



# Controlling Translation Purpose through Prompt Engineering: A Skopos-Based Study of AI Poetry Translation

Li Haijin<sup>1,\*</sup>

Yogyakarta State University, Indonesia

Received: 03.04.2026 • Accepted: 02.06.2026 • Published: 05.07.2026 • Final Version: 30.07.2026

**Abstract:** Large language models (LLMs) have rapidly expanded the horizon of AI-assisted literary translation, yet the mechanisms of how prompts construct translation purpose are theoretically underexplored. This study, based on Vermeer’s Skopos Theory, conceptualizes prompts as a “digital translation brief” and examines the systematic regulation of four function-oriented prompts: baseline (P0), reader-oriented (P1), form-oriented (P2), and culture-oriented (P3) in DeepSeek’s translation strategies for English-to-Indonesian poetry translation, with Emily Dickinson’s *Hope is the Thing with Feathers* as the source text. The study finds, through comparative close reading across the four prompt conditions, as well as against two published human translations, that each prompt orientation produces distinct and observable shifts in diction, imagery construction and formal expression: P1 enhances literary register and emotional intensity at the cost of collocational stability; P2 replicates structure but may trigger unprompted source-language retention, which undermines target-reader adequacy; P3 allows discourse-level metaphorical coordination and selective dependence on the source text. Human translators’ choices emerge from cultural memory, aesthetic intentionality, and poetic agency. DeepSeek’s functional adaptability is largely a matter of probabilistic generation and prompt compliance. The findings re-conceptualize prompts not only as technical input instructions, but as purposive regulators of translation, offering a translation-theoretic framework for analysing LLM behaviour and practical guidance for designing Skopos-informed prompts for AI-assisted literary translation.

**Keywords:** prompt engineering, Skopos theory, poetry translation, DeepSeek, English-Indonesian translation

## 1. Introduction

The fast development of large language models (LLMs) is having a profound impact on translation practice and research. Generative AI systems like ChatGPT and GPT-4 have evolved from language processors to intelligent systems able to perform complex translation tasks (Kwok et al., 2025). But the effectiveness of LLMs in literary translation, particularly in poetry translation, which is one of the most challenging fields, remains controversial.

Translating poetry is not just a matter of conveying meaning but rather of a deep restructuring of metaphorical density, systems of images, rhythms and cultural associations (Bassnett, 2002). While LLMs have been shown to perform well on automatic evaluation metrics (Jiao et al., 2023), works have shown that they have severe limitations when it comes to emotional expression, cultural

\* Corresponding Author: [lihai.2025@student.uny.ac.id](mailto:lihai.2025@student.uny.ac.id)

metaphor reproduction, and stylistic control (Ed-Dali, 2025; Pratiwi & Gusthini, 2025). Hence, when the aim of translation changes, as towards reader orientation, form preservation or cultural adaptation, can AI match poetic strategic decisions like human translators? This question is largely unanswered.

The basic tenet of Skopos Theory “the purpose determines the strategy” (Vermeer, 1978) serves as a powerful theoretical lever to analyze how prompts control the AI translation behavior. Therefore, this study employs Skopos Theory as the theoretical framework, proposing a new perspective of “prompts as a digital translation brief”, and examines the control of DeepSeek’s behavior by various function-oriented prompts in translating English-Indonesian poetry. This study aims to comparatively analyze AI-generated translations and human translations in order to: (1) empirically explore the systematic shift in the model’s translation strategies under different prompts (reader-oriented, form-oriented, culture-oriented); and (2) draw the demarcation line between the functional flexibility of prompt-controlled LLMs and the inception of human poetic agency (the translator’s capacity to make intentional aesthetic, rhythmic, and culturally interpretive decisions beyond literal semantic transfer). This study reconceptualizes prompt design as a functional regulatory mechanism, providing a theoretical reference for prompt optimization in AI-assisted literary translation.

## **2. Literature Review**

### **2.1 Functionalist Translation Theory: A Theoretical Framework**

Skopos Theory was proposed by Vermeer. In his 1989 work “Skopos and Commission in Translational Action,” Vermeer argues that the primary determinant of any translation action is its purpose (Skopos)—that is, the function the target text is intended to fulfill in the target context. Different purposes lead to different translation strategies, linguistic forms, and functional orientations.

Vermeer’s work was later extended by Nord (1997), who introduced the concept of the “translation brief.” The brief, provided by the client, specifies the target readers, usage scenario, and expected function of the translation. Based on this, the translator decides on the balance between fidelity, readability, and cultural adaptation. At the same time, Nord’s “loyalty principle” emphasizes that while pursuing functional goals, the translator must maintain ethical responsibility toward the source text author, target readers, and client, thus preventing functionalism from sliding into relativism.

Applying Skopos Theory to AI translation research has unique theoretical value. In the LLM translation environment, users specify target readers, style, and cultural orientation through prompts. This means that prompts effectively take on the function of a translation brief, influencing the model’s translation strategies by presetting the translation purpose. Yamada (2026), in his study “Agentic AI Translate,” uses human-AI collaboration to guide users in constructing a structured translation brief based on Skopos theory, audience, and register. He directly converts translation meta-language, such as Skopos, audience, domestication, and foreignization into instruction codes for LLMs. Kolat (2025), in a study on poetry and song translation, reveals that different functional purposes directly affect the choice of translation strategies. Wang (2025), using political texts as a case study, compares DeepSeek with official translations from a Skopos perspective, confirming the feasibility of applying Skopos theory to analyze DeepSeek translations.

### **2.2 Literary Translation and Prompt Engineering Research on LLMs**

The application of large language models (LLMs) in the field of literary translation has evolved from general texts to very complex fields such as poetry. While research demonstrates that LLMs are effective at leveraging discourse-level context for literary translation, they still make major errors, particularly in their treatment of culture-specific metaphors and emotional nuances (Karpinska & Iyyer, 2023). Ed-Dali (2025) systematically evaluated the performance of DeepSeek R1 and ChatGPT 4.5 in Arabic-English literary translation. While DeepSeek outperformed ChatGPT in automatic metrics such as BLEU, COMET, and BERTScore, the qualitative analysis highlighted its weaknesses in pragmatic coherence and emotional depth, especially in poetry and drama texts. Klumpp (2026) further quantitatively analysed rhyme and meter in English-to-German translations of Shakespeare's sonnets. They found that LLMs are still behind professional human translators, but also that their performance is heavily affected by prompt strategies.

