Humanities and Educational Sciences Journal
ISSN: 2617-5908 (print)
مجلـــة العلـــوم التربـــوية والدراســـات الإنســـانيـــة
ISSN: 2709-0302 (online)

The Influence of AI-Based Translation Tools on the Translation of Dr. Ghazi Al-Qusaibi's Poetry by Saudi EFL Learners(*)

Mustafa Abdulmuhsen Thabet
Assistant Professor of English Literature
Faculty of Language Studies
Arab Open University, Riyadh, Saudi Arabia
m.thabet@arabou.edu.sa
ORCID: 0009-0003-1038-1564

Adel Mohammed Qadha
Associate Professor of Applied Linguistics,
Department of English, Faculty of Education,
Zabid, Hodeidah University, Hodeidah, Yemen
aadel75q@gmail.com
ORCID: 0000-0002-0967-5448

*) تاريخ تسليم البحث 2/3/2024 تاريخ قبوله للنشر 18/4/2024
*) موقع المجلة: http://hesj.org/ojs/index.php/hesj/index

تأثير أدوات الترجمة القائمة على الذكاء الاصطناعي على ترجمة شعر الدكتور غازي القصيبي بواسطة متعلمي اللغة الإنجليزية السعوديين كلغة أجنبية

مصطفى عبد المحسن ثابت
أستاذ الأدب الإنجليزي المساعد، كلية الدراسات اللغوية
الجامعة العربية المفتوحة، الرياض، السعودية

عادل محمد قدحه
أستاذ اللغويات التطبيقية المشارك، قسم اللغة الإنجليزية،
كلية التربية - زبيد، جامعة الحديدة، الحديدة، اليمن

الملخص

هذه الدراسة تتناول تأثير أدوات الترجمة القائمة على استخدام الذكاء الاصطناعي على قدرة متعلمي اللغة الانجليزية كلغة أجنبية على ترجمة قصيدة من تأليف الشاعر السعودي المعروف د. غازي القصيبي من اللغة العربية إلى اللغة الانجليزية. شارك في البحث خمسون طالب سعودي جامعي يدرسون اللغة الإنجليزية كلغة أجنبية، تم تقسيمهم بشكل متساوي إلى مجموعة تجريبية استخدمت أدوات الذكاء الاصطناعي ومجموعة ضابطة لم تستخدم هذه الأدوات. طُلب من المجموعتين ترجمة إحدى قصائد الدكتور غازي القصيبي، ثم تمت مراجعة الترجمات وفقاً لمعايير مثل الدقة والسلاسة واستخدام أدوات شعرية والالتزام بأنماط النص. أظهرت نتائج التحليل الإحصائي أن ترجمات المجموعة التجريبية حصلت على درجات أعلى بشكل كبير في معظم المعايير مقارنة بالمجموعة الضابطة. في حين أظهر كلا المجموعتين القدرة على نقل المعنى الأساسي، حققت المجموعة المساعدة بالذكاء الاصطناعي مزايا في مجالات مثل إنتاج ترجمات أكثر سلاسة وتماسكاً مع وجود أخطاء نحوية أقل. وفي حدود التركيز على شاعر واحد، تقدم الدراسة نظرة عن كيفية استخدام أدوات الذكاء الاصطناعي بشكل استراتيجي لمساعدة المتعلمين في تخطي التحديات المتأصلة في ترجمة الشعر الذي يعتمد بشكل كبير على السياق الثقافي والتأويل الإبداعي للغة الرمزية. يمكن أن تكون الأبحاث المستقبلية التي تزيد حجم العينة وتنوع مصادر النصوص أكثر غزارة في النتائج.

الكلمات المفتاحية: شعر الدكتور غازي القصيبي، ترجمة الذكاء الاصطناعي، متعلمي اللغة الانجليزية السعوديين كلغة أجنبية، الدقة، عناصر الشعر، الترجمة البشرية.

The Influence of AI-Based Translation Tools on the Translation of Dr. Ghazi Al-Qusaibi's Poetry by Saudi EFL Learners

Mustafa Abdulmuhsen Thabet
Assistant Professor of English Literature
Faculty of Language Studies
Arab Open University, Riyadh, Saudi Arabia

Adel Mohammed Qadha
Associate Professor of Applied Linguistics,
Department of English, Faculty of Education,
Zabid, Hodeidah University, Hodeidah, Yemen

Abstract

This study examines the impact of artificial intelligence-based translation tools on English as a Foreign Language (EFL) learners' ability to translate a poem written by renowned Saudi Arabian poet Dr. Ghazi Al-Qusaibi from Arabic to English. Fifty Saudi undergraduate EFL students participated in the research and were divided equally into an experimental group that used AI tools and a control group that did not. Both groups were asked to translate one of Dr. Al-Qusaibi's poems, which was then evaluated based on parameters like accuracy, fluency, use of poetic devices, and adherence to genre conventions. Results of statistical analyses found the translations produced by students in the experimental group scored significantly higher on most parameters compared to the control group. While both groups demonstrated accuracy in conveying the core meaning, the AI-assisted group showed advantages in areas like producing more fluent and coherent translations with fewer syntactic errors. Within the limitations of focusing on just one poet, the study provides insights into how strategic integration of AI during the translation process could help learners overcome challenges inherent to translating poetry, which relies so heavily on cultural contextualization and creative interpretation of symbolic language. Further research expanding the sample size and variety of source texts could yield a richer understanding.

Keywords: Dr. Ghazi Al-Qusaibi's poetry, AI Translation, Saudi EFL learners, Accuracy, Poetic elements, Human translation.

Introduction

Poetry has long been hailed as a powerful form of expression, encapsulating the rich cultural heritage, emotions, and nuances of a language. The intricate nature of poetry, particularly Arabic poetry, presents a unique challenge for translators seeking to convey the same depth, beauty, and essence of the original work in a different language. However, with the rise of artificial intelligence (AI), the possibilities for bridging the gap between languages and cultures have expanded exponentially. In this paper, we will delve into the realm of using AI technology to translate Arabic poetry, exploring its potential, limitations, and the implications it holds for preserving and appreciating the art form.

Arabic poetry has a long and illustrious history, spanning centuries and encompassing a diverse range of themes, forms, and styles. Renowned Arab poets such as Al-Mutanabbi, Ibn Arabi, and Mahmoud Darwish have left an indelible mark on the literary world. Yet, the power and subtleties of Arabic poetry are often lost in translation. The inherent complexities of the Arabic language, including its unique grammar, metaphors, and cultural references, make it a formidable challenge for human translators. This is where the potential of AI in translating Arabic poetry comes to the fore.

