

Research Article

# Prompt-oriented Output of Culture-Specific Items in Translated African Poetry by Large Language Models: An Initial Multi-layered Tabular Review

Adeyola Opaluwah\* 

Department of Translation Studies, University of Antwerp, Antwerp, Belgium

## Abstract

This paper examines the output of culture-specific items (CSIs) generated by ChatGPT 3.5 and ChatGPT Pro in response to three prompts to translate three anthologies of African poetry. The first prompt was broad, the second focused on poetic structure, and the third emphasized cultural specificity. To support this analysis, five comparative tables were created. The first and second tables presents the results of the CSIs produced by Chat GPT 3.5 and ChatGPT Pro respectively after the three prompts; the third table categorizes the unchanged CSIs based on Aixelá's framework of "Proper nouns and Common expressions"; the fourth summarizes the CSIs generated by the human translators, a custom-built translation engine (CTE), and the two versions of a Large Language Model (LLM). The fifth table shows how the seven CSIs that were repeated in translation in French were rendered after the three prompts. The sixth table shows the strategies employed by ChatGPT 3,5 and ChatGPT Pro after the culture-specific prompt on the CSIs that were not translated unrepeated. Compared to the outputs of CSIs from the reference human translation (HT) and the CTE in prior studies, the findings indicate that the culture-oriented prompts used with ChatGPT Pro did not yield significant enhancements in the CSIs during the translation of the three African poetry from English to French. On evaluation however, ChatGPT Pro scored better in BLEURT than ChatGPT 3.5. A combined total of 20 CSIs were generated by the LLM versions, where 13 were repeated as the source word. The repeated CSIs were inconsistent with the outcome of the HT and CTE; some of the translations of the remaining seven unrepeated CSIs were also inaccurate compared to the reference HT and CTE. While the corpus of this investigation is small, the results show that the data used to build LLMs has not been French-centric nor poetry domain-specific and thus LLMs could benefit from a higher and better performance when tailored to other languages and specific domains.

## Keywords

LLMs, ChatGPT 3.5, ChatGPT Pro, Structured Prompts, African Poetry, Translation

## 1. Introduction

Large Language Models (LLMs) are AI systems trained on vast amounts of text data to understand and generate human-like language. Utilizing neural networks and machine

learning algorithms, LLMs have transformed natural language processing (NLP) by enabling accurate and efficient language comprehension and generation. These models are trained on

\*Corresponding author: [adeyola.opaluwah@student.uantwerpen.be](mailto:adeyola.opaluwah@student.uantwerpen.be) (Adeyola Opaluwah)

**Received:** 17 March 2025; **Accepted:** 15 May 2025; **Published:** 18 June 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

diverse text sources such as Wikipedia, news, books, and social media, allowing them to identify linguistic patterns and relationships. This capability makes LLMs versatile, supporting applications like translation, chatbots, text summarization, and sentiment analysis. A significant advantage of LLMs is their ability to perform tasks without task-specific training, making them adaptable across fields like finance, healthcare, and education [11].

This paper's investigation extends prior research that analyzed the translation of Culture-specific items (CSIs) from English to French by General Machine Translation (GMT) and a custom machine translation engine (CTE) in three African anthologies. It evaluates explicitly how ChatGPT, a Large Language Model (LLM), translates CSIs in the same anthologies of African poetry authored by Wole Soyinka. The study employs three different prompts: the first prompt is a general translation request into French; the second prompt focuses on maintaining the poem's structure by preserving the stanzas; and the third is culture-oriented, asking for a translation that emphasizes cultural vocabulary. Initially, ChatGPT could not translate an entire anthology from English to French, leading to an upgrade to ChatGPT Pro after receiving notifications that only a few poems could be translated daily. It was believed that the Pro version would facilitate the translation of more poems, which prompted the request for a complete anthology translation. The response received from ChatGPT Pro is presented below.

*"To translate an anthology into French, you would typically need to follow these steps:*

- 1) Select the Text: Choose the specific excerpts or poems you want to translate.*
- 2) Context Understanding: Understand the themes and nuances of the original text, as literary translation often requires capturing tone and style.*
- 3) Translation: Translate the text into French, paying attention to vocabulary, grammar, and idiomatic expressions.*
- 4) Review and Edit: Revisit your translation for clarity and accuracy. It might be helpful to have a native speaker or professional translator review it.*
- 5) Finalize: Format the translated anthology appropriately, ensuring it's accessible to readers.*

*If you have a specific excerpt or text from the anthology you would like translated, feel free to share it!"*

Previous research has noted that DeepL Pro can translate a complete African anthology from English into French, leading to the belief that ChatGPT Pro would also possess the capability to translate an entire anthology. To identify and analyze the CSIs generated by the two versions of LLM, Aixelá's 1996 model was utilized. This model was previously applied in research [8] to analyze CSIs in human translations (HT), General Machine Translations (GMT), and translations from a custom-built Microsoft Azure translation model.

## 2. Context of the Translated Anthologies

Wole Soyinka's three anthologies *Early Poems*, *Mandela's Earth and Other Poems*, and *Idanre and Other Poems* were translated into French by Etienne Galle and André Bordeaux into *Cycles sombres*, *La terre de Mandela* and *Idanre: po ène*. "Galle's French translation collection, *Cycles sombres*, consists of 30 poems drawn from *Idanre and Other Poems*, *A Shuttle in the Crypt* and *Ogun Abibimāñ*. *Idanre and Other Poems* is a collection of 37 poems by Soyinka, from which Étienne Galle translated six poems, and André Bordeaux one. Galle translated 23 poems from *A Shuttle in the Crypt*, a collection of 34 poems by Soyinka. *Ogun Abibimāñ* is a long single poem of three parts" [9]. Opaluwah [9] in her work, showed that while most of the readership of Etienne Galle and André Bordeaux's translated poetry were the French, Etienne Galle also responded in a private communication that his primary goal in translating Wole Soyinka's poetry into French was out of interest and to bring the work to the French audience. Etienne Galle also noted that while he spent most of his time in the Niger Republic, as a lecturer, he did not work with other poets or any translation theory in the course of translating Wole Soyinka's poetry.

## 3. ChatGPT's Functionalities

This section mainly examines how ChatGPT's function has evolved and how it interprets and translates literary texts. Ogundare and Araya [7], in the comparative Analysis of CHATGPT and the evolution of language models, utilized spontaneous quality (SQ) scores to compare the performance of ChatGPT on many NLP tasks such as machine translation, machine summarization, question answering, and language generation, compared with other mainstream algorithms. They propose that (LLMs) like ChatGPT have made many improvements in understanding and modelling language and bridging the gap between language models and knowledge models. With an increasing amount of available data for systematic training and advancements in the computational theory of machine learning, it is expected that language models will continue to improve and push the boundaries of what is possible in natural language processing, particularly in industrial automation applications.