In the realm of English-Indonesian poetry translation, Pratiwi and Gusthini (2025) pitted ChatGPT against the version of a Dickinson poem translated by Indonesian Abdul Mukhid, and proved that AI is far inferior to humans in the expression of metaphor and emotion. Karjo (2025) compared QuillBot, Google Translate, and ChatGPT for translating the works of the Indonesian poet Sapardi Djoko Damono into English and concluded that none of the three could fully capture the aesthetic qualities of the original poem. Alowedi and Al-Ahdal (2023) in their study of Arabic-English poetry translation, further point out that machine translation cannot fully understand the socio-cultural background of poetic creation, a finding that also applies to literary translation practices in other non-major languages such as Indonesian.

Prompt engineering has evolved from the early days of zero-shot and few-shot prompting to a complex system involving multiple strategies, including role-setting, chain-of-thought reasoning, and theory embedding (Hooper, 2025). The mechanism of prompt influence in poetry translation is especially hard. Klumpp (2026) showed that iterative prompt refinement can improve the ability of LLMs to retain poetic form, but this strategy often comes at the cost of grammatical correctness and semantic fidelity. In the scope of modern poetry translation, Wang et al. (2024) also point out that augmenting the number of examples does not necessarily lead to improved translation outcomes, which reveals the challenge in consistently applying traditional prompting models due to the highly creative nature of poetry translation. However, the current prompt research mainly concentrates on translation quality assessment and output optimisation, e.g., the GEMBA-MQM framework proposed by Kocmi and Federmann (2023) for LLM translation error detection, rather than exploring from the perspective of translation theory how prompts are involved in the construction of translation purposes.

Synthesizing the existing literature, this study identifies the following gaps: Current research mostly targets general LLMs such as ChatGPT, and has not systematically examined DeepSeek's specific performance in English-Indonesian poetry translation. Moreover, prompt engineering studies largely remain at the level of outcome evaluation (translation quality, accuracy, etc.), with little in-depth analysis from a translation theory perspective on how prompts "construct" AI's translation purpose. Although scholars recognize that prompts can "control" translation output, there is still no translation-theoretical explanation of how this control mechanism systematically affects translation strategies, linguistic choices, and functional orientations. Meanwhile, the application of Skopos theory to AI translation remains at an initial conceptual validation stage, with no controlled experimental study that systematically tests the effects of different function-oriented prompts (reader-oriented, form-oriented, culture-oriented) on poetry translation strategies.

### 3. Methods

This study employs a qualitatively oriented comparative text analysis method, supplemented by limited counts of lexical and translation strategies, to investigate how different prompts influence the translation strategies and functional orientations of large language models in poetry translation. Drawing on Vermeer's Skopos Theory, prompts are defined in this paper as the digital counterpart of a translation brief in the LLM translation environment. Accordingly, the study examines how the AI model adjusts its translation strategies, registers, imagery processing, and form preservation in response to the functional requirements embedded in different prompts, as well as the resulting poetic differences across distinct functional orientations.

#### 3.1. Corpus Selection

The source text is Emily Dickinson's classic poem *Hope is the Thing with Feathers*, taken from *The Complete Poems of Emily Dickinson* (Dickinson, 2016). This poem has the following features:

- (1) Dense metaphorical and symbolic imagery;
- (2) Simple syntax but distinct formal features;
- (3) Strong literary register and interpretative space for culture.

These features make it suitable for observing how AI adjusts its strategies and functions under different translation purposes.

In order to explore the relationship between prompts and translation purpose construction, the study generates AI translations under four conditions: a baseline translation with no additional prompts (P0), a reader-oriented prompt emphasizing emotional and poetic readability (P1), a form-oriented prompt emphasizing structural preservation and formal fidelity (P2), and a culture-oriented prompt emphasizing cultural adaptation and literary reconstruction (P3). The design of these prompt types refers to common functional orientations in literary translation studies, including reader reception, formal fidelity, and cultural adaptation.

The prompts are as follows:

- (1) P0 (Baseline Prompt)  
"Please translate the following English poem into Indonesian."
- (2) P1 (Reader-oriented Prompt)  
"Please translate the following English poem into Indonesian for Indonesian high school students who have no background in English literature. Prioritize readability and emotional impact."
- (3) P2 (Form-oriented Prompt)  
"Please translate the following English poem into Indonesian, preserving the original line breaks, sentence structure, and literal meaning of key images as closely as possible."
- (4) P3 (Culture-oriented Prompt)  
"Please translate the following English poem into Indonesian, adapting cultural references and imagery to resonate with contemporary Indonesian readers and cultural context."

The analysis proceeds mainly through comparisons across the four prompt conditions (P0–P3). In examining the translations, the study focuses on differences in diction, imagery construction, and form preservation, as these factors are closely related to functional orientation and poetic expression in poetic texts. These dimensions are not treated as separate analytical categories but as observational

indicators for analyzing how different prompts affect the functional orientation and stylistic choices of the translations.

As large language models generate outputs probabilistically, output stability was considered during the research process. For each prompt condition, the translation was generated twice using DeepSeek's web interface under the default settings available at the time of data collection. The styles of two outputs were highly similar, with only minor lexical variations that did not affect the overall functional orientation of the translations. Therefore, the first output was selected as the primary dataset for detailed analysis.

The study does not seek to establish deterministic prompt–output relationships. Rather, it examines recurring functional tendencies associated with different prompt designs. Given the proprietary and continuously evolving nature of commercial LLMs, exact replication of outputs cannot be guaranteed. The findings should therefore be interpreted as observations of prompt-induced functional orientations rather than permanently stable characteristics of the model.