Artificial Intelligence, particularly machine learning and natural language processing, has made remarkable strides in recent years. AI-powered translation tools such as Google Translate, DeepL, and Microsoft Translator have become increasingly sophisticated, and capable of understanding and translating complex linguistic structures. The application of AI to translate Arabic poetry holds immense promise in capturing the essence, rhythm, and cultural context of the original work, enabling a wider audience to appreciate the artistry of poets who wrote in Arabic (Dessi et al., 2021).

Despite its potential, AI translation of Arabic poetry faces several challenges. One of the primary hurdles is the intricate nature of poetic language, which often involves wordplay, metaphor, and cultural allusions. Capturing these nuances accurately requires a deep understanding of the cultural and historical context, something that AI models may struggle with (Van der Meer, 2019). The inherent subjectivity and ambiguity of poetry also pose a challenge for AI translation systems, as they often prioritize literal

translations over capturing the intended emotional impact. Moreover, the lack of availability of large-scale, high-quality training data for Arabic poetry further hampers the development of AI models specific to this domain.

While AI can assist in the translation process, it is crucial to recognize and value the role of human expertise in preserving the integrity and spirit of Arabic poetry. Human translators possess a deep understanding of the intricacies of the Arabic language, cultural references, and the nuances of poetic expression (Risku, 1998). They bring their creativity, intuition, and sensitivity to the task of translation, ensuring that the beauty and essence of the original work are not lost. AI should be seen as a tool to augment human effort, providing suggestions, aiding in the linguistic aspects, and facilitating the translation process, rather than replacing the human translator.

The use of AI in translating Arabic poetry holds broader implications for the preservation and dissemination of cultural heritage. By making Arabic poetry more accessible to non-Arabic speakers, AI translation can foster cross-cultural understanding and appreciation. It can facilitate dialogue, enhance literary exchange, and serve as a bridge between diverse linguistic and cultural communities. However, it is essential to strike a balance between embracing AI's potential and preserving the authenticity and cultural richness of the original work. Ongoing research, collaboration between AI experts and human translators, and the compilation of high-quality datasets are crucial for developing AI models tailored specifically to the unique challenges of translating Arabic poetry.

The rise of artificial intelligence has opened up new possibilities for translating Arabic poetry, a task that has long been considered complex and challenging. While AI holds immense potential in capturing the beauty and essence of Arabic poetry, it is vital to acknowledge its limitations and the importance of human expertise in preserving the cultural richness and nuances of the art form. Through a balanced approach that leverages AI as a powerful tool in the hands of skilled translators, we can foster a deeper appreciation and understanding of Arabic poetry across linguistic boundaries, enriching the world of literature and cultural exchange.

Poetry translation involves particular difficulties because of the intricate analogies, symbolism, and poetic elements used. The original author's

creative essence must be conveyed while keeping in mind the subtleties of both language and culture. To help students overcome these obstacles while translating the aesthetic spirit present in Arabic literary works into English, educators have long-researched strategies. The strategic integration of artificial intelligence (AI) translation technologies is one promising area that has gotten little attention.

Due to increased processing power and vast datasets used to train machine learning models, AI systems have achieved significant advancements in generic text translation (Koehn, 2009). But because poetry depends so heavily on original language, it presents unique challenges. While AI might have trouble understanding symbolic intent, it has skills like part-of-speech recognition, rhyme scheme detection, and numerous interpretations of ambiguous sentences that could be useful to a human translator (Moussalem,2019).

AI could assist translators at different phases of their work. AI may assist in surfacing cultural or historical allusions during pre-translation analysis to provide the appropriate context. When writing first drafts, the depiction of poetic elements discovered by AI could scaffold learners' analysis. AI's capacity to recognize and explain metaphors, suggest rhyme schemes or offer synonyms for complex symbolic language may enhance rather than replace human creativity. Then, feedback and remark features could direct improvement. A post-translation assessment that uses AI to identify places that need to be changed or human judgment of relevance could improve the results.

Instead of just outsourcing the activity, effective integration would position AI as a collaborative assistant focused on the natural translation talents of working learners. Prior studies on AI-assisted language learning show that skill development is most successfully facilitated when technology is appropriately integrated into pedagogical objectives and activities (Heilesen,2010). However, comprehensive research into the best methods for translating poetry still has to be done, especially when it comes to works in languages like Arabic that are culturally significant. As educators explore artificial intelligence to tap its potential while avoiding pitfalls, valuable insights could emerge from examining AI’s impact on translating poetry with profound cultural resonance, such as the outstanding works of Dr. Ghazi Al-Qusaibi.

Study Objectives

The study aims to compare the quality of translations produced by Saudi students when translating one of Dr. Ghazi Al-Qusaibi's poems with and without the assistance of AI tools. This will provide insights into how AI may help or hinder learners' ability to accurately convey symbolic meaning, cultural references, and poetic elements.

Significance of the Study

This study is significant as it explores the impact of AI-based translation tools on Saudi English as a Foreign Language (EFL) learners' translation of Dr. Ghazi Al-Qusaibi's poetry. It sheds light on the effectiveness and potential benefits of using AI in the translation process, providing valuable insights for language learners and educators.

Research Questions

To address the gaps in the existing studies, the current study seeks to answer the following questions:

RQ1. How effective is AI compared to human translation in translating poems by Ghazi Al-Qusaibi from Arabic into English?

RQ2. How does AI-based translation compared to human translation influence different parameters?

Literature Review

Arabic poetry poses significant challenges for machine translation because of its complex linguistic structure and cultural nuances. Early studies revealed AI had difficulty rendering stylistic characteristics of figurative Arabic poetry while translating it literally (Ghareb & Elimam, 2018). Arabic poetry has a complicated morphology, syntax, and rhetorical devices, which Alotaibi and Hussein (2017) examined as obstacles in AI translation. However, including human judgment appears to have the potential to maximize both the advantages of humans and technology.

One of the pioneering studies was conducted by Hassan and Mahyoub (2015), who built a statistical machine translation (SMT) system using the Moses toolkit. When trained on a generic Arabic-English parallel corpus and evaluated on 20 poems ranging from 3-20 lines, the translations often failed to preserve intended meanings or poetic devices despite word-level accuracy. This highlighted SMT's inability to capture poetry's deep contextual and

cultural aspects. Alhendawi et al. (2017) developed a rule-based system incorporating syntactic parsing and lexical selection rules. Evaluated on 20 short poems, it showed limited capability for metaphorical language but provided initial proof-of-concept. However, both rule-based and early data-driven models struggled with poetry's nuanced semantic and aesthetic qualities. This revealed the need for more advanced approaches leveraging larger language understanding.