Partha Pratim Ray [11] comprehensively reviews ChatGPT's background, applications, key challenges, bias, ethics, limitations, and future scope in his work. Some of the issues he outlined that ChatGPT can work towards solving include maintaining context, handling ambiguity, personalization, common sense reasoning, emotional intelligence, ethical considerations, robustness and security, multi-modal interactions, out-of-distribution queries, scalability, and efficiency. He concludes that over the years, LLMs have become larger and more powerful, with impressive natural language understanding and generation capabilities, and they have also been used in a wide range of natural language processing tasks,

from language translation to question-answering systems. However, as these models become more complex, they also raise ethical and societal concerns, such as biases and limitations, which must be carefully addressed; their development requires substantial computational resources, including high-performance GPUs and extensive memory, and the training process often takes weeks or months.

In Servais Martial Akpaca's [2] work on the Syntactic, Semantic, and Pragmatic Evaluation of the Translation of an Ethnographic Text by ChatGPT, the results of his study reveal that contrary to the current propaganda, ChatGPT primarily engages in literal translation. It does not engage in oblique translation. Indeed, errors and mistakes of syntactic, semantic, and pragmatic nature are abundant. Procedures such as transposition, modulation, foreignization, domestication, adaptation, transediting, etc., are almost unknown to it. According to him, ChatGPT is a tool containing many words and can effectively assist translators in their work. However, it is too early to envision a scenario where this technology would replace experienced translators. His evaluation involves identifying translation mistakes and subsequently revising the translation accordingly. After a questionnaire was submitted to ChatGPT, which it answered, he concluded that ChatGPT contains many words, but it is not the number of words a translator knows that makes them a good translator. This syntactic, semantic, and pragmatic evaluation of ChatGPT's translation has yielded disappointing results, leading to the conclusion that ChatGPT is merely a translation aid, even though it is more advanced than other tools such as Systran, Google Translate, Microsoft Translator, etc.

Shijing Si et al. [14] explored ChatGPT's capabilities on ancient Chinese via two tasks: translating ancient Chinese to modern Chinese and recognizing ancient Chinese names. They evaluated the ability of ChatGPT on an ancient Chinese book, *Shi Shuo Xin Yu*, which is largely ignored by previous research. They also studied the performance of ChatGPT on ancient-to-modern translation by varying the input length for each query. Additionally, personal name recognition is rarely explored in ancient Chinese processing, given that ancient Chinese is challenging in natural language processing due to its complex grammatical structures, cultural nuances, and polysemy language. After designing proper prompt instructions for the translation task, it was discovered that the prompt rarely produced errors or superfluous information. Then, they extracted text content that contained only modern Chinese translations of ancient Chinese. They conclude that ChatGPT's capacity for ancient-to-modern Chinese translation and personal name recognition still leaves much to be desired. One possible explanation might be that ChatGPT is mainly pre-trained on English corpora with a few Chinese corpora.

Yu Wei [18], in his work on a comparative study between manual and ChatGPT translations of Literary Texts, uses *Kung I-chi* as an Example in testing ChatGPT's literary translation ability at lexical, syntactic, and textual levels.

While the research results showed that ChatGPT's literary translation has generally progressed, ChatGPT can hardly rival sophisticated human translators at the lexical level.

Faten Khoshafah [3] also evaluated the accuracy of ChatGPT for Arabic-to-English translation. Her study focuses on different types of texts, including general texts from the media, literary, scientific, religious, and legal texts, to compare the accuracy of ChatGPT translation in these different contexts. From the results presented, it was discovered that the translations by ChatGPT and human translation carry the same meaning to some extent but with differences in style and word choice. Some texts translated by ChatGPT are grammatically correct but lack proper Arabic phrasing and word order. Idiomatic expressions are another challenge for ChatGPT. Arabic contains many colloquial terms that are not present in English. ChatGPT may also struggle with Arabic diacritical marks, which indicate the sounds of vowels and pronunciation. False cognates, words that appear or sound similar but have different meanings in both languages, can cause translation problems. Moreover, the rules, sentence structure, and cultural background differ between Arabic and English. She concludes that ChatGPT may not understand the cultural significance of some phrases or terms because it lacks the cultural background and context in which these expressions or terms were used.

Shanshan Wang et al. [17] in Their study examine ChatGPT's capabilities in English-Chinese poetry translation tasks, utilizing targeted prompts and small sample scenarios to ascertain optimal performance. Despite promising outcomes, our analysis reveals persistent issues in the translations generated by ChatGPT that warrant attention. They used all 10 poems in the test set of the Explanation-Assisted Poetry Machine Translation (EAPMT) experiment to conduct data verification experiments on GPT-4 to verify the originality of ChatGPT's responses. They invited six judges to evaluate the experimental results of eight modern English poems translated under eight different prompts. These judges are native Chinese speakers and Chinese professional poets born between the 1960s and 1990s, who are members of the Chinese Writers Association who understand English, with some having engaged in poetry writing or English-Chinese poetry translation for many years. They asked the judges to vote on the output translations of GPT-3.5 and GPT-4 under three human prompts through a questionnaire survey to select the most appropriate prompt. At the end of their investigation, they observed that the decrease in the performance of the models GPT3.5 and GPT 4 was unexpected, as more samples did not promote the models to learn the characteristics of poetry. The superior performance in the zero-shot scenario is likely due to the inherent nature of modern poetry, which is characterized by its freedom and lack of constraints, concluding that the optimum performance of GPT-3.5 on the poetry translation task occurs in the 0-shot scenario using the third prompt (H3). The best translation result is denoted as GPT3.5-Best. The optimum performance of GPT-4 occurs in

the 0-shot scenario using the second prompt (H2), and the best translation result is recorded as GPT4-Best.

However, Peng et al. [10] investigate in their paper how to improve ChatGPT's translation performance at different temperatures to verify its effect and find the optimal temperature for task- and domain-specific prompts. They presume that inconsistent task-oriented prompts will somewhat limit Chat GPT's ability. In response to this, they proposed task-specific prompts (TSP) to further emphasize the task information to bridge the task gap. At the domain information level, they argue that the flexible interaction of ChatGPT to incorporate additional information, like human interactions, may alleviate some classical MT challenges like cross-domain generalizations. They therefore propose Domain-Specific Prompts (DSP) to introduce the domain navigation information to elicit ChatGPT's generalization ability across different domains. They compared Chat GPT with Google translator and for multilingual translation and in-context learning, they evaluated the performance of the models on the Flores-200 4 test sets, which consists of 1012 sentences; to evaluate the effect of cross-domain translation, they adopted the test set of WMT19 Biomedical, News Translation Task and WMT22 E-Commerce task. For evaluation, they used the reference-based metric COMET-20 plus BLEU score and Chrf using Sacrebleu for completeness.