To further explore the differences between AI literary translation and human literary translation, the study includes two published Indonesian translations as reference texts (Catatan kecil, 2016). Abdul Mukhid's translation is included for its clear tendency toward rhythm and emotional expression, while Wawan Eko Yulianto's translation is included for its notable features of cultural adaptation and imagery reconstruction. These two translations are selected to compare the different poetic and functional strategies employed by human translators beyond literal semantic transfer.

**Table 1.** Human Translator Reference Translation Comparison

| Abdul Mukhid (HT1)   | Wawan Eko Yulianto (HT2)  |
|--|---|
| <i>Asa</i>   | <i>Harap</i>  |
| <i>Asa itu bagaikan burung dan sayapnya<br/>yang bersemayam di jiwa,<br/>dengarkan irama<br/>tiada putusnya.</i>   | <i>Harap ialah sesuatu bersayap<br/>yang bertengger di jiwa,<br/>dan berdendang tanpa kata,<br/>dan tanpa putus-putusnya,</i>   |
| <i>Dalam terpaan angin kencang kukecap merdunya<br/>Dalam amukan badai serasa perihnya<br/>namun si burung mungil tetap setia<br/>tiada henti sebar hangatnya.</i>           | <i>dan terdengar merdu di deru topan;<br/>dan badai sungguhlah ganas<br/>jika sampai mengusir burung kecil itu,<br/>burung yang sebarakan hangat.</i>                                   |
| <i>Kudengar suaranya di negeri paling gigit<br/>dan di samudera paling musykil<br/>namun tak sedikit pun dariku<br/>ia pinta walau hanya secuil,<br/>walau hanya secuil.</i> | <i>Pernah kulihat ia di sedingin-dinginnya daratan,<br/>juga di seasing-asingnya lautan;<br/>tapi, biar cuaca seganas apa, tak pernah<br/>mulut menadah padaku, meski demi seremah.</i> |

### 3.2. Analytical Procedure

This study combines close reading and comparative textual analysis to examine the functional differences in diction, imagery construction, and formal expression across different prompt conditions, and further analyzes the different performances of AI and human translators in terms of poetic agency, aesthetic judgment, and cultural reconstruction.

## 4. Results and Discussion

The translations generated under different prompts are shown below. The analysis will be based on the following texts:

**Table 2.** Comparison of Translations Generated with Different Prompt Words

| English Original  | P0  | P1  | P2   | P3  |
|---|---|---|--|---|
| “Hope” is the thing with feathers -<br>That perches in the soul -<br>And sings the tune without the words -<br>And never stops - at all - | “Harapan” adalah sesuatu yang berbulu -<br>Yang bertengger di dalam jiwa -<br>Dan menyanyikan lagu tanpa kata -<br>Dan tak pernah berhenti - sama sekali -            | “Harapan” adalah makhluk berbulu<br>Yang bertengger di relung jiwa —<br>Melantunkan lagu tanpa kata —<br>Dan tak pernah berhenti — sama sekali —                            | “Hope” adalah makhluk berbulu -<br>Yang bertengger di jiwa -<br>Dan menyanyikan lagu tanpa kata -<br>Dan tak pernah berhenti - sama sekali -                               | “Harapan” adalah sesuatu yang berbulu —<br>Bersemayam di relung jiwa —<br>Melantunkan lagu tanpa kata —<br>Dan tak pernah berhenti — sama sekali —                                |
| And sweetest - in the Gale - is heard -<br>And sore must be the storm -<br>That could abash the little Bird<br>That kept so many warm -   | Dan paling merdu - terdengar di Badai -<br>Dan sungguh dahsyat badai itu -<br>Yang bisa membuat malu si Burung kecil<br>Yang telah menghangatkan begitu banyak jiwa - | Dan terdengar merdu — di tengah badai —<br>Betapa ganasnya badai yang menerjang —<br>Hingga mampu buat burung kecil itu gentar —<br>Yang selama ini hangatkan banyak jiwa — | Dan terdengar paling manis - di saat Badai -<br>Dan pastilah badai yang hebat -<br>Yang bisa membuat malu si Burung kecil<br>Yang telah menghangatkan begitu banyak hati - | Dan terdengar paling manis — saat badai menerpa —<br>Dan betapa ganasnya angin ribut —<br>Yang bisa membuat malu si Burung kecil<br>Yang telah menghangatkan begitu banyak hati — |
| I’ve heard it in the chilliest land -<br>And on the strangest Sea -<br>Yet - never - in Extremity,<br>It asked a crumb - of me.           | Aku mendengarnya di tanah paling dingin -<br>Dan di Laut paling asing -<br>Namun - tak pernah - dalam penderitaan,<br>Ia meminta sekerat roti pun dariku.             | Aku mendengarnya di negeri paling dingin —<br>Dan di lautan paling asing —<br>Namun — dalam kepedihan apa pun —<br>-Tak pernah — ia meminta sepuccuk roti dariku —          | Aku mendengarnya di negeri paling dingin -<br>Dan di Laut paling asing -<br>Namun - tak pernah - dalam kepedihan,<br>Ia meminta remah - dariku.                            | Kudengar kicauannya di negeri paling dingin —<br>Dan di lautan paling asing —<br>Namun — tak pernah — dalam penderitaan,<br>Ia meminta sedikit pun — dariku.                      |

#### 4.1. P0 (Baseline Prompt): Default Translation Purpose

Under the baseline condition (P0, no functional prompt; see Methods for the exact instruction), DeepSeek exhibits its default translation behavior, as shown below.