Researchers began applying powerful neural models to larger parallel corpora. Alqahtani et al. (2017) used over 12 million United Nations documents to train an attention-based NMT model, finding significantly better performance than Hassan and Mahyoub's (2015) SMT baseline both quantitatively and qualitatively. Extending such work, Elmadany and Wayne (2018) created a new parallel poetry corpus by scraping existing translations and utilized it to fine-tune Transformer and LSTM models. Automatic and human evaluations showed that fine-tuned systems preserved more subtleties than generic baselines. However, limitations remained around structural fluency and conveying nuanced cultural references.

To address poetry's resource scarcity, researchers leveraged transfer learning and domain adaptation techniques. Amarin and AlNajjar (2021) employed unsupervised NMT to translate between monolingual poetry collections without parallel data. Self-learning techniques yielded improved BLEU scores over baselines, demonstrating generative capabilities when transferring learning across languages. However, conveying intricate artistic spirit remained difficult without direct pairings.

Recent research examined the interaction between humans and AI. An AI-assisted instructional framework was created by Al-Ghamdi et al (2023) that uses examples of annotated Arabic poetry to explain translation techniques to learners. To enhance the depiction of poetic elements, Oudah and Shaalan (2020) presented a hybrid approach integrating rule-based and statistical techniques with human post-editing. Al-Zoghby and Elshamy (2021) used automatic methods combined with human verification to rate the readability of Arabic poetry that had been machine-translated.

To enhance AI models, larger parallel corpora are also required. To create annotated datasets, Khalifa et al. (2019) automatically identified poetic

devices in Arabic poetry. Similar to this, isolating common metaphors from poetry that is aligned in Arabic and English aids in pattern recognition (Al-Zoghby & Elshamy, 2021). By emphasizing important elements in instructional texts, AI may help students learn more effectively (Khalifa et al., 2019).

Evaluation of aesthetic aspects in translation remains difficult due to the challenge of quantifying them objectively. Baquees & Mohan (2020) found that incorporating Part-of-Speech tags and syntactic parsing improved AI translation of Arabic literature. Alotaibi (2021) also used Transliterated Arabic text to facilitate AI training for resource-poor languages. In conclusion, a collaboration between human expertise and AI computational abilities is crucial for optimizing machine translation of poetry. Further efforts should be made to develop specialized Arabic language resources and evaluation metrics to better represent the aesthetic qualities of translated Arabic verse globally.

Early studies primarily used BLEU and human evaluations to assess meaning accuracy. Tawfik et al. (2022) proposed a multidimensional framework incorporating metrics like semantic similarity alongside dimensions of fluency, faithfulness, and creativity/artistic merit.

Applying this to various models fine-tuned on poetry data revealed limitations in capturing complex stylistic elements and generating content preserving emotional resonance, despite strong semantics. This emphasized evaluation must account for poetry's subjective, non-literal nature beyond just correctness. Recent works have explored integrating poetic form priors during decoding. Hasan et al. (2022) proposed a generative model conditioning on source-side rhythmic patterns via constrained beam search, qualitatively improving cadence in translations. Chen et al. (2023) pre-trained large vision-language models on image-text pairs depicting Arabic poems' cultural contexts. Fine-tuning for translation showed potential for nuanced renditions by grounding generations in historical understanding beyond text alone. However, full stylistic control across languages remains elusive.

Methodology

The influence of AI on translating Dr. Ghazi Al-Qusaibi's poem entitled "Tell Her" was evaluated in this study by utilizing a research design that combined experimental methods. A group of 50 undergraduate Saudi EFL

students was randomly assigned to either use AI translation tools (experimental group) or not use any tools (control group). Both groups translated the poem with similar complexity and theme.

Robert Frost once remarked: "Poetry is what is lost in translation. It is also what is lost in interpretation" (Untermeyer, 1964, p. 18). The evaluation of poetry translation is as challenging as the translation of poetry. Keeping in mind the evaluation approaches of literary translation, we have used an eclectic practical model for evaluating the features of both the form and the content of the target text to gain an objective summative translation quality assessment of the translated poem in this study. This model draws upon ten parameters which have been widened due to considering the nature of translating poetry. The first six parameters are general translation assessment parameters, namely accuracy of meaning, fluency, syntactic features, cohesion and coherence, style, and diction (Al-Qinai 2000, Banat and Abu Adla, 2023). Four more parameters have been taken into account, viz. figurative language, rhyme and rhythm, tone, and genre. These criteria are necessarily used to assess the elements of poetry in the translated poem. We hope that these ten parameters will be sufficient to make a comprehensive assessment of the target poetic text.

Participants

A total of 50 EFL undergraduate students from Saudi Arabia with a comparable degree of translation proficiency participated in this study. All the individuals were Arabic native speakers and specialized in English-related disciplines. The present study only recruited college seniors to ensure that they had engaged in the related classes and were equipped with AI-based translation technology.

Procedure and Data Analysis

One of Al-Qusaibi's poems entitled "Tell Her" was given to students enrolled in translation departments at one of the Saudi universities. The participants in the experimental group translated the poem using AI and then did post-editing before submitting the final product of the translation. The participants in the control group translated the poem using traditional methods such as dictionaries before submitting their final product of translation. The output of the two groups was scored according to a rubric. The scores in each

group were calculated and compared to determine the effectiveness of AI in translating the poems.

Parameters

To gain a valid reliable assessment, the used evaluation parameters were defined precisely to make their concepts clear and specific. The definitions are as follows:

Table 1. Assessment Parameters and Their Definitions

No.	Parameters	Definitions
1.	Accuracy of meaning	The degree to which the translation conveys the intended meaning of the original text.
2.	Fluency	It refers to how natural the translation sounds in the target language.
3.	Syntactic features	It refers to the grammatical features such as word order, sentence structure, verb-subject agreement, voice, number, gender, person, modality, tense, and aspect.
4.	Cohesion and coherence	Cohesion and coherence assess how well the translation is organized and connected. Coherence measures the degree of thematic symmetry and cohesion deals with reference (anaphora, cataphora), ellipsis, deixis, and conjunctions.
5.	Style	The suitability of the translation for the intended audience and purpose.
6.	Diction	It refers to the choice of words. It examines the use of the most appropriate words and expressions to convey the meaning of the original text accurately.
7.	Figurative language	The use of rhetorical devices like metaphor, irony, metonymy, personification, imagery, etc.
8.	Rhyme and rhythm	A rhyme is a repetition of similar-sounding words, occurring at the end of lines in poems while rhythm refers to meter and patterns of stressed and unstressed syllables in poem lines.
9.	Tone	It refers to the literary speaker's emotion and attitude and the ability of the translator to transfer them into the TT.
10.	Genre	It designates a literary type that determines how a work is ordered and organized.