To investigate the influence of diversity, they compared the performance of ChatGPT in different temperature settings, including 0, 0.2, 0.4, 0.6, 0.8, and 1, across three translation directions: English⇒Romanian, English⇒Chinese, and English⇒German. They noted that ChatGPT's sensitivity to temperature varies depending on the language pair: the impact of temperature was relatively small when translating to high-resource languages, such as German, while for complex languages, such as Chinese, it resulted in a significant degradation in performance (−4.3 COEMT points and −3.7 BLEU points for Chinese) when the temperature was changed from 0 to 1. For the TSPs, they noted that when tackling non-English-centric MT language pairs, ChatGPT tends to generate translation hallucinations and that lower temperature can reduce the number of hallucinations (especially in distant languages, e.g., Chinese), and their TSP method can further reduce the number of hallucinations.

They also explored the effects of advanced in-context learning (ICL) strategies by investigating ChatGPT's few-shot ICL and Chain-of-Thought (CoT) abilities on MT tasks. Peng et al [10]. observed that in-context learning with random examples consistently improves performance in lexical metrics (BLEU) and COMET scores compared to the zero-shot approach, and that increasing the number of shots could lead to further improvement. They also randomly selected 20 samples from the test set and adopted the zero-shot Chain-of-Thought (CoT) technique (Kojima et al., 2022) and the 1-shot CoT technique to investigate the influence of CoT on machine translation tasks. Their results showed that there was a significant degradation in the COMET score with the

zero-shot Cot setting, especially in the English⇒Chinese direction.

Zan et al. [19] also investigated that accurate translation-tailored LLMs can be built with language-aware instruction tuning. They designed a two-stage fine-tuning algorithm to enhance the instruction-following ability (especially the translation direction) of LLMs. Using the International Conference on Spoken Language Translation (IWSLT) and Workshop on Machine Translation (WMT) benchmarks on the Large Language Model Meta AI (LLaMA) model spanning 16 zero-shot directions, they tuned LLMs with maximum likelihood estimation loss on the translation dataset to elicit basic translation capabilities in the first stage. In the second stage, they constructed instruction-conflicting samples by randomly replacing the translation directions with incorrect ones within the instruction and then introduced an additional unlikelihood loss to learn from these samples.

They utilised multilingual translation examples for the first-stage pre-tuning to unlock the translation capabilities of the LLM; the model is trained to predict output based on provided instructions. They then enhanced the zero-shot translation ability through a dual optimization approach, incorporating Maximum Likelihood Estimation (MLE) loss on multilingual translation samples and unlikelihood loss on instruction-conflicting samples. To mitigate the off-target problem with unlikelihood training, they created negative candidate samples by replacing the instruction with a different one while keeping the input and output unchanged. Based on the instruction-conflicting samples, they generalised the unlikelihood training to zero-shot translation of translation-tailored LLMs. They fed instruction samples into the model trained after stage 1, optimizing the unlikelihood loss.

The results obtained showed that their model outperformed the considered baselines across 16 translation directions. Furthermore, the effect of the unlikelihood training steps was evident, as their model produced fewer incorrect language translations and higher-quality translations. When they further investigated the influence of a larger model size, they observed that the model even performed better.

From the few works reviewed about ChatGPT's functions as an LLM, it can be concluded that LLMs are still evolving in their understanding of more human languages and cultures, as well as the more varied or minority language data and instructions used to train and retrain them. The better its output regarding language translation is directly related to the more diverse or minority language data and instructions used to train and retrain it.

## 4. Methodology

This investigation aims to analyze the interpretation of CSIs by large language models (LLMs) using ChatGPT 3.5 and ChatGPT Pro in the English-French translations of three African anthologies: *Early Poems*, *Mandela's Earth and Other Poems*, and *Idanre and Other Poems*. This analysis



used structured prompts in tabular format, with a comparative assessment against human translations and a custom translation model.

#### 4.1. Data Collection

The CSIs generated by general machine translation systems (DeepL and Amazon Translate) and a custom translation engine were identified in an ongoing larger project (*Man Versus Machine: Translating Culture-Specific Items in Wole Soyinka's Poetry*).

To identify the CSIs in the reference HT, GMT, and CTE, Aixelá's (1996) model for identifying and categorizing CSIs and identifying and/or inputting strategies is used. According to the definition, a CSI is

"A result of a conflict arising from any linguistically represented reference in a source text which, when transferred to a target language, poses a translation problem due to the nonexistence or to the different value whether determined by ideology, usage, frequency, etc.) of the given item in the target language culture" [1].

And also that

"There is a common tendency to identify CSIs with those items, especially linked to the most arbitrary area of each linguistic system - its local institutions, streets, historical figures, place names, personal names, periodicals, works of art, etc., which will typically present a translation problem in other languages" [1].

From an initial trial, it was observed that ChatGPT Pro could only translate six stanzas at once out of the 19 stanzas (not counting the refrains) of the second part of the *Ogun Abibimañ* poem. In the order of *Early Poems*, *Ogun Abibimañ*, *Mandela's Earth and other Poems*, and *Idanre and Other Poems* each six-stanzas embedding the CSI (three stanzas above and three stanzas below the CSI including the stanza embedding the CSI provided that the stanzas count does not jump into a new poem) was transferred into the chat interface of ChatGPT 3.5 and ChatGPT Pro; followed by the provision

of structured prompts for translation into French, the results of the translations of the CSIs by both ChatGPT 3.5 and ChatGPT Pro were documented.

Three structured prompts were submitted to ChatGPT 3.5 and ChatGPT Pro. The selection of ChatGPT 3.5 was influenced by its previously noted favorable results in the reviewed literature, whereas ChatGPT Pro was chosen under the assumption that it could process entire anthologies due to its unlimited access to features. The three prompts are a general prompt: *Translate this poem into French*, a poetic-structure-oriented prompt: *Translate this poem into French, keeping the stanzas*, and a culture-oriented prompt: *Translate this poem into French, paying attention to the cultural words*.

#### 4.2. Data Analysis

Six tables were developed for both versions of LLM to facilitate this investigation. The initial table outlines the results of the CSIs generated after three distinct levels of prompts directed at ChatGPT 3.5 and ChatGPT Pro. The subsequent table categorizes the output of CSIs according to Aixelá's framework of "Proper nouns and Common expressions" [1], which will assist in evaluating the frequency of repetition within each category as produced by Human Translation (HT), the custom-built Microsoft Azure translation engine (CTE), and LLM (ChatGPT 3.5 and ChatGPT Pro). The third table summarizes the outputs of CSIs from the human translators, CTE, and LLM. Finally, a concluding table provides an overview of the strategies employed by ChatGPT 3.5 and ChatGPT pro, in translating CSIs, particularly following Aixelá's guidelines, after the culture-oriented prompt.