**Table 3.** Examples of Key Translation Choices In P0

| Source Text                         | P0 Translation                          | Explanation  |
|-------------------------------------|---|--|
| “Hope” is the thing with feathers - | “Harapan” adalah sesuatu yang berbulu - | Uses the more common standard Indonesian expression, rather than the more literary “Asa”                             |
| That perches in the soul -          | Yang bertengger di dalam jiwa -         | Uses basic spatial structure “di dalam jiwa” (in the soul), not the more poetic “relung jiwa” (recesses of the soul) |
| It asked a crumb - of me.           | Ia meminta sekerat roti pun - dariku.   | Turns the abstract “crumb” into “sekerat roti” (sliver of bread), enhancing clarity                                  |

The corpus comparison shows that P0, in the baseline condition with only a basic translation prompt, has an implicit translation tendency of “semantic fidelity first, restricted poetic reconstruction”. At the lexical level the model prefers frequent and accessible expressions in the target language. For example, the core concept “Hope” is translated into the everyday Indonesian word “Harapan”, rather than the more literary and aesthetically marked “Asa”. Similarly, “in the soul” is translated as “di dalam jiwa”, a neutral spatial term, rather than the more poetic “relung jiwa”.

The preference for linguistic stability and accessibility points to the model’s preference for a low risk translation strategy in the absence of an explicit translation brief. When faced with the weakly metaphorical image “crumb” P0 does not choose the more literary lexical equivalent “remah”, but

the semantically clearer and more widely accessible expression “sekerat roti”. It is not mainly for the sake of poetic aesthetics that this adjustment is made, but for the sake of semantic transparency and comprehensibility in the target language.

In terms of imagery and syntax, P0 also reflects a conservative retention of semantic structure from the source text. First, the model is prone to preserve only the surface semantic components of the original imagery, without significant metaphorical reconstruction. For example, “the thing with feathers” is rendered as “sesuatu yang berbulu” (feathered creature) and keeps the lexical item “feathers” in a mechanical way. However, in present-day Indonesian, this expression can also bring to mind animal fur or body hair, and does not necessarily activate the metaphorical image of a bird. The model thus preserves the lexical form but only partially recovers the underlying poetic imagery.

Second, P0 was very obedient to the source-text syntax organization and employed standard Indonesian word orders throughout the translation. The syntax is neutral and stable, with no major inversion, rhythmic rearrangement or literary distortion. Even formally, where P0 mechanically preserves visible features of punctuation such as the dashes at the ends of poetic lines. This is largely at the level of surface textual layout and not deeper poetic rhythm or aesthetic recreation.

As a general rule, P0 is characterized by a restrained register of written Indonesian, where the main function is to produce a comprehensible and semantically coherent poetic text, rather than a literary translation that is based on aesthetic reconstruction. In this sense, P0 gives an important baseline for understanding the default translational tendency of LLM under minimally regulated conditions. It also offers a comparative basis for the study of the functional and aesthetic changes that occur under the prompt-guided Skopos orientations of P1, P2 and P3.

## 4.2. Prompt Regulation of Translation Skopos

### Reader-oriented Skopos (P1)

When the translation prompt specifies a target readership and functional expectation (P1: “Translate the following English poem into Indonesian for Indonesian high school students who have no background in English literature. Prioritize readability and emotional impact”), DeepSeek’s translation strategy has a significant functional shift. The results show that, contrary to the conventional expectation that reader-oriented translation should imply lexical simplification or decreased linguistic complexity, the model does not reduce the stylistic register of the translation. But it systematically elevates the literary register of the target text with emotionally heightened and highly literary diction. Thus, in the model’s prompt-conditioned translation behaviour, “reader-friendliness” is not so much understood in terms of lexical accessibility, but in terms of the enrichment of literary and emotional experience. In other words, the model appears to associate reader engagement with poetic intensity and not linguistic simplification.

**Table 4.** Examples and Comparisons of Key Translation Choices In P1

| Source Text                            | P0  | P1   |
|--|---|--|
| That perches in the soul -             | Yang bertengger di dalam jiwa -           | Yang bertengger di relung jiwa —           |
| And sings the tune without the words - | Dan menyanyikan lagu tanpa kata -         | Melantunkan lagu tanpa kata —              |
| I’ve heard it in the chilliest land -  | Aku mendengarnya di tanah paling dingin - | Aku mendengarnya di negeri paling dingin — |

|                                  |  |   |
|----------------------------------|--|---|
| And sore must be the storm -     | Dan sungguh dahsyat badai itu -        | Betapa ganasnya badai yang menerjang —        |
| That could abash the little Bird | Yang bisa membuat malu si Burung kecil | Hingga mampu buat burung kecil itu gentar —   |
| It asked a crumb - of me.        | Ia meminta sekerat roti pun - dariku.  | Tak pernah — ia meminta sepucuk roti dariku — |

At the micro-lexical level, P1 shows a strong tendency to poetic elevation. In P1, compared to the baseline condition (P0), there is a frequent use of more literary expressions with higher emotional density and deeper spatial imagery. For example, the neutral spatial phrase “di dalam jiwa” is elevated to the more emotionally resonant “relung jiwa”; the common verb “menyanyikan” (sing) is replaced by the more poetic and chant-like “melantunkan” (chant); and the general term “tanah” (land) is transformed into the more elevated and macro-imagistic “negeri” (state). These lexical changes are highly coherent with the skopos-oriented function of the prompt as the model attempts to increase the emotional impact through literary restructuring of the target-language poetic environment.

Moreover, P1 is highly susceptible to emotional explicitation in both syntactic and semantic rephrasing. For example, the relatively descriptive “sungguh dahsyat” (really terrible) is replaced by the highly affective exclamatory structure ‘betapa ganasnya’ (how fierce). The more abstract psychological expression “membuat malu” (embarrassing) is reformulated as “gentar” (tremble), which projects a more immediate and embodied emotional response. These changes indicate that in a reader-centric functional framework the model is inclined to externalize the emotional implications encoded in the source text and transform them into more explicit affective signals in the target language, thereby enhancing the accessibility of emotional contents to readers.

To provide additional support for this qualitative observation, a simple lexical count was conducted across key translation points in the poem. Lexical choices were classified as Elevated/Literary (E), Neutral (N), or Marked (M). Here, *Marked* refers to expressions that are understandable in context but deviate from conventional collocational or stylistic patterns in contemporary Indonesian. The results are presented in Table 5.