Assessment Scale

Moreover, we have applied a five-level scale that can ensure objectivity and manageability as recommended by Pollitt (1991). This proposed scale is adopted from several sources including Waddington (2001), Williams (2001 & 2004), and Medadian and Mahabadi (2015). The scale is as follows:

Table 2. Five-Level Assessment Scale

No.	Parameters	Levels				
		1	2	3	4	5
1.	Accuracy of meaning	Inadequate transfer of ST meaning, undermined by serious inaccuracies	Adequate transfer of general ST meaning, but with many lapses in accuracy.	Adequate transfer of general ST meaning, but with a few lapses in accuracy	Almost complete transfer of ST meaning; there may be one or two insignificant inaccuracies	Complete transfer of St meaning
2.	Fluency	Mostly unable to render the poem adequately in TL	Almost the entire poem reads like a translation.	Certain parts of the poem read like a piece originally written in TL, but others read like a translation.	Most of the translated poem reads like a piece originally written in TL.	The whole translated poem reads like a poem originally written in TL.
3.	Syntactic features	The grammatical features of the poem are mostly inaccurate leading to communication failure in many parts of the poem.	The translated poem contains frequent grammatical mistakes that disrupt communication.	The translated poem is syntactically good but there are a few grammatical mistakes that hinder communication.	The translated poem is syntactically good but there are occasional grammatical errors that do not hinder communication.	The whole translated poem is syntactically perfect.
4.	Cohesion and coherence	The poem lacks cohesion and is mostly incoherent.	The poem has thematic discontinuity and lack of cohesion at the level of sentence and stanza.	The poem has some sentence-level cohesion and a frequent lack of stanza-level coherence and cohesion.	The poem has good sentence-level cohesion, and some stanza-level coherence and cohesion.	The translated poem is fully coherent, and it is cohesive on both sentence and stanza levels.
5.	Style	The translated poem is mostly unsuitable for the intended audience and purpose.	Many parts of the translated poem are not suitable for the intended audience and purpose.	Some parts of the translated poem are not suitable for the intended audience and purpose.	The translated poem is mostly suitable for the intended audience and purpose.	The translated poem is fully suitable for the intended audience and purpose.

No.	Parameters	Levels				
		1	2	3	4	5
6.	Diction	Mostly inaccurate vocabulary, communicating few clear meanings, and mostly irrelevant to content	frequently inaccurate vocabulary communicating some clear ideas, occasionally relevant to the content	occasionally inaccurate words, mainly communicating ideas clearly, and overall relevant to the content.	One or two inaccurate words. Vocabulary mainly conveys the ideas clearly, and overall relevant to content.	Accurate word choice conveys meanings. Vocabulary is relevant to the content.
7.	Figurative language	Figurative language pieces are rarely rendered in the translated poem.	Many figurative language pieces are not rendered in the translated poem.	Some Figurative language pieces are not rendered in the translated poem.	Figurative language pieces are mostly transferred into the translated poem.	Figurative language pieces are successfully rendered into the translated poem.
8.	Rhyme and rhythm	The translated poem has neither rhyme nor rhythm.	Most of the lines do not have rhyme and rhythm.	Some lines have rhyme and rhythm.	Most of the lines have rhyme and rhythm.	The translated poem has perfect rhyme and rhythm as per the TL poetic rules.
9.	Tone	The transfer of the literary speaker's emotion and attitude has almost failed.	The literary speaker's emotion and attitude are not mostly transferred into the target poem.	Some literary speaker's emotions and attitudes are not transferred into the target poem.	The literary speaker's emotion and attitude are mostly transferred into the target poem.	The literary speaker's emotion and attitude are transferred into the target poem successfully.
10.	Genre	The translated poem is mostly not ordered and organized as per the rules of the poetic genre of the TL.	Many parts of the translated poem are not ordered and organized as per the rules of the poetic genre of the TL.	Some parts of the translated poem are not ordered and organized as per the rules of the poetic genre of the TL.	The translated poem is mostly ordered and organized as per the rules of the poetic genre of the TL.	The translated poem is ordered and organized as per the rules of the poetic genre of the TL.

Results

RQ1. How effective is AI compared to human translation in translating poems by Ghazi Al-Qusaibi from Arabic into English?

To answer the first research question, descriptive and inferential analyses were performed. The results are shown in Table 1. These results compare the effectiveness of human translation without AI-based assistance versus AI-based translation. The mean values indicate the average effectiveness of each method. The mean score for AI-based translation (33.4) is higher than that of human translation without AI assistance (27.67), suggesting that, on average, AI-based translation performs better in translating the poems used in the study. The standard deviation values show the variability or spread of the data around the mean. In this case, both methods have relatively similar standard deviations (4.42 for human translation and 4.49 for AI-based translation), indicating that the variability in performance within each group is comparable.

To find out if the difference was significant between the two groups, a t-test was performed. The t-statistic was 12.428. It indicated the ratio of the variance between groups to the variance within groups. In this case, the t-value was significant, indicating that there were significant differences between the groups. The significance value associated with the t-statistic was .001, which is less than the conventional alpha level of .05. This suggested that the differences between the groups are statistically significant.

Table 3. Descriptive analysis of the total scores for each group

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		T	Sig.
					Lower Bound	Upper Bound		
Control	15	27.6667	4.41858	1.14087	25.2197	30.1136	12.428	.001
Experiment	15	33.4000	4.48888	1.15902	30.9141	35.8859		
Total	30	30.5333	5.25871	.96010	28.5697	32.4970		

RQ2. How does AI-based translation compared to human translation influence different parameters?

To find out the effect of AI-based translation on each category, descriptive and inferential analyses were performed. The results are shown in Table 2.

Each aspect was rated on a scale of 5 scores. The experimental group (Mean=3.43) scored higher than the control group (Mean = 3.13) in accuracy of meaning, suggesting that the experimental condition led to better accuracy in conveying meaning. However, the difference in accuracy of meaning between groups was not statistically significant ($t(1, 28) = 2.181$, $p = .151$). The experimental group (Mean = 3.3667) outperformed the control group (Mean = 2.8667) in syntactic features, indicating better adherence to formal writing conventions.