Prompt Definition

1<sup>st</sup> Prompt = General prompt (GP): *Translate this poem into French*

2<sup>nd</sup> Prompt = Structure-oriented prompt (SoP): *Translate this poem to French, keeping the stanzas*

3<sup>rd</sup> Prompt = Culture-oriented prompt (CoP): *Translate this poem to French, paying attention to the cultural words*

**Table 1.** Overview of CSIs output by ChatGPT 3.5 After Prompts.

N0.	CSI	Prompts	Output
1	Bean-cake	GP	ruche de galettes de haricots galettes gravées
		SoP	ruche de gâteau de haricot gâteau de haricot marqués
		CoP	ruche de galettes de haricots galettes marquées
2	Ogun	GP	Ogun
		SoP	Ogun
		CoP	Ogun

N0.	CSI	Prompts	Output
3	Soshangani	GP	Soshangani
		SoP	Soshangani
		CoP	Soshangani
4	Lobengula	GP	Lobengula
		SoP	Lobengula
		CoP	Lobengula
5	Mfekane	GP	Mfekane
		SoP	Mfekane
		CoP	Mfekane
6	Shaka	GP	Shaka
		SoP	Shaka
		CoP	Shaka
7	Amazulu	GP	Amazulu
		SoP	Amazulu
		CoP	Amazulu
8	Mickey Mouse	GP	Un souris
		SoP	Une souris
		CoP	Une souris
9	Esu (1988)	GP	Esu
		SoP	Esu
		CoP	Esu
	Esu (1967)	GP	Esu
		SoP	Esu
		CoP	Esu
10	Egungun	GP	Egungun
		SoP	Egungun
		CoP	Egungun
11	Jigida	GP	jigidas
		SoP	jigida
		CoP	jigidas
12	Scottsborough Boys	GP	garçons de Scottsborough
		SoP	garçons de Scottsborough
		CoP	Scottsboro Boys
13	Asantehene	GP	Asantehene
		SoP	Asantehene
		CoP	Asantehene
14	Sango (1967)	GP	Sango
		SoP	Sango
		CoP	Sango

N0.	CSI	Prompts	Output
15	Atunda	GP	Atunda
		SoP	Atunda
		CoP	Atunda
16	Orisanla	GP	Orisanla
		SoP	Orisanla
		CoP	Orisanla
17	Orunmila	GP	Orunmila
		SoP	Orunmila
		CoP	Orunmila
18	Ire	GP	Colère
		SoP	Colère
		CoP	Colère
19	Iron One	GP	Homme de fer, Le Fer, Fer-Unique, Ô Fer, l'Iron One
		SoP	L'Iron One, l'
		CoP	Iron, Fer-Roi, Ô Homme de Fer, À l'Être de Fer Maître du Fer; l'Homme de Fer; l'Homme de Fer; O Fer, l'Unique; le Fer

*Table 2. Overview of CSIs translated by ChatGPT Pro After Prompts.*

No.	CSI	Prompts	Output
1	Bean-cake (Soyinka, 1998, p. 133)	GP	galettes de haricot galette
		SoP	galettes de haricots, galettes
		CoP	gâteau de fève le gâteau de fève
2	Ogun (Soyinka, 1976, pp. 9, 11, 19, 22) (Soyinka, 1967, pp. 61, 63, 64, 65, 67-72, 74, 75, 80, 83, 85)	GP	Ogun Ogun
		SoP	Ogun Ogun
		CoP	Ogun Ogun
3	Soshangani (Soyinka, 1976, p. 12)	GP	Soshangani
		SoP	Soshangani
		CoP	Soshangani
4	Lobengula (Soyinka, 1976, p. 12)	GP	Lobengula
		SoP	Lobengula
		CoP	Lobengula
5	Mfekane (Soyinka, 1976, p. 19)	GP	Mfekane

No.	CSI	Prompts	Output
6	Shaka (Soyinka, 1976, pp. 9, 10, 11-17)	SoP	Mfekane
		CoP	Mfekane
		GP	Shaka
		SoP	Shaka
7	amaZulu (Soyinka, 1976, pp. 12, 14, 16, 17)	CoP	Shaka
		GP	amaZulu
		SoP	amaZulu
		CoP	amaZulu
8	Mickey Mouse (Soyinka, 1988, p. 41)	GP	Mickey
		SoP	Mickey Mouse
		CoP	souris en peluche
		GP	Lèvres à Louiseville
9	Louisville Lips	SoP	Lèvres à Louiseville
		CoP	Louiseville Lips
		GP	Esu
		SoP	Esu
10	Esu (Soyinka, 1967, pp. 70 & 78) <i>Esu</i> (Soyinka, 1988, p. 48)	CoP	Esu
		GP	<i>Esu</i>
		SoP	<i>Esu</i>
		CoP	<i>Esu</i>
11	<i>egungun</i> (Soyinka, 1988, p. 69)	GP	<i>egungun</i>
		SoP	<i>egungun</i>
		CoP	<i>egungun</i>
		GP	<i>jigida</i>
12	<i>jigida</i> (Soyinka, 1988, p. 45)	SoP	<i>jigida</i>
		CoP	<i>jigida</i>
		GP	Scottsborough Boys
		SoP	Scottsborough Boys
13	Scottsborough Boys (Soyinka, 1988, p. 40)	CoP	Scottsboro Boys
		GP	<i>Asantehene</i>
		SoP	<i>Asantehene</i>
		CoP	<i>Asantehene</i>
14	<i>Asantehene</i> (Soyinka, 1988. P. 17)	GP	Sango
		SoP	Sango
		CoP	Sango
		GP	Atunda
15	Sango (Soyinka, 1967, pp. 61, 70)	SoP	Atunda
		CoP	Atunda
		GP	Orisanla
		SoP	Orisanla
16	Atunda	GP	Orisanla
		SoP	Orisanla
		CoP	Orisanla
		GP	Orisanla
17	Orisanla	GP	Orisanla
		SoP	Orisanla
		CoP	Orisanla
		GP	Orisanla



No.	CSI	Prompts	Output
18	Orunmila	SoP	Orisanla
		CoP	Orisanla
		GP	Orunmila
		SoP	Orunmila
		CoP	Orunmila
19	Ire	GP	Colère
		SoP	Colère
		CoP	Colère
20	Iron One	GP	Homme de fer, Le Fer, Fer-Unique, Ô Fer, l'Iron One
		SoP	L'Iron One, l'Iron, Fer-Roi, Ô Homme de Fer, À l'Être de Fer
		CoP	Maître du Fer; l'Homme de Fer; l'Homme de Fer; O Fer, l'Unique; le Fer

It is observed that while ChatGPT 3.5 records 19 CSIs, the Pro version records 20. This is because of the CSI *Louseville Lip*, which is not attributed as a CSI for the 3.5 version because the translation output by ChatGPT 3.5 is the expected output. But it is a CSI for the Pro version because it was not left untranslated. As noted, “aesthetically, this should have been left untranslated as it is the name Muhammad Ali is known by in certain circles.” [9].

## 5. Unchanged (Repeated) CSIs

With all three types of prompts given to LLM, a total of 13 CSIs out of a combined total of 20 CSIs are translated unchanged by LLM: *Ogun, Esu, Mfekane, Shaka, Soshangani, Lobengula, amaZulu, Sango, Atunda, Orisa-nla, Orunmila, Egungun*, and *Asantehene* were rendered by both ChatGPT 3.5 and ChatGPT Pro, unchanged. In other words, the above CSIs were rendered through the strategy of Repetition.