**Table 5.** Lexical Register Differences between P0 and P1

| Lexical Item                     | P0                                     | P1                                   | Classification |
|----------------------------------|--|--------------------------------------|----------------|
| Hope                             | Harapan                                | Harapan                              | N              |
| thing with feathers              | sesuatu yang berbulu                   | mahluk berbulu                       | N              |
| perches in the soul              | bertengger di dalam jiwa               | bertengger di relung jiwa            | E              |
| sings the tune without the words | menyanyikan lagu tanpa kata            | melantunkan lagu tanpa kata          | E              |
| sweetest in the Gale is heard    | paling merdu terdengar di badai        | terdengar merdu di tengah badai      | E              |
| sore must be the storm           | sungguh dahsyat badai itu              | betapa ganasnya badai yang menerjang | E              |
| abash the little Bird            | membuat malu si burung kecil           | buat burung kecil itu gentar         | E              |
| chillest land                    | tanah paling dingin                    | negeri paling dingin                 | E              |
| that kept so many warm           | telah menghangatkan begitu banyak jiwa | selama ini hangatkan banyak jiwa     | M              |

Note: E = Elevated/Literary; N = Neutral; M = Marked. For example, *relung jiwa* (“recesses of the soul”), *melantunkan* (“chant”), and *negeri* (“realm”) were classified as elevated due to their higher literary register, whereas *hangatkan banyak jiwa* was classified as marked because of its non-standard collocational pattern.

As shown in Table 5, two lexical items remain largely neutral across the two prompt conditions, while seven display noticeable shifts in register or stylistic markedness. Among these, six involve movement toward more elevated or literary expressions, and one results in a stylistically marked collocation. Although descriptive rather than statistical, this pattern provides additional support for the observation that the reader-oriented prompt tends to encourage register elevation.

But this aggressive strategy of intensification also produces a kind of stylistic overcompensation, resulting in local linguistic instability. The most obvious example is the translation of “crumb” into “sepucuk roti” (a piece of bread). However, the classifier *sepucuk* in Indonesian is typically used with nouns such as letters (*surat*), firearms (*senjata*), or abstract hopes (*harapan*), but not conventionally with bread (*roti*). The combination with the concrete noun “roti” (bread) thus represents a marked collocational deviation. This odd couple, on the macro-literary level, produces a certain effect of poetic defamiliarization, but on the micro-linguistic level, it also produces a certain semantic awkwardness. This is what also happens in the sentence “hangatkan banyak jiwa” where the direct quantifying of “jiwa” (soul) as a substitute for human beings makes the collocation sound less natural in the standard Indonesian usage.

In general, the textual patterns observed for P1 indicate that prompts oriented to the reader can significantly activate the model’s tendencies for literary generation and move it beyond the conservative orientation for information transfer observed in P0. At the same time, however, this “emotion-first” approach also reveals a weakness of LLM literary translation: in pursuit of emotional and aesthetic intensification, the model may sacrifice collocational stability and local textual naturalness in the target language. This tension reveals the challenge posed for LLMs to reconcile the achievement of a function and the linguistic convention of the target language in the context of highly affective literary translation tasks.

### Form-oriented Skopos (P2)

P2 advocates for “keeping the original line breaks, sentence structure and literal meaning of key images as close as possible.” P2’s formal instructions are more explicit, visible, and operationalizable than the more abstract affect-oriented goal of P1. Hence, DeepSeek responds more directly and consistently to this prompt condition. This suggests that large language models are more adept at carrying out identifiable structural and formal characteristics, but struggle to form stable generation strategies for abstract aesthetic goals such as “emotion” or “poeticity”.

**Table 6.** Examples and Comparisons of Key Translation Choices In P2

| Source Text                         | P0                                      | P2                              |
|-------------------------------------|---|---------------------------------|
| “Hope” is the thing with feathers - | “Harapan” adalah sesuatu yang berbulu - | “Hope” adalah makhluk berbulu - |
| It asked a crumb - of me.           | Ia meminta sekerat roti pun - dariku.   | Ia meminta remah - dariku.      |

P2 preserves very much the structural appearance of the original poem at the level of visual form. For example, the English word “Hope” is kept in quotation marks (“Hope” is ...) and the original line breaks must be strictly preserved, so that each English line corresponds to one Indonesian line and vice versa, without merging or splitting. Moreover, the translation consistently maintains the dash “-” as a visual cue, thus producing a layout very similar to the source text. These features indicate that DeepSeek can execute low-level and explicitly recognizable formal instructions with significant consistency and precision.

At the lexical level, P2 clearly increases the tendency towards literal equivalence. In line with this form-oriented Skopos, DeepSeek replaces the reader-oriented explanatory rendering “sekerat roti” found in P0 with the more literal “remah” for “crumb”. Although “remah” is semantically closer to the original lexical form, in the context of contemporary Indonesian poetry it is more likely to be understood as a physical bread crumb, thus weakening the abstract symbolic implication carried by the source text. This suggests the model’s capacity to mitigate naturalization and literary accommodation when the prompts explicitly foreground formal and literal fidelity, to the detriment of closer lexical correspondence with the source text.

What is especially interesting is that P2 preserves the English word “Hope” in the opening line. The form-oriented prompt only requested the retention of line breaks, sentence structure, and the literal meaning of key images, but not the retention of source-language lexical items. The untranslated “Hope” can therefore be read as an instance of unprompted source language retention rather than prompt over-compliance. This behavior suggests that when formal fidelity is foregrounded, DeepSeek may associate source-text preservation not only with structural features but also with greater visibility of source language elements. The hybridized expression “*Hope*” *adalah makhluk berbulu* maintains a strong link with the source text but at the same time introduces a certain degree of hybridity into the target text. This example indicates that prompt-controlled LLM translation might, under certain circumstances, prefer source-language preservation even if not explicitly asked in the prompt.