There was a statistically significant difference in syntactic features between groups ($t(1, 28) = 5.268$, $p = .029$). The experimental group (Mean=3.23) scored higher than the control group (Mean = 2.53) in fluency, suggesting smoother and more coherent writing in the experimental condition. The difference in fluency between groups was statistically significant ($t(1, 28) = 11.873$, $p = .002$). Similar to fluency, the experimental group (Mean = 3.23) performed better than the control group (Mean = 2.73) in cohesion and coherence, indicating stronger connections between ideas. There was a statistically significant difference in cohesion and coherence between groups ($t(1, 28) = 5.921$, $p = .022$). The experimental group (Mean=3.4000) exhibited a higher score in style compared to the control group (Mean = 2.7333), suggesting a more sophisticated or polished writing style. The difference in style between groups was statistically significant ($t(1,28) = 11.618$, $p = .002$). The experimental group (Mean = 3.3333) had a higher rating in diction compared to the control group (Mean = 2.6000), indicating more precise and appropriate word choice. There was a statistically significant difference in diction between groups ($t(1, 28) = 15.193$, $p= .001$) The experimental group (Mean = 3.5667) scored higher than the control group (Mean = 3.1333) in the use of figurative language, suggesting greater creativity or effectiveness in using figures of speech. There was a statistically significant difference in the use of figurative language between groups ($t(1,28) = 5.915$, $p = .022$). The experimental group (Mean = 3.8000) had a higher score in tone compared to the control group (Mean = 3.2000), indicating a more effective or appropriate tone in writing. The difference in tone between groups was statistically significant ($t(1, 28) = 8.591$, $p = .007$). The experimental group (Mean = 2.6000) outperformed the control group (Mean = 1.8000) in rhyme and rhythm, suggesting better use of these

elements in the experimental condition. There was a statistically significant difference in rhyme and rhythm between groups (t(1, 28) = 13.440, p = .001). The experimental group (Mean = 3.4333) scored higher than the control group (Mean = 2.9333) in adhering to genre conventions, indicating better alignment with the expectations of a particular writing genre. There was a statistically significant difference in genre between groups (t(1, 28) = 4.245, p = .049).

Table 4. Scores of control and experimental groups for each category

		Mean	Std. Deviation	95% Confidence Interval for Mean		Df	T	Sig.
				Lower Bound	Upper Bound			
Accuracy of meaning	Control	3.1333	.63994	2.7789	3.4877	1, 28	2.181	.151
	Experiment	3.4333	.45774	3.1798	3.6868			
Syntactic features	Control	2.8667	.63994	2.5123	3.2211	1, 28	5.268	.029
	Experiment	3.3667	.54989	3.0621	3.6712			
Fluency	Control	2.5333	.51640	2.2474	2.8193	1, 28	11.873	.002
	Experiment	3.2333	.59362	2.9046	3.5621			
Cohesion and coherence	Control	2.7333	.59362	2.4046	3.0621	1, 28	5.921	.022
	Experiment	3.2333	.53005	2.9398	3.5269			
Style	Control	2.7333	.59362	2.4046	3.0621	1, 28	11.618	.002
	Experiment	3.4000	.47056	3.1394	3.6606			
Diction	Control	2.6000	.50709	2.3192	2.8808	1, 28	15.193	.001
	Experiment	3.3333	.52327	3.0436	3.6231			
Figurative language	Control	3.1333	.51640	2.8474	3.4193	1, 28	5.915	.022
	Experiment	3.5667	.45774	3.3132	3.8202			
Tone	Control	3.2000	.67612	2.8256	3.5744	1, 28	8.591	.007
	Experiment	3.8000	.41404	3.5707	4.0293			
Rhyme and rhythm	Control	1.8000	.56061	1.4895	2.1105	1, 28	13.440	.001
	Experiment	2.6000	.63246	2.2498	2.9502			
Genre	Control	2.9333	.79881	2.4910	3.3757	1, 28	4.245	.049
	Experiment	3.4333	.49522	3.1591	3.7076			

In sum, the results indicated that the use of AI-based translation had a significant effect on syntactic features, fluency, cohesion and coherence, style, diction, figurative language, tone, rhyme and rhythm, and genre. However, there were no statistically significant differences in accuracy of meaning between groups.

Discussion

The results of the descriptive quantitative analysis show that, on average, the AI-based translation is better than the human translation. They also indicate that the differences between the groups are statistically significant and that the variability in performance between groups is comparable. In this section, we will shed some light on the performance variability of both study groups in each aspect of translation as per the criteria that are used in the assessment.

Accuracy of Meaning

Meaning is the basic unit of translation; therefore, it cannot be compromised in any way in the process of rendering the SL meaning into the TL text. Though the AI-aid translation group scored higher than the human translation group, the difference in the accuracy of meaning between groups is statistically insignificant. This proves that both study groups seriously focused on transferring the meaning into the translated poem. The below translations of the following line are evidence of this fact:

Table 5. Accuracy of Meaning

Original Line	Human Translation	AI-based Translation
سكب الدهر من أساه رحيقا فتحساه جُرعة إثْر جُرعةْ	1.a Of his sorrow many cups rejuvenated, And he drank the biggest glass.	1. b Time poured essence from his wounds, Opening them with each drop.
	2.a Time poured from his vessel a fermenting drink, He tasted it in draught after draught.	2.b Time has poured its sorrows on him, He feels it with every sip.
	3.a Eternity poured from its grief nectar, He sipped it dose after dose.	3.b Time poured from its chalice a potion, He tasted drop by drop.

The above translation examples clearly convey the meaning of the SL poetic line; moving from a freer translation in which the focus is on the spirit of the text, not the form of the text (Newmark 1988), as in 1.a and 1.b; to a direct translation in which the focus is on transferring the meaning and taking into account the form of the TL (Ghazala 1995), as in 2.a and 2.b; to a one-to-one literal translation in which besides transferring the meaning, each SL word or phrase has an identical equivalent in the TL text with the same

number, grammatical class, and type of language (Newmark 1988), as in 3.a and 3.b. The common feature in all these translations is that all of them efficiently render the meaning from the SL to the TL. Based on the above illustration, we can confidently conclude that the difference between the human translation and the AI-based translation in terms of meaning accuracy is insignificant.