To categorize the above CSIs, Aixelá's model distinguishes “Two basic categories from the translator's point of view: proper nouns and common expressions (for want of a better term, the world of objects, institutions, habits, and opinions restricted to each culture and that cannot be included in the field of proper names)” [1].

Thus, the 13 CSIs that have retained their forms as in the source texts are grouped below, giving 11 proper nouns and two common expressions.

**Table 3.** Overview of Unchanged CSIs After the Three Levels of Prompts.

No.	Proper Nouns	Common Expressions
1	Ogun	Mfekane
2	Soshangani	Egungun
3	Lobengula	
4	amaZulu	
5	Shaka	
6	Esu	
7	Atunda	
8	Orisa-nla	
9	Orunmila	
10	Asantehene	
11	Sango	

## 6. Human Translation of CSIs

The human translations (reference translations) of the three anthologies were done separately by two French translation scholars: Etienne Galle and André Bordeaux. While Etienne Galle translated poems from *Early Poems, Mandel's Earth and Other Poems, A Shuttle in the Crypt, Ogun Abibima ñ* and *Idanre and Other Poems* into French, André Bordeaux translated one long seven-part poem, *Idanre*, from the col-

lection of *Idanre and Other Poems*, making a total of 47 Soyinka poems translated into French by the two translators.

Adeyola Opaluwah [9] appraised these European translators in her article, detailing their expertise by evaluating their translations of African poetry, their publications about the African Nobel laureate Soyinka, and the awards and recommendations they received. As Etienne Galle stated in a private communication, his aim in translating Soyinka's poetry was to bring the translation to the French readership.

Based on prior investigation, following Aixelá's model of translating CSIs, it was noted that, despite the intended French readership, European translators encountered constraints in their choices of how best to render the 18 CSIs.

## 7. Custom Translation Engine (CTE) of CSIs

In a previous work, a custom translation engine was developed to investigate the output of CSIs. The CTE was built on Mohamed Elghazali's model, which "involved creating a translator resource key" [6] on Microsoft Azure and creating a workspace and a project.

Once the project is successfully created, document uploading, which includes the three anthologies of African poetry, begins. The Microsoft Azure custom model requires 10,000 training, tuning, and testing sentences. Training, tuning, and testing were realised using 14,065 sentences of bilingual African poetry. During document upload, each poetry file is labelled to notify the model of its function as either training, tuning or testing data. For example, Wole Soyinka's three anthologies were labelled "testing data" when uploading them to the custom model.

Training starts after the minimum requirement of 10,000 parallel sentences of African poetry is uploaded. At the end of the training, the CTE outputs the model's training details, including the BLEU score. Elghazali [6] suggests that the size of the BLEU score could be improved by enhancing the alignment of the parallel sentences; the model aligns the sentences during the upload process.

The CTE was fully trained on 20 document sets comprising 23,344 sentences of strictly African poetry, spanning 14 hours and 7 minutes. The details are as follows: Training = 16,234 sentences, Testing = 2,476 sentences, Tuning = 854 sentences, and Dictionary = 11. The CTE automatically calculated the BLEU score to 35.29.

**Table 4.** Overview of CSIs output by Man (HT), CTE and LLM.

	Source word	Human translation	CTE output of CSIs	LLM's (Gpt 3.5, Gpt Pro) translation after the culture-oriented prompt
1	My impi (Soyinka, 1976, p. 11)	Mes impi (Soyinka 2019:160)	Mes impi	Not a CSI for LLM
2	Bean-cake (Soyinka, 1998, p. 133) x 2	Gâteau de fèves (Soyinka 2019, p. 77)	gâteau de fèves	galettes de haricots, gâteau de fève
		Gâteau de fèves (Soyinka 2019, p. 77)	gâteau de fèves	galettes, le gâteau de fève
3	Ogun (Soyinka, 1976, pp. 9, 11, 19, 22)	Ogun (Soyinka 2019, pp. 151, 154, 156, 158, 160, 170, 173-178)	Ogun	Ogun
	(Soyinka, 1967, pp. 61, 63, 64, 65, 67-72, 74, 75, 80, 83, 85)	Ogoun Soyinka, 1982, p. 7, 8, 10, 12, 15, 17, 19-21, 23, 26, 27, 28, 32	Ogoun	Ogun
4	Soshangani (Soyinka, 1976, p. 12)	Soshangani (Soyinka 2019, p. 162)	Soshangani	Soshangani
5	Lobengula (Soyinka, 1976, p. 12)	Lobengula (Soyinka 2019, p. 162)	Lobengula	Lobengula
6	Mfekane (Soyinka, 1976, p. 19)	Mfëkané (Soyinka 2019, p. 169)	Mfekane	Mfekane
7	Shaka (Soyinka, 1976, pp. 9, 10, 11-17)	Not a CSI in HT	Not a CSI for CTE	Shaka
8	amaZulu (Soyinka, 1976, pp. 12, 14, 16, 17)	Not a CSI in HT	amaZulu	amaZulu
9	Mickey Mouse (Soyinka, 1988, p. 41)	Mickey Mouse (Soyinka,	Mickey Mouse	souris en peluche

	Source word	Human translation	CTE output of CSIs	LLM's (Gpt 3.5, Gpt Pro) translation after the culture-oriented prompt
		1989, p. 63)		
10	Louisville Lips (Soyinka, 1988, p. 47)	lèvres... à Louisville (Soyinka, 1989, p. 71)	Lèvres... à Louisville	Not a CSI for LLM
	Esu (Soyinka, 1967, pp. 70 & 78)	Not a CSI in HT	Eshou	Esu
11	Esu (Soyinka, 1988, p. 48)	Esu (Soyinka, 1982, pp. 27 & 34)	Esu	Esu
12	Egungun (Soyinka, 1988, p. 69)	egungun (Soyinka, 1989, p. 103)	egungun	egungun
13	jigida (Soyinka, 1988, p. 45)	jiguida (Soyinka, 1989, p. 68)	jiguida	Jigidas, jigida
14	Scottsborough Boys (Soyinka, 1988, p. 40)	Not a CSI in HT	Not a CSI for CTE	Scottsboro Boys
15	Asantehene (Soyinka, 1988, P. 17)	Not a CSI in HT	Not a CSI in HT	Asantehene
	Sango (Soyinka, 1967, p. 61)	Chango (Soyinka, 1982, p. 17)	Chango,	Sango
16	Sango (Soyinka, 1967, p. 70)	Shango (Soyinka, 1982, p. 27)	Shango	Sango
17	Atunda (Soyinka, 1967, pp. 81, 83)	Atunda (Soyinka, 1982, p. 37 & 38)	Atunda	Atunda
18	Orisa-nla (Soyinka, 1967, p. 70)	Orisanla (Soyinka, 1982, p. 27)	Orisa-Nla	Orisa-nla
19	Orunmila (Soyinka, 1967, p. 70, 83)	Orunmila (Soyinka, 1982, p. 27 & 38)	Orunmila	Orunmila
20	Ire (Soyinka, 1967, p. 71)	Ire (Soyinka, 1982, p. 27)	Ire	Colère
	Iron One (Soyinka, 1967, p. 61)	Fer en personne (Soyinka, 1982, p. 17)	Fer en personne	le Fer
21	Iron One (Soyinka, 1967, pp. 68, 70, 74)	dieu du fer (Soyinka, 1982, pp. 25, 27, 32)	Dieu du fer, Ogoun	Le Fer, l'Iron One, Ô Fer,
	Iron One (Soyinka, 1967, p. 78)	Ogoun (Soyinka, 1982, p. 34)	Dieu de fer	Fer One

## 8. LLM's Strategies for Unrepeated CSIs

This section examines how LLM translated CSIs that were not literary repetitions, utilising Aixelá's model for attributing

strategies to translating CSIs. Tables 1 and 2 show that seven CSIs were mostly translated without repetition. The CSIs repeated in translation are shown in Table 5 below, while the strategies are included in Table 6 for a clearer understanding.