According to Skopos Theory, this phenomenon shows a conflict between the visibility of the source text and the accessibility of the target reader. The preservation of the title “Hope” enhances the feeling of formal closeness to the original poem, but it also risks reducing immediate intelligibility for Indonesian readers not familiar with English literary discourse. This is not only a matter of formal fidelity, but of different translation priorities. The translation does not fully weaken communicative transparency in the target language, but succeeds in preserving aspects of source-text form.

More generally, the P2 results imply that prompt-controlled literary translation can achieve a high degree of structural fidelity when formal instructions are explicit. However, the preservation of “Hope” also suggests that the final translation is determined not only by the prompt, but also by implicit model biases. This example is not a straightforward case of prompt compliance but it illustrates the interaction of prompt guidance and model specific translation behaviour in the process of creating literary translations. The resulting text is thus a negotiated result between the functional orientation encoded in the prompt and the model’s own tendencies towards source-text preservation.

### **Culture-oriented Skopos (P3)**

When the translation prompt is oriented toward macro-cultural adaptation (P3: “adapting cultural references and imagery to resonate with contemporary Indonesian readers and cultural context”), DeepSeek’s output shows a significant shift from sentence-level correspondence to discourse-level reconstruction. As opposed to P0, which is still largely constrained by local semantic equivalence, and P2, which is focused on structural replication, P3 adopts a more dynamic generation strategy based on metaphorical coherence and target-cultural reception at the discourse level. The translation that results is therefore much more strongly oriented to literary and cultural norms in the target language.

**Table 7.** Examples and Comparisons of Key Translation Choices In P3

| Source Text | P0 | P3 |
|-------------|----|----|
|-------------|----|----|

|                                       |   |   |
|---------------------------------------|---|---|
| “Hope” is the thing with feathers -   | “Harapan” adalah sesuatu yang berbulu -   | "Harapan" adalah sesuatu yang berbulu —       |
| That perches in the soul -            | Yang bertengger di dalam jiwa -           | Bersemayam di relung jiwa —                   |
| I've heard it in the chilliest land - | Aku mendengarnya di tanah paling dingin - | Kudengar kicauannya di negeri paling dingin — |
| It asked a crumb - of me.             | Ia meminta sekerat roti pun - dariku.     | Ia meminta sedikit pun — dariku.              |

P3 considerably broadens the literary and aesthetic register of the target language at the level of lexical regulation and stylistic selection, approaching the conventions of contemporary Indonesian poetic discourse. The lexical contrasts in Table 7 clearly elevate the poetic register. For example, the verb “bersemayam” (reside) has a more solemn and spiritual literary tone than the relatively neutral “bertengger” (perched), and is more commonly related to elevated poetic or ceremonial contexts. Similarly, the spatial phrase “relung jiwa” (the depths of the soul) evokes a much stronger sense of psychological depth and emotional resonance than the more neutral “di dalam jiwa”. Furthermore, the auditory image “kicauan” (chirping) concretely and explicitly reinforces the underlying metaphor of “hope as a bird” than the neutral “lagu” (song). These micro-level adjustments suggest that the culturally oriented prompt effectively reorients the model’s translational Skopos, promoting not only literary elevation but also more metaphorical coherence for the Indonesian target readers.

Furthermore, P3 exhibits considerable discourse-level control in the domain of what can be called cross-line imagery coordination. In the original poem, the third stanza includes the relatively neutral auditory expression “I’ve heard it” without specifying birdsong. But creatively P3 adds the noun “kicauannya” (its chirping) to the translation: This addition is not a straightforward linear transfer of vocabulary from the source text, but a sort of metaphorical explicitation, rooted in the wider imagistic network of the poem. The model improves the continuity and perceptibility of the bird metaphor throughout the poem on the discourse level, not on the line-by-line lexical correspondence. Here, metaphorical coherence at the level of discourse takes precedence, to a certain point, over strict lexical fidelity.

P3 also shows a greater orientation to pragmatic equivalence in dealing with culturally marked imagery. P2 retains the literal “remah” before the highly materialized Western image “crumb”, emphasizing formal fidelity but weakening the abstract symbolic implication of the source text. P3, on the other hand, removes the surface visibility of the food image and recasts it as the more natural Indonesian pragmatic expression “sedikit pun” (not even a little). This adjustment removes the concrete food-related imagery, but it is more effective in communicating the implied meaning that “hope never asks for anything in return”. This means that in the case of culture-oriented prompt regulation the model prefers the naturalness in communication and pragmatic accessibility over the accurate preservation of source-text imagery.

However, such a local reconstruction does not fully free the translation from residual source-text dependence and a unique hybrid functional strategy is found. At the central idea of the poem, P3 successfully translates “Hope” into the usual Indonesian “Harapan” and thus avoids the sudden code-switching in P2. However, the translation in the immediately subsequent metaphorical structure still retains the more literal expression “sesuatu yang berbulu” (feathered creature) rather than a more interpretive metaphorical reconstruction such as Abdul Mukhid’s “bagaikan burung dan sayapnya” or Wawan Yulianto’s “sesuatu bersayap”. This pattern indicates that although the model is quite

flexible in coordinating local imagery and restructuring metaphors at the discourse level, its lexical generation still tends to be somewhat limited by linear correspondence in the source text.

In general, P3 shows a translation tendency focusing on cultural adaptation and discourse level metaphorical alignment. P3 emphasizes more on metaphorical continuity, pragmatic naturalness and target-oriented literary reconstruction, when compared to the emotional intensification of P1 and the structural replication of P2. These characteristics indicate that under the culture-oriented prompt regulation, LLM literary translation can partially break free from strict sentence-level literalism and present a more interpretative mode of poetic reconstruction. The persistence of selective source-text dependency also manifests the tension between cultural adaptation and lexical fidelity in prompt-regulated AI literary translation at the same time.