Fluency

The second aspect of comparison is fluency. It gauges the naturalness of the translation in TL. The statistical analysis shows that the AI-based translation group outperformed the human translation group. The following translation extracts highlight the differences between the two groups:

Table 6. Fluency

Original Line	Human Translation	AI-based Translation
قل لـها .. إنه يـهيم .. وأخـشى أن تـواريـه رحـلة دون رجـعةْ	4.a Tell her that I see him torn, And fear his journey deems no return.	4.b Tell her he is infatuated and I fear he will Encounter a departure without a return.
	5.a Tell her that he is lost and I am afraid That he will take a trip that will never return.	5.b Tell her that he is overwhelmed, and he fears That his journey might end without return.
	6.a Tell her that he is lost and I'm worried That he might go away without returning	6.b Tell her that he wanders and fears That a journey might shroud him without a return

The examples of AI-based translation are more natural than the examples of human translation in terms of vocabulary, grammatical structures, and style. The reason behind this is that AI-based translation is a human translation supported by ChatGPT that can generate text with a high level of fluency. The findings of this study on the aspect of fluency are in line with those of Banat and Abu Adla (2023). However, the first example of human translation (4.a) is a free translation which is as natural as any other example of AI-based translation, but it remains an exceptional translation by a talented participant when we consider the overall performance of both groups.

Syntactic Features

The third aspect of comparison is the syntactic features. Using the appropriate TL grammatical forms is an important factor that can affect the translated text. The results of the statistical analysis show that the AI-based translated poem has fewer grammatical mistakes as compared to the human-translated poem. The examples below highlight this point:

Table 7. Syntactic Features

Original Line	Human Translation	AI-based Translation
قـل لـهـا .. إنه تـأمَّـل في دنـيـاه حـيـناً فـعـاد يـحـضـنُ دمـعـه	7.a Tell her he is meditate in his world once He came embraces his tears.	7.b Tell her, he contemplated his worldly life, For a while and then started holding back his tears.
	8.a Tell her he is meditating in his world When he hugs his tears.	8.b Tell her that he contemplated his worldly life For a while, then he returned embracing his tears.
	9.a Say to her he ruminate in his life for a second Then recur embracing his tears.	9.b Tell her that he contemplated his life, For a while, and then he embraced his tears.

However, the AI-based translations are not quite syntactically perfect, but they are better than the human translations.

Cohesion and Coherence

The statistical analysis and the nine examples (1-9) provided above prove that AI helped in producing more coherent and cohesive translated poems in the TL. The AI-aid translated poems are more organised and connected. They have better referential compatibility and thematic symmetry. These findings are in line with those of Banat and Abu Adla (2023) which confirm that ChatGPT is capable of producing TL translations that are contextually relevant and coherent.

Style

The suitability of the translated text for the intended audience and purpose is another significant criterion in translation assessment. The statistical analysis shows that the AI-based poem translations are more suitable for the TL poetic purpose and audience. The following examples support this finding:

Table 8. Style

Original Line	Human Translation	AI-based Translation
قل لـها .. إنّه يـفيق على جرح وتـغفـو سـنيـنه فـوق لـوعـةْ	10.a Tell her that he wakes up from an injury And his years will pass above his pain.	10.b Tell her he awakens to a wound's sting, While years doze atop the pain they bring.
	11.a Tell her that he recovers from a wound, And his years cover his pain.	11.b Tell her, he awakens to a wound, And slumbers with years above its ache.
	12.a Say to her that he wakes up on a hurt, And sleep years deeply on hurt.	12.b Tell her that he awakens to a wound, And dozes off above his agony for years.

The AI-based translations are more poetic in style than the human translations. Moreover, they serve the poetic purpose and convey the speaker's long-lasting agony. The TL audience can easily understand the meaning of the AI-based translations and feel the emotion they have. On the other hand, it is difficult for the TL audience to capture the speaker's idea and pain in the human translations due to their poor TL style.

Diction

The next parameter is diction which refers to the appropriate choice of words. The words that are used in the AI-based translations are more precise and appropriate. For example, in the last line of the SL poem, the word "يـهـيم" is translated as "lost, torn, wandering, adorning, deeming, whispering, and enchanted" in the human TL translations; whereas in the AI-based translations, it is translated mainly as "wanders" and in three translations as "infatuated, overwhelmed, and lost." Another instance is the translation of the word "جرح" which is translated as "wound" in the AI-based translations and as "wound, hurt, and injury" in the human TL translations. The diction of the AI-based translations of the poem is more literary and thus fits the poetic genre of the TL.

Figurative Language

The language of poetry is necessarily figurative. Transferring the figures of speech from the SL to the TL is quite challenging because certain figures of speech are idiomatic, culturally specific, and connotative. Certain images lose their poetic beauty and powerful effect when they are translated. Robert

Frost once remarked: "Poetry is what is lost in translation. It is also what is lost in interpretation" (Untermeyer, 1964, p. 18). Take for instance the SL image in the expression "يحضنُ دمعه" which is translated literally by most of the participants as "embracing his tears" instead of "holding back his tears" to maintain the image and the effect it creates at the expense of fluency, making them possibly unattainable by the native TL reader. Another example is the simile in the second line of the SL poem:

Table 9. Figurative Language

Original Line	Human Translation	AI-based Translation
راعه أنَّ عمره يتلاشى مثل ما تُخمد الأعاصير شمعةْ	13.a He fears that his life is fading away, Like the storms extinguish a candle's flame.	13.b He was alarmed that his life was fading away, Like a candle extinguished by a storm.
	14.a He was terrified that his life was vanishing Like hurricanes blowing out a candle.	14.b Fearful that his life was fading away, Like a candle extinguished by time's whim.
	15.a He realized that his life was fading away, Like storms putting down a candle.	15.b He worries that his life fades away, Like how storms extinguish a candle.

The statistical analysis draws our attention to the significant impact of using ChatGPT as an AI-translation tool and how it helped students achieve greater creativity and effectiveness in rendering the SL figures of speech in the TL poem translations.