**Table 5.** Overview of CSIs that were not repeated in translation by LLM.

No.	Source CSI	Prompt	GPT 3.5's translation	GPT Pro's translation
1	Bean-cake	GP	galettes de haricots galettes	galettes de haricot galette
		SoP	gâteau de haricot	galettes de haricots,

No.	Source CSI	Prompt	GPT 3.5's translation	GPT Pro's translation
2	Mickey Mouse		gâteau de haricot	galettes
		CoP	galettes de haricots	gâteau de fève
			galettes	le gâteau de fève
		GP	Un souris	Not a CSI here
3	Louisville Lips	SoP	Une souris	Mickey Mouse
		CoP	Une souris	souris en peluche
		GP	Not a CSI here	lèvres de Louisville
		SoP	Not a CSI here	lèvres de Louisville
4	Jigida	CoP	Not a CSI here	Not a CSI here
		GP	jigidas	Not a CSI here
		SoP	Not a CSI here	Not a CSI here
		CoP	jigidas	Not a CSI here
5	Scottsborough Boys	GP	Not a CSI here	Scottsborough Boys
		SoP	Not a CSI here	Scottsborough Boys
		CoP	Scottsboro Boys	Scottsboro Boys
		GP	Colère	Colère
6	Ire	SoP	Colère	Colère
		CoP	Colère	Colère
		GP	Homme de fer, Le Fer, Fer-Unique, Ô Fer, l'Iron One	Homme de fer, Le Fer, Fer-Unique, Ô Fer, l'Iron One
		SoP	L'Iron One, l' Iron, Fer-Roi, Ô Homme de Fer, À l'Être de Fer	L'Iron One, l' Iron, Fer-Roi, Ô Homme de Fer, À l'Être de Fer
7	Iron One	CoP	Maître du Fer; l'Homme de Fer; l'Homme de Fer; O Fer, l'Unique; le Fer	Maître du Fer; l'Homme de Fer; l'Homme de Fer; O Fer, l'Unique; le Fer

## 9. Aixelá's Model of Strategies for CSIs

Aixelá discusses methods for translating culture-specific elements between languages, noting that the selection of these methods reflects, among other considerations, the tolerance level of the receiving culture and its robustness. He defines culture-specific items as those closely associated with the unique aspects of each linguistic system, such as local institutions, streets, historical figures, place names, personal names, periodicals, and works of art, which typically present challenges in translation to other languages. Furthermore, the frequent emergence of textual elements that do not appear particularly arbitrary, whose translation difficulties can only be understood through intercultural differences, compels translation students to broaden their perspectives. He explains that conflicts may arise from any reference represented lin-

guistically in a source text, creating a translation challenge when rendered into a target language due to the absence or differing significance, shaped by ideology, usage, frequency, and other factors, of the item within the target culture [1].

### 9.1. Summary of Aixelá's Strategies Used by LLM

Aixelá divides the proposed strategies into conservation and substitution strategies. The conservation strategies are:

- 1) Repetition: Rendered precisely or almost precisely as the source word. The translators retain as much of the original reference as possible.
- 2) Orthographic adaptation: This strategy includes procedures like transcription and transliteration.
- 3) These are mainly used when the original reference is expressed in a different alphabet from the target read-

ers.

- 4) Linguistic (non-cultural) translation: Linguistic transparency of CSI; supports pre-established translations within the intertextual corpus of the target language, or making use of the linguistic transparency of the CSI, the translator chooses in many cases a denotatively very close reference to the original, but increases its comprehensibility by offering a target language version which can still be recognized as belonging to the cultural system of the source text.
- 5) Extratextual gloss: Provides an explanation of the meaning or implications of the CSI as a footnote or endnote.
- 6) Intratextual gloss: The translators feel they can or should include their gloss as an indistinct part of the text. So, the explanation is within the text and not separated as in extratextual gloss.
- 7) The substitution strategies are:
- 8) Synonymy: This strategy is based on stylistic grounds linked with recurrence.
- 9) Limited universalization: CSI is too obscure for readers, or there is another, more usual possibility, and they decide to replace it; for credibility, they seek another reference belonging to the source language culture but closer to another of their readers' CSI.

10) Absolute Universalization: The basic situation is identical to Limited universalization, but the translators do not find a better-known CSI or prefer to delete any foreign connotations and choose a neutral reference.

11) Naturalization: The translator brings the CSI into the intertextual corpus, which is felt to be specific by the target language culture. This strategy is infrequently used in the literature.

12) Deletion: The translators consider the CSI unacceptable on ideological or stylistic grounds, believe it is not relevant enough for the comprehension effort required of their readers, or find it too obscure, and are either not allowed or do not want to use procedures such as glosses. Therefore, they decide to omit it in the target text.

13) Autonomous creation: This is a rarely used strategy in which the translators (or their initiators) decide it would be interesting for their readers to include a nonexistent cultural reference in the source text.

Of the 11 strategies outlined above, LLM employed Repetition, Linguistic (non-cultural) translation, Limited universalization, Absolute universalization, and Autonomous Creation, which comprise two conservation and three substitution strategies. The table below shows this. The use of "NA" (not applicable) indicates that no applicable strategy exists.

