. *Informatics*, *7*(2), 12. https://doi.org/10.3390/informatics7020012

Mahbub, R., Khan, I. T., Anuva, S. S., Shahriar, M. S., Laskar, M. T. R., & Ahmed, S. (2023). Unveiling the essence of poetry: Introducing a comprehensive dataset and benchmark for poem summarization. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 14878–14886. https://doi.org/10.18653/v1/2023.emnlp-main.920

Melgar Hernández, R. (2022). *Neural machine translation as a translation tool: The case study of Spanish at the translation service of the Council of the European Union* (Master thesis, National University of Distance Education).

Mulatsih, M. V. E. (2020). *Introduction to prose in English language teaching*. Yogyakarta: Sanata Dharma University Press.

Nababan, M., Nuraeni, A., Sumardiono. (2012). Pengembangan model penilaian kualitas terjemahan. *Kajian Linguistik dan Sastra, 24*(1), 39-57.

Ni, J., & Wang, C. (2022). Research on translation of hyperbole based on effects studies--Taking selected poems of Li Bai in library of Chinese classics as an example. *International Journal of Social Science and Education Research*, *5*(5), 411–416. https://doi.org/10.6918/IJOSSER.202205_5(5).0058

Nuryadi, N. (2021). Denotation and connotation in London's William Blake. *Makna: Jurnal Kajian Komunikasi, Bahasa, dan Budaya*, *9*(2), 48–57. https://doi.org/10.33558/makna.v9i2.2976

Pallavi, K., & Mojibur, R. (2018). A preliminary pragmatic model to evaluate poetry translation. *Babel*, *64*(3), 434–463. https://doi.org/10.1075/babel.00046.pal

Pudjiati, D., & Zuriyati. (2022). Students' perception of cultural values in "travel" poem through youtube. *IJLECR - International Journal of Language Education and Culture Review*, *8*(1), 41–50. https://doi.org/10.21009/IJLECR.081.06

Qassem, M., & Aldaheri, M. M. (2023). Can machine translate dialogue acts: Evidence from translating dialogues from English to Arabic. *3L: Language, Linguistics, Literature*, *29*(4), 63-81. https://doi.org/10.17576/3L-2023-2904-05

Rahman, M. S. (2020). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review. *Journal of Education and Learning 6*(1), 102-102. https://doi.org/10/5539/jel.v6n1p102

Rustandi, P. (2020). Connotative and denotative meaning in poem "who am I, without exile?" by Mahmoud Darwish. *TEXTURA*, *1*(2), 30–36.

Sedlanić, R. (2022). *Machine translation vs. human translation: Semantic distinctions in English-Croatian translations* (Thesis, University of Rijeka, Rijeka, Croatia).

Seljan, S., Dunđer, I., & Pavlovski, M. (2020). Human quality evaluation of machine-translated poetry. *43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1040–1045. https://doi.org/10.23919/MIPRO48935.2020.9245436

Stahlberg, F. (2020). Neural machine translation: A review. *Journal of Artificial Intelligence Research*, *69*, 343–418. https://doi.org/10.1613/jair.1.12007

Stockwell, P. (2019). *Cognitive poetics: An introduction*. Oxfordshire: Routledge.

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences*, *39*(2), 163–173. https://doi.org/10.1080/01490400.2015.1127189

Appendices

Appendix 1. Evaluator instrument of translation accuracy

Translation Category	Score	Qualitative Parameter
Accuracy	3	The accurate transposition of lexical meanings, technical terms, phrases, sentences, or entire textual passages from the source language to the target language is of paramount importance. It is imperative that this process remains devoid of any semantic distortion or misrepresentation.
Less Accuracy	2	The majority of lexical meanings, technical terms, phrases, clauses, sentences, or texts from the source language have been accurately transferred into the target language. However, there still exist distortion in meaning or instances of multiple meanings (polysemy) in translation, or meanings that have been omitted, which compromise the integrity of the message.
No Accuracy	1	The meaning of words, technical terms, phrases, clauses, sentences, or text from the source language is inaccurately transferred to the target language or omitted.

Appendix 2. Evaluator instrument of translation acceptance

Translation Category	Score	Qualitative Parameter
Acceptance	3	The translation feels natural; technical terms commonly used and familiar to readers; phrases, clauses, and sentences used adhere to the norms of the Indonesian language.
Less Acceptance	2	Generally, the translation feels natural; however, there are minor issues related to usage or occasional grammatical errors.
Not Acceptance	1	The translation feels unnatural or resembles a translated work; technical terms used are uncommon and unfamiliar to readers; phrases, clauses, and sentences employed do not conform to the norms of the Indonesian language.

Appendix 3. Evaluator instrument of translation readability

Translation Category	Score	Qualitative Parameter
Readability	3	The translation feels natural; technical terms commonly used and familiar to readers; phrases, clauses, and sentences used adhere to the norms of the Indonesian language.
Less Readability	2	Generally, the translation is comprehensible to readers; however, there are certain sections that require multiple readings for full understanding.
Not Readability	1	The translation is difficult for readers to understand.

Appendix 4. Research data

Source Text	Data	Human Translator	Machine System
Pantun terang bulan di midwest	1	Full moon in midwest	Moon light poetry in the midwest
Sebuah bulan purnama	2	A perfect moon	A full moon
Bersinar agak merah	3	Sheds of rosy light	Glowing slightly red
Lingkarannya di sana	4	An aureole around it	The circle is there
Awan menggaris bawah	5	An underline off clouds	Clouds underline
Sungai Mississippi	6	The Mississippi	Mississippi River
Lebar dan keruh	7	Wide and muddy	Wide and murky
Bunyi-bunyi sepi	8	Sounds of silence	Quiet sound

Source Text	Data	Human Translator	Machine System
Awan Gemuruh	9	Roll of thunder	Thunderous clouds
Ladang-ladang jagung	10	Fields of corn	Corn fields
Rawa-rawa dukana	11	Sensual swamps	Ducana swamps
Serangga mendengung	12	Insect buzzing	Insects buzz
Sampaikan suara	13	Do you hear the sound?	Convey sound
Cuaca musim gugur	14	In the autumn air	Autumn weather
Bukit membisu	15	Hills lie mute	Silent Hill
Asap yang hancur	16	Smoke transform	Shattered smoke
Biru abu-abu	17	To greyish blue	Gray blue
Danau yang di sana	18	On the lake beyond	That lake over there
Seribu burung belibis	19	A thousand duck	A thousand grouse
Lereng pohon pina	20	The hillside pines	Pine tree slopes
Angina pun gerimis	21	Even the wind is wet	The wind was drizzling
Adakah suara cemara	22	Is it the sound of the pines	Is there a sound of fir
Adakah suara cemara	23	Is it the sound of the pines	Is there a spruce sound
Mendesing padamu	24	That hiss and roar at you	Whooshing at you
Adakah melintas sepintas	25	Is it the fleeting passage	Did anyone pass by?
Gemersik dedaunan lepas	26	Of rusting leaves	The rustle of loose leaves
Deretan bukit-bukit biru	27	A line of blue hills	Rows of blue hills
Menyeru lagu itu	28	Kneel the song	Shouting that song
Gugusan mega	29	A cluster of cloud	Mega group,
Lalah hiasan kencana	30	Is the bracelet's jewel?	Go for gold decoration
Adakah suara cemara	31	Is it the sound of the pines	Is there a spruce sound
Mendesing padamu	32	That hiss and roar at you	Whooshing at you
Adakah lautan ladang jagung	33	Is it the sea of corn fields	Is there a sea of corn fields
Mengombakkan suara itu	34	Throwing up waves of sound	Waves the sound