Rhyme and Rhythm

One of the prominent features of poetry is the use of rhyme and rhythm. However, this feature is language-specific which makes its rendering from the SL to the TL quite difficult during the process of translation. Moreover, the SL poem is classical and strictly follows a single rhyme making it hard for the student to render this feature in the TL poem. Another challenge that the participants in this study faced is that they are not poets. Only two participants could create poem translations that are rhythmic and rhymed to a great extent, while other participants could partly achieve that. In general, the AI-based poem translations have better rhyme and rhythm as compared to the human poem translations. Below are two examples of the poem translations in the TL:

Table 10. Rhyme and Rhythm

Original SL Poem	Human TL Translation	AI-based TL Translation
قل لها .. إنه تـأمَّـل في دنـيـاه حـيـناً فـعـاد يـحـضـنُ دمـعـه راعه أنَّ عـمـره يـتـلاشـــى مثل ما تُخمد الأعاصير شمـعةْ وصباه يـضيع منه .. كـما ضاع نـداء .. تطوي المتاهات رجعه قل لـها .. إنّه يـفـيق على جرح وتـغـفـو سـنـيـنه فـوق لـوعـةْ سـكـب الـدهر من أسـاه رحـيـقا فـتـحـسـاه جُـرعـة إثْـر جُرعـةْ قل لـها .. إنه يـهـيم .. وأخـشى أن تـواريـه رحـلة دون رجـعةْ	Confess to her that he contemplated, A life of tears and joys and laughs. His tears had fallen desolated, Thinking of his time and past. Let her thoughts be regulated, For his years are deemed to pass. As a storm when infuriated, Eats the candle and the ash. The youth he lost and demonstrated, Is an echo in an uneven path. Tell her he awakens desolated, As his passions sleep on the snath. Of his sorrow many cups rejuvenated, And he drank the biggest glass. Tell her that I see him torn, And fear his journey deems no return.	Tell her that he contemplated his world, As he embraced his tears once more. Beware that his life is fading away, Like hurricanes extinguish a candle's core. His youth slips away, lost and astray, A call folds the mazes, retracing the way. Tell her he awakens to a wound's sting. While years doze atop the pain they bring. Time poured from his woes like honey, Each drop tasting the aftermath, so uncanny. Tell her he wanders, fearing the chance. She might hide him in a departure's dance.

The rhyme scheme of the human TL translation is *ab ac ad ae af af ac gg* and the rhyme scheme of the AI-based TL translation is *ab cb cc dd ee ff*. The lines have unequal lengths in both translations and thus have an irregular meter. The rhythm patterns are varied yet mostly iambic. It is worth mentioning here that the creation and maintenance of the rhyme scheme in these two translations were sometimes achieved at the expense of meaning. Overall, the AI-based poem translations are better than the human poem translations in terms of rhyme and rhythm.

Tone

The transfer of the emotion(s) and attitude(s) from the SL poem to the TL poem is very crucial in a successful poetry translation. The SL poem has deep feelings of worry, sadness, agony, despair, and longing. Both study groups could convey these emotions in the TL poem. The examples 1 – 15 provided in the tables above reflect these poetic feelings and how they were transferred by both groups. The speaker's feelings are reflected in the TL words and expressions used in the translations, like "fear, worried, pain, agony, torn, lost, wandering, wound, terrified, ache, etc. Anyway, they also show that the AI-based translation group outperformed the human translation group in this regard. This finding is confirmed by the statistical analysis which illustrates the significant impact of the ChatGPT in enabling the students to successfully transfer the literary speaker's emotions and attitudes into the TL poem.

Genre

The order and organization of a TL text are determined by the genre of the SL text. The SL poem that is used in this study is lyrical. The poet speaks to an unspecified person. The English poetry genre that suits this kind of poem is the ode. The six long lines of the SL poem are transformed into six couplets. The division of the TL couplets follows the SL line which is divided into two halves. Another possible way of organization can be according to the thematic order indicated by the phrase "قـل لـهـا" (Tell her). Accordingly, the TL poem will have two stanzas of six and four lines respectively, followed by a couplet. Most of the students used the first division. Some students in the human translation group fell into the trap of translating the SL poetry line as a TL prose sentence, which consequently affected the poetic language and structure of the TL poem. Another poetic feature of the SL poem is that all lines have the same rhyme which is quite rare in English language poetry and thus difficult to be transferred into the TL poem. As indicated in the findings of the statistical analysis, ChatGPT helped the AI-aide translation group to overcome these challenges and thus produce well-organized TL poems that match the poetic TL genre of the poem.

Conclusion, Limitations, and Future Research Direction

Conclusion

This study investigated the impact of AI-based translation tools on Saudi EFL learners' ability to translate poetry by renowned Arab poet Dr. Ghazi Al-Qusaibi from Arabic to English. Overall, the findings provide valuable insights into how strategic integration of AI during the translation process can help students overcome some of the inherent complexities of poetry translation. The results indicated that using AI tools led to statistically significant improvements in several dimensions of translation quality compared to translating without AI assistance. It was noticed that the translations created by learners in the experimental group who used AI scored statistically higher than those in the control group who did not use AI. This held true across most dimensions evaluated including syntactic features, fluency, cohesion/coherence, style, diction, figurative language usage, tone, rhyme/rhythm, and genre conventions. The only aspect lacking a statistically significant difference was the accuracy of meaning, indicating both groups focused heavily on preserving core semantics. This confirms that AI tools can enhance many technical and creative aspects of literary translation for learners. Thus, the study demonstrated that AI can play an augmented role for language learners rather than a replacement one when translating literary works. By aiding comprehension of poetic devices, vocabulary, style and providing feedback, AI tools have the potential to strengthen learners' translation skills if appropriately incorporated into the learning process. However, rigorous human verification is still crucial to preserve cultural authenticity and assess subjective aesthetic impact.

In conclusion, this study offered initial empirical evidence that AI translation tools can positively impact EFL learners when appropriately positioned as collaborative partners within instruction. With balanced utilization and consideration of limitations, AI shows promise to augment human efforts in translating Arabic poetry for wider cross-cultural understanding and literary exchange. By leveraging each other's relative strengths, the combined efforts of skilled human translators and AI technologies can help foster a deeper appreciation of the art form beyond linguistic boundaries. However, additional research expanding sample populations, source texts, and assessment methods could provide a more generalizable understanding of AI's applications and boundaries.

Limitations

One of the main limitations of this study is that it focused on translating only one Arabic poem by Dr. Ghazi Al-Qusaibi. Evaluating the translation of a single poem limits the generalizability of the findings, as the effects of using AI translation tools may vary depending on the complexity, style, theme, or other characteristics of the original text. Translating multiple poems across different genres and topics would have provided more robust insights. Another constraint of this study stems from its reliance on undergraduate participants, resulting in unforeseen outcomes. Subsequent researchers are encouraged to enlist professional translators to ensure more accurate findings. Another limitation is that the study only involved Saudi EFL learners from one university. Using a larger and more diverse sample of translators from multiple universities and language proficiency levels could affect the results.