### 4.3. Differences Between AI and Human Translators

The P0–P3 analyses show that different types of function-oriented prompts can dramatically affect the translation performance of DeepSeek in terms of diction, imagery construction, and formal presentation, systematically modulating the language-generation behavior of the model. Yet whether this high level of functional adaptability can be equated with the poetic agency that characterizes human translators remains a critical question in defining the boundary between AI-assisted and human literary translation. To deepen the discussion of this issue, this study offers two representative Indonesian human translations (HTs) as comparative references to analyze the significant difference between LLM translation and human literary translation on the aspects of aesthetic judgment, cultural mediation, and poetic function realization.

While DeepSeek is able to rapidly shift stylistic tendencies in response to different translation briefings, its manifest “creativity” still depends to a great extent on probabilistic association and conditioned generation, rather than autonomous aesthetic intentionality. For example, the use of “kicauannya” (its chirping) in P3 is effective in reinforcing the metaphorical network of “hope as a bird” at the discourse level. However, this adjustment seems to be more related to statistically reinforced literary associations typical of poetic corpora, rather than to an independent artistic intention in the human literary sense. In other words, while AI may produce a literary expression that seems aesthetically plausible, its decision-making process is still largely determined by the probabilistic textual patterns encoded in the training data.

In contrast, human translators display a more purposeful kind of creative transgression and culturally embedded literary interpretation.

#### (1) Abdul Mukhid’s rhythmic reconstruction and formal deviation

Mukhid restructures the final lines as:

“namun tak sedikit pun dariku  
     ia pinta walau hanya secuil,  
     walau hanya secuil.”

The repetition of “walau hanya secuil” (even if only a tiny bit) is not present in the original poem and therefore cannot be explained in terms of literal correspondence. In the context of Indonesian poetry, the repetition serves as a rhythm and sound device which builds a long emotional rhythm, intensifying the subtle tenderness implied in the phrase “hope never asks for anything”. Abdul Mukhid’s choice is not a byproduct of AI-generated repetition coming from prompt-driven stylistic

generation, but a deliberate orchestration of rhythm and regulation of emotion for a higher-order aesthetic effect through selective deviation from the source structure.

## (2) Wawan Eko Yulianto's cultural behavior substitution and deep mediation

Wawan Eko Yulianto translates the final line as:

“mulut menadah padaku, meski demi seremah”

Here, “mulut menadah” (raising one's mouth in expectation or opening one's mouth to receive) essentially deviates from the literal object-based representation of “crumb” and rebuilds it as a culturally embodied social behavior. In Indonesian cultural memory and social imagination, this phrase conjures up associations with poverty, dependence, and silent pleading, and engenders an emotional resonance far more powerful than a direct lexical equivalent could ever hope to. Thus the translator ceases to be a mere decoder of linguistic signs and becomes a cultural mediator who reconstructs poetic experience through localized social imagery.

This distinction also points to a deeper difference between AI and human translation with respect to cultural mediation. In the whole corpus of MT versions the translation of “crumb” does not go much beyond the lexical level and local metaphorical adaptation. P0 chooses “sekerat roti” for semantic clarity at the expense of poetic subtlety; P2 keeps the literal “remah,” favoring formal equivalence but reducing symbolic implication; and P3 opts for the highly abstract “sedikit pun” to maximize communicative fluency. However, these variants are functionally different but still operate primarily within the framework of lexical substitution and local imagery adjustment.

In contrast, Wawan Eko Yulianto's “mulut menadah” does not simply substitute one lexical item for another, but rather reconstructs an entire socio-cultural scene. This kind of cultural mediation goes beyond the standard lexical equivalence or metaphorical substitution since it involves the recontextualization of socially and emotionally recognizable lived experience in the cultural environment of the target language.

Therefore, the functional differences between AI-powered poetry translation and human poetry translation can be summarized as follows.

**Table 8.** Functional Differences between AI and Human Poetry Translation

| Analytical Aspect      | DeepSeek (AI Translation)  | Human Translators   |
|------------------------|--|---|
| Translation Motivation | Responds to externally provided prompts and functional instructions    | Forms translation purpose through subjective literary interpretation            |
| Stylistic Adjustment   | Adjusts register and tone according to prompt constraints              | Actively rewrites, reshapes, or intensifies expression for poetic effect        |
| Imagery Construction   | Performs limited imagery expansion based on semantic association       | Reconstructs imagery through cultural imagination and contextual interpretation |
| Formal Handling        | Excels at replicating visible structural features and line arrangement | Negotiates between poetic rhythm, readability, and literary aesthetics          |
| Cultural Mediation     | Relies on probabilistic associations within training data              | Draws upon social experience, cultural memory, and reader awareness             |
| Poetic Agency          | Generates text through statistical prediction and prompt compliance    | Makes intentional aesthetic and emotional judgments                             |

The distinction between AI and human translation is not only a matter of the capacity to generate linguistic output but, more fundamentally, of the presence or absence of translator subjectivity.

DeepSeek demonstrates high functional flexibility in various translation briefs, especially in tasks with well-operationalized semantic and formal requirements. However, such flexibility is largely an artifact of the computational compliance to prompt requirements. By contrast, human translators move multidimensionally and nonlinearly between cultural memory, embodied experience, emotional cognition, reader expectation and poetic purpose. Literary translation therefore consists not only of interlingual transfer but also of the reconstruction of literary experience and socio-cultural meaning across contexts. Prompts can simulate different functional orientation well, but can not simulate the poetic agency of human literary translation.

## 5. Conclusion

This paper conceptualize prompts as a “digital translation brief” based on Skopos Theory and empirically show that function-oriented prompts systematically reshape DeepSeek’s translation strategies in English–Indonesian poetry translation. Reader-oriented prompts (P1) enhance emotional register at the cost of collocational stability; form-oriented prompts (P2) foster structural replication but may trigger unprompted source-language retention, which undermines target-reader adequacy; culture-oriented prompts (P3) allow for discourse-level metaphorical compensation and selective source-text dependence (see Section 4.2 for a detailed comparison). Compared to human translators, DeepSeek’s prompt-driven adaptability is statistically grounded, lacking the aesthetic subjectivity and cultural judgment characteristic of human literary translation.