**Table 6.** Translation Strategies of LLMs for unrepeated CSIs after prompts.

No.	Source word	GPT 3.5			GPT Pro		
		Prompt	Translation output	Translation Strategy	Prompt	Translation output	Translation Strategy
1	Bean-cake	GP	galettes de haricots	Linguistic (non-cultural) translation	GP	galettes de haricot	Linguistic (non-cultural) translation
			galettes	Linguistic (non-cultural) translation		galettes	Linguistic (non-cultural) translation
			gâteau de haricot	Linguistic (non-cultural) translation		galettes de haricots,	Linguistic (non-cultural) translation
		SoP	gâteau de haricot	Linguistic (non-cultural) translation	SoP	galettes	Linguistic (non-cultural) translation
			galettes de haricots	Linguistic (non-cultural) translation		gâteau de fève	Linguistic (non-cultural) translation
		CoP	galettes	Linguistic (non-cultural) translation	CoP	le gâteau de fève	Linguistic (non-cultural) translation
2	Mickey Mouse	GP	Un souris	Limited Universalization	GP	Not a CSI here	NA

No.	Source word	GPT 3.5			GPT Pro		
		Prompt	Translation output	Translation Strategy	Prompt	Translation output	Translation Strategy
3	Louisville Lips	SoP	Une souris	Limited Universalization	SoP	Mickey Mouse	Repetition
		CoP	Une souris	Limited Universalization	CoP	souris en peluche	Linguistic (non-cultural) translation
		GP	Not a CSI here	NA	GP	Lèvres à Louisville	Linguistic (non-cultural) translation
		SoP	Not a CSI here	NA	SoP	Lèvres à Louisville	Linguistic (non-cultural) translation
4	Jigida	CoP	Not a CSI here	NA	Cop	Not a CSI here	NA
		GP	Jigidas	Repetition	GP	Not a CSI here	NA
		SoP	Not a CSI here	NA	SoP	Not a CSI here	NA
		CoP	Jigidas	Repetition	CoP	Not a CSI here	NA
5	Scottsborough Boys	GP	Not a CSI here	NA	GP	Scottsborough Boys	Repetition
		SoP	Not a CSI here	NA	SoP	Scottsborough Boys	Repetition
6	Ire	CoP	Scottsboro Boys	Limited Universalization	CoP	Scottsboro Boys	Limited Universalization
		GP	Colère	Autonomous Creation	GP	Colère	Autonomous Creation
		SoP	Colère	Autonomous Creation	SoP	Colère	Autonomous Creation
		CoP	Colère	Autonomous Creation	CoP	Colère	Autonomous Creation
7	Iron One	GP	Homme de fer,	Linguistic (non-cultural) translation,	GP	Homme de fer,	Linguistic (non-cultural) translation,
			Le Fer,	absolute Universalization,		Le Fer,	absolute Universalization,
			Fer-Unique,	Linguistic (non-cultural) translation,		Fer-Unique,	Linguistic (non-cultural) translation,
			Ô Fer,	Absolute Universalization,		Ô Fer,	Absolute Universalization,
			l'Iron One	Repetition		l'Iron One	Repetition
			L'Iron One,	Repetition,		L'Iron One,	Repetition
			l'Iron,	Limited Universalization		L'Iron,	Limited Universalization,
		SoP	Fer-Roi,	Linguistic (non-cultural) tran-	SoP	Fer-Roi,	Linguistic (non-cultural) transla-



GPT 3.5					GPT Pro		
No.	Source word	Prompt	Translation output	Translation Strategy	Prompt	Translation output	Translation Strategy
				slation,			tion,
			Ô Homme de Fer,	Linguistic (non-cultural) translation,		Ô Homme de Fer,	Linguistic (non-cultural) translation,
			l'Être de Fer	Linguistic (non-cultural) translation		l'Être de Fer	Linguistic (non-cultural) translation
			Maître du Fer;	Absolute Universalization,		Maître du Fer;	Absolute Universalization,
			l'Homme de Fer;	Linguistic (non-cultural) translation		l'Homme de Fer;	Linguistic (non-cultural) translation,
		CoP	l'Homme de Fer;	Linguistic (non-cultural) translation	CoP	l'Homme de Fer;	Linguistic (non-cultural) translation
			O Fer, l'Unique;	Absolute Universalization,		O Fer, l'Unique;	Absolute Universalization
			le Fer	Absolute Universalization		le Fer	Absolute Universalization,

## 9.2. Fluency and Adequacy Between LLMs (GPT 3.5 and GPT Pro)

Since the investigation focuses on how CSIS has been translated into French by LLM, comparing it with the reference HT might provide an idea of the quality of CSIs rendered by the LLM versions. According to Snover et al. [15], fluency measures whether a translation is fluent, regardless of the correct meaning, while adequacy measures whether the translation conveys the correct meaning, even if it is not fully fluent. The BLEURT metric will, however, be used for an objective and concise evaluation of the versions.

### Analysis

Table 3 indicates that the LLMs produced 13 CSIs in repetition, translating them as they appear in the source text. Among these, 11 are classified as Proper Nouns, while two are identified as Common Expressions. In a prior study, CTE repeated 12 CSIs in their original form from the source anthology, comprising 11 Proper Nouns and one Common Expression. In contrast, the reference (human) translation preserved 14 CSIs unchanged, with 10 categorized as Proper Nouns and four as Common Expressions.

An examination of how the language model translated certain CSIs that were neither repeated nor left untranslated reveals inconsistencies in its approach to translating CSIs from the three anthologies of African poetry from English to

French following three distinct levels of prompts. For instance, after the initial general prompt, ChatGPT translated "Bean-cake" as "galettes de haricot" and "galette" in the subsequent translation. Following the structure-oriented prompt, the same term was rendered as "galettes de haricots" and "galettes" in the second instance. After the culture-oriented prompt, ChatGPT provided "gâteau de fève" in the first instance and "le gâteau de fève" in the second.

The term "Ire" was translated as "Colère" following three prompts. Ire refers to "a town in Ekiti" [13], yet ChatGPT has transformed this place noun into an abstract noun. In the case of "Iron One," five different translations were produced, including "L'homme de Fer," "Le Fer," "L'homme de Fer," "Ô Fer," and "L'Iron One" after the initial prompt. The second prompt yielded "L'homme de Fer," "Le Fer," "L'homme de Fer," "Ô Fer," and "l'Acier." Following the culture-specific prompt, the translations for "Iron One" were "Le Fer," "Le Fer," "l'Iron One," "O Fer," and "Fer One." The most frequently occurring translations are "L'homme de Fer" and "Le Fer." Notably, even after the culture-oriented prompt, ChatGPT's translation of "Iron One" does not align closely with the reference human translations such as "Fer en personne" [16], "dieu du fer" [16], and "Ogoun" [16]. Two other cases to note are LLM's translation of "Mickey Mouse" as "souris en peluche" and "Scottsborough Boys" as "Scottsboro Boys" (without the ough) after the culture-oriented prompt.

As illustrated in Table 6, the two versions of LLM utilised

five strategies for translating CSIs, as per Aixela's framework: encompass Repetition, Linguistic (non) cultural translation, Autonomous creation, Absolute Universalization, and Limited Universalization. In a previous study, the strategies associated with CTE included Repetition, Linguistic (non) cultural translation, Absolute Universalization, Limited Universalization, and Synonymy while the human translators adopted four strategies of Repetition, Linguistic (non) cultural translation, Limited Universalization, and Synonymy.