Future Research Direction

This paper recommends a balanced approach that leverages AI as a powerful tool to assist human translators in translating Arabic poetry. While AI holds great potential in capturing the nuances and essence of Arabic poetry, the paper emphasizes that human expertise is vital in preserving the cultural richness and authenticity of the original work. Besides, the paper suggests that ongoing research, collaboration between AI experts and human translators, and the compilation of high-quality datasets are crucial for developing AI models that are tailored specifically to the unique challenges of translating Arabic poetry. It also recommends incorporating human feedback and judgment at various stages of the translation process, such as pre-translation analysis, first draft writing, and post-translation assessment. Additionally, the paper calls for an integral relationship between AI and human translators, where technology augments and enhances the translation process rather than replacing the creative and cultural expertise of the human translator. It's important to highlight that the parameters and scales employed in analyzing the results were devised by the researchers specifically for this paper. Enhancing these parameters and scales in future studies is highly recommended.

Acknowledgment

This research received grant no. (120/2023) from the Arab Observatory for Translation (an affiliate of ALECSO), which is supported by the Literature, Publishing & Translation Commission in Saudi Arabia.

References

Alhendawi, A., Way, A., & McCreath, E. (2017). A rule-based machine translation approach for translating Arabic poetry. *Procedia Computer Science*, 117, 11-18. https://doi.org/10.1016/j.procs.2017.10.003

Al-Ghamdi, S., Al-Khalifa, H., & Al-Salman, A. (2023). Fine-tuning Bert-based pre-trained models for Arabic dependency parsing. *Applied Sciences*, 13(7), 4225.

Alotaibi, M. S. (2021). Improving Arabic Machine Translation Using Transliterated Data. *Journal of Computational Science*, 54, 101-213.

Alotaibi, S., & Hussein, A. H. (2017). Challenges in Arabic poetry machine translation: a survey. *Ain Shams Engineering Journal*, 8(4), 311-318.

Alqahtani, S., Aslam, F., & Mubarak, H. (2017). Neural machine translation of Arabic poetry. *Procedia Computer Science*, 117, 42-49. https://doi.org/10.1016/j.procs.2017.10.007

Al-Qinai, J. (2000). Translation Quality Assessment. Strategies, Parameters and Procedures. *Meta*, 45(3), 497–519. https://doi.org/10.7202/001878ar

Al-Zoghby, A. A., & Elshamy, A. A. (2021). Readability assessment for Arabic poetry translation. *Future Computing and Informatics Journal*, 6(1), 100063.

Amarin, R., & AlNajjar, S. (2021). Unsupervised neural machine translation of Arabic poetry. *Computers and Electrical Engineering*, 93, 107258. https://doi.org/10.1016/j.compeleceng.2021.107258

Banat, M., & Abu Adla, Y. (2023). Exploring the Effectiveness of GPT-3 in Translating Specialized Religious Text from Arabic to English: A Comparative Study with Human Translation. *Journal of Translation and Language Studies*, 4(2), 1-23.

Baquees, M. N., & Mohan, M. (2020). Leveraging syntactic information for improving neural machine translation of Arabic literature. *Procedia Computer Science*, 172, 339-346.

Chen, L., Yang, Z., Roth, D., & Xing, E. P. (2023). Vision-language models to generate culturally-grounded literary translations. *Transactions of the Association for Computational Linguistics*, 11, 56-72. https://doi.org/10.1162/tacl_a_00500

Dessì, D., Recupero, D. R., & Sack, H. (2021). An assessment of deep learning models and word embeddings for toxicity detection within online textual comments. *Electronics*, 10(7), 779.

Elmadany, A., & Wayne, A. (2018). Machine translation of Arabic poetry: Developing a poetry translation corpus and domain adaptation. *Procedia Computer Science*, 142, 160-168. https://doi.org/10.1016/j.procs.2018.10.365

Ghareb, I., & Elimam, A. (2018). Evaluation of Google Translate in translating figurative Arabic poetry into English. *Advances in Language and Literary Studies*, 9(1), 142-148.

Ghazala, H. (1995). *Translation As Problems and Solutions*. Beirut: Dar El-Ilm Lilmalayin.

Hasan, F., Datt, A., Sarkar, A., Neubig, G., & Bangalore, S. (2022). Poetic language generation with rhythmic and semantic constraints. *Computer Speech & Language*, 76, 101405. https://doi.org/10.1016/j.csl.2022.101405

Hassan, H., & Mahyoub, A. (2015). Statistical machine translation for Arabic poetry: Challenges and limitations. *Procedia Computer Science*, 65, 687-694. https://doi.org/10.1016/j.procs.2015.09.082

Khalifa, H., Al-Ajlan, A., Abdulghani, M. A., Idris, K. B. M., & Al-Madani, H. (2019). Automatic detection and visualization of poetic devices in Arabic poems. *Information*, 10(3), 92.

Koehn, P. (2009). A process study of computer-based translation. *Machine Translation*, 23(4), 241–263.

Medadian, G., and Mahabadi, D., (2015). A Summative Translation Quality Assessment Model for Undergraduate Student Translations: Objectivity Versus Manageability. *Studies About Languages*. (26), 41-53.

Moussalem, D. S. (2019). Towards a Better Automatic Translation for Arabic Poetry. *Procedia Computer Science*, 150, 577-584.

Oudah, M., & Shaalan, K. (2020). A hybrid approach for improving Arabic poetry machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3), 1-22.

Pollitt, A. (1991). Response to Charles Alderson's Paper 'Bands and Scores'. In Alderson, Ch., North, B. (eds.), 1991. *Language Testing in the 1990s*. London: Macmillan, pp.87–94.

Risku, H. (1998). Transediting: A new concept for translation quality research. *Translation and Meaning*, 2(3), 393-413.

Tawfik, M., Samih, Y., Kholy, A. E., Elmadany, A., & Magdy, W. (2022). Towards a systematic framework for evaluating machine translation of poetry. *Computer Speech & Language*, 74, 101317. https://doi.org/10.1016/j.csl.2022.101317

Untermeyer, L. (1964). *Robert Frost: A Backward Look*. Washington: Library of Congress.

Van der Meer, J. (2019). Can AI solve poetry translation?. *Nature*, 575(7781), 12-14.

Waddington, C. (2001). Should Translations be Assessed Holistically or Through Error Analysis? Hermes, *Journal of Linguistics*, 26 (2001), pp.15–38.

Williams, M. (2001). The Application of Argumentation Theory to Translation Quality Assessment. *Meta: Journal des traducteurs/ Translators' Journal*, 46(2), pp.326–344.

____, (2004). *Translation Quality Assessment: An Argumentation-Centred Approach*. University of Ottawa Press.