Theoretically, this study extends Skopos Theory to AI-mediated environments by operationalising the “translation brief” as machine-readable prompts, providing a controlled framework to examine how functional orientations influence LLM generation. Methodologically, it offers an analytical model for prompt engineering beyond quality optimisation, viewing prompts as purposive regulators rather than mere input instructions. Practically, explicit Skopos-oriented prompts (e.g., specifying target readers or cultural adaptation) yield more differentiated outcomes than generic instructions, though form-oriented prompts may trigger unprompted source-language retention that risks reducing target-reader accessibility.

Crucially, this study differentiates between “functional realisation” and “subjective creation” in AI literary translation. LLMs can mimic various Skopoi, but cannot duplicate the interpretive agency of human translators who draw on cultural experience and aesthetic intention. This limitation opens opportunities for human–AI hybrid translation models that combine AI’s functional flexibility with human poetic judgment.

There are some limitations to note. Firstly, it is based on a single source text, Emily Dickinson’s *Hope is the thing with Feathers*. While suitable for the observation of prompt-driven functional variation, the poem’s relatively simple structure and universal theme may not reflect the full range of challenges in poetry translation. Second, the analysis is restricted to the English-Indonesian language pair. Therefore, the functional patterns identified here can be subject to the linguistic and cultural idiosyncrasies of language pairs and should not be generalized to the field of AI poetry translation in general. While a restricted stability check found consistent patterns, the variability of the output over multiple generations was not systematically studied. Future research should examine different literary genres, language pairs and LLM architectures, and should employ larger-scale generation experiments and reader-response measures to evaluate the reception of prompt-controlled translations.

## References

- Alowedi, N. A., & Al-Ahdal, A. A. M. H. (2023). Artificial Intelligence based Arabic-to-English machine versus human translation of poetry: An analytical study of outcomes. *Journal of Namibian Studies : History Politics Culture*, 33. <https://doi.org/10.59670/jns.v33i.800>
- Bassnett, S. (2002). *Translation studies*. Routledge.
- Catatan kecil: Puisi Hope – Emily Dickinson – Terjemahan dan Analisisnya. (2016). Catatan Kecil. <http://antonharyadi.blogspot.com/2016/05/puisi-hope-emily-dickinson-terjemahan.html>
- Dickinson, E. (2016). *The Complete Poems of Emily Dickinson: With Linked Table of Contents*. Simon and Schuster.
- Ed-Dali, R. (2025). Assessing DeepSeek R1 and ChatGPT 4.5 in Arabic-English literary translation: performance, challenges, and implications. *Cogent Arts & Humanities*, 12(1). <https://doi.org/10.1080/23311983.2025.2531183>
- Hooper, A. (2025). Prompt Engineering and In-Context Learning: A Comprehensive Survey of Techniques, Applications, and Future Directions. [Preprint]. *ResearchGate*. <https://doi.org/10.13140/RG.2.2.36738.59843>
- Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023). Is ChatGPT A Good Translator? A Preliminary Study. *ArXiv Preprint ArXiv:2301.08745*. <https://doi.org/10.48550/arXiv.2301.08745>
- Karjo, C. H. (2025). A Comparative Analysis Of Poetry Translation Using Three Ai Tools. *Prosiding Konferensi Linguistik Tahunan Atma Jaya (KOLITA)*, 23(23). <https://doi.org/10.25170/kolita.v23i23.7147>
- Karpinska, M., & Iyyer, M. (2023, December). Large language models effectively leverage document-level context for literary translation, but critical errors persist. In *Proceedings of the Eighth Conference on Machine Translation* (pp. 419-451).
- Klumpp, J. F. (2026, March). Quantitative Analysis of Rhyme and Metre in LLM-generated Translations of Poetry. In *Proceedings of the 10th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature 2026* (pp. 189-201).
- Kocmi, T., & Federmann, C. (2023, December). GEMBA-MQM: Detecting translation quality error spans with GPT-4. In *Proceedings of the Eighth Conference on Machine Translation* (pp. 768-775).
- Kolat, H. (2025). Göstergeler Arası Yolculuk: Şiir ve Şarkı Çevirilerinde Stratejiler ve İşlevsellik. *Söylem Filoloji Dergisi, Çeviribilim Özel Sayısı II*, 847–866. <https://doi.org/10.29110/soylemdergi.1602192>
- Kwok, H. L., Shi, Y., Xu, H., Li, D., & Liu, K. (2025). GenAI as a Translation Assistant? A Corpus-Based Study on Lexical and Syntactic Complexity of GPT-Post-Edited Learner Translation. *System*, 130, 103618–103618. <https://doi.org/10.1016/j.system.2025.103618>
- Nord, C. (1997). *Translating as a Purposeful Activity*. Routledge. <https://doi.org/10.4324/9781315760506>
- Pratiwi, R. I., & Gusthini, M. (2025). ChatGPT sebagai Alat Penerjemahan AI untuk Puisi “Hope is the Thing with Feathers.” *Jurnal Ilmu Sosial Dan Humaniora.*, 3(4), 513–526. <https://doi.org/10.57248/jishum.v3i4.589>
- Vermeer, H. J. (1978). Ein Rahmen für eine allgemeine Translationstheorie. *Lebende Sprachen*, 23(3), 99–102. <https://doi.org/10.1515/les.1978.23.3.99>
- Vermeer, H. J. (1989). Skopos and commission in translational action. In A. Chesterman (Ed.), *Readings in translation theory* (pp. 173–187). Finnlectura.
- Wang, L. (2025). A Comparative Analysis of AI Translation Quality from the Perspective of Skopos Theory: A Case Study of The Governance of China IV. *Lecture Notes on Language and Literature*, 8(2). <https://doi.org/10.23977/langl.2025.080201>
- Wang, S., Wong, D., Yao, J., & Chao, L. (2024, August). What is the best way for ChatGPT to translate poetry?. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 14025-14043).
- Yamada, M. (2026). Agentic AI Translate: An Agentic Translator Prototype for Translation as Communication Design. *arXiv preprint arXiv:2605.17041*.