An attempt to evaluate the two versions of LLMs manually was also made. In Table 3, ChatGPT 3.5 translated *Bean-cake* as *galettes de haricots* and *gâteau de fève* while GPT Pro rendered it as *galettes* and *gâteau de fève*; Meanwhile, the reference translation is *gâteau de fèves*. While GPT 3.5's outputs are judged correct, the missing letter *s* in *gâteau de fève* does not make it fluent; the output of *galettes* by GPT Pro without the addition of *haricots* or *fèves* after it, makes the first translation incomplete; the second translation *gâteau de fève* also lacks the letter *s* in *fèves*. In this example, GPT3.5's translation will be judged more fluent and adequate than GPT Pro.

*Mickey Mouse* was translated as "*souris*" by GPT 3.5 and as "*souris en peluche*" by GPT Pro after the culture-oriented prompt. Although both LLM versions did not output the French cultural version, "Mickey" [5], GPT Pro's output will generally be more acceptable.

A third example of manually evaluating the fluency and adequacy of LLM after the CoP is in the translation of *Louisville Lips*. While ChatGPT 3.5 translates it correctly as *Louisville Lips*, GPT Pro translates it as *Lèvres à Louisville*. GPT Pro's translation is inadequate here because.

*Louisville Lips* is the name by which Mohammed Ali, the famous boxer, is known in certain circles [12]. GPT Pro's translation is inadequate here. This could have been left untranslated, as ChatGPT 3.5 did.

## 10. Conclusion

In this article, three distinct prompts, a general prompt (GP), a structure-oriented prompt (Sop), and a culture-oriented prompt (CoP), were given to two versions of LLM to translate three anthologies of African poetry to investigate how LLM would render the culture-specific items (CSIs) in the anthologies.

Table 1 shows that ChatGPT 3.5 rendered 19 CSIs after the prompts, while ChatGPT Pro rendered 20 CSIs. According to Aixela's model for the definition, identification, and classification of CSIs, the more CSIs produced by the translator, the more translation issues the translator (in this case, LLM) encountered. From the results in the initial tables, ChatGPT Pro is expected to encounter more issues than ChatGPT 3.5 in the translation of CSIs.

The BLEURT score, however, showed that the Pro version scored better than ChatGPT 3.5 after 29 random prediction sentences bearing the CSIs were used alongside the reference sentence on the GP.

The analysis demonstrates that CoP applied to ChatGPT

Pro, an LLM, did not result in any meaningful differences in the output of culture-specific items in translating African poetry from English to French. In translating "Iron One," the version suggested "L'homme de Fer" as a translation option after utilizing general, structural, and culture-oriented prompts. Still, the CoP did not significantly improve the translation quality. Moreover, the translation of "Scottsboro Boys" yielded a less acceptable result, being rendered as "Scottsboro Boys" without the necessary "ugh" in the word. In the case of "Mickey Mouse," the CoP led to the translation "souris en peluche". While the anthologies used for the investigation are African poetry, there were CSIs in some poems about personalities in the West, like *Muhammad Ali at The Ringside, 1985*, and *The Most Expensive Anchorman in U.S.A.*, which LLM could not capture.

While LLMs present challenges such as inherent biases, ethical concerns, and interpretability issues due to their size and complexity limiting their application in some areas and highlighting the need for responsible development and deployment, LLMs continue to evolve, becoming more powerful and capable, with an expanding role in enhancing human-computer interaction; they have potentials for great improvement as investigated by Peng et al. [10] and Zan et al. [19]. This development trajectory has led to the creation of increasingly sophisticated models, enabling advanced NLP tasks such as language translation and question-answering systems. However, ethical and societal concerns regarding their use require careful consideration to ensure fairness and mitigate biases and as research progresses, LLMs are expected to become even more integral to various industries and aspects of daily life (Ray, 2024).

A comparison of the CSI output from a CTE for translating African poetry from English to French reveals that the CTE rendered superior CSIs compared to LLMs, even when the CoP was given to the LLMs. This conclusion is consistent with the findings of Kuzman et al. [4], who argue that models customised for specific authors (the CTE was customised to Wole Soyinka's poetry) generally outperform those developed using a more extensive and varied literary corpus. Peng et al. [10] also noted that LLMs were built as general task solvers and may not respond ideally to specific tasks. Peng et al. [10] also proved that LLMs have the potential for improved translation output when trained with in-context domain data.

Peng et al. [10] noted that most non-English-centric MT language pairs tend to generate hallucinations. The French-centric language pair of this investigation could explain the translation of "Ire", a town in Ekiti State, Nigeria, as "colère" in French. Zan et al. [19] also demonstrated the possibility that a translation-tailored LLM can be built with Language Aware Instruction Tuning; this can further mitigate the effect of generating hallucinations, and their unlikelihood training steps were critical to ensuring their model produced fewer incorrect language translations and higher-quality translations.

This paper's investigation is a small part of a broader research

project, utilising the same corpus from the three anthologies employed in the larger project: *Early Poems*, *Mandela's Earth and Other Poems*, and *Idanre and Other Poems*. Other anthologies from world regions outside of Africa were not investigated. Additional investigation is required to understand how LLMs interpret translation prompts across domains like drama and social media, from different regions into other languages like Arabic, French, Chinese, etc.

## Abbreviations

CoP	Culture-Oriented Prompt
CSI	Culture-Specific Items
CTE	Custom Translation Engine
GMT	General Machine Translation
GP	General Prompt
HT	Human Translation
LLMs	Large Language Models
NA	Not Applicable

## Author Contributions

Adeyola Opaluwah is the sole author. The author read and approved the final manuscript.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] Aixelá, J. F. (1996). Culture-Specific Items in Translation. In R. Alvarez and M. Vidal (Eds.), *Translation, Power, Subversion* (52-78). Multilingual Matters.
- [2] Akpaca, S. M. (2024). A Syntactic, Semantic, and Pragmatic Evaluation of the Translation of an Ethnographic Text by ChatGPT. *London Journal of Research in Humanities & Social Science*, 24(12). Retrieved from [https://journalspress.com/LJRHSS\\_Volume24/A-Syntactic-Semantic-and-Pragmatic-Evaluation-of-the-Translation-of-an-Ethnographic-Text-by-ChatGPT.pdf](https://journalspress.com/LJRHSS_Volume24/A-Syntactic-Semantic-and-Pragmatic-Evaluation-of-the-Translation-of-an-Ethnographic-Text-by-ChatGPT.pdf)
- [3] Khoshafah, F. (2023). ChatGPT for Arabic-English Translation: Evaluating the Accuracy. *Research square*. Retrieved from [https://assets-eu.researchsquare.com/files/rs-2814154/v1\\_cover.pdf](https://assets-eu.researchsquare.com/files/rs-2814154/v1_cover.pdf)
- [4] Kuzman, T. et al. (2019). *Neural Machine Translation of Literary Texts from English to Slovene*. [Conference Presentation]. Dublin Machine Translation Summit Aug 2019. W19-73.pdf (aclweb.org).
- [5] Liebman, S. (2018). *How to Say Mickey Mouse in 27 Different Languages*. Disney Diary. Retrieved from <https://disneydiary.com/2018/09/how-to-say-mickey-mouse-in-27-different-languages/>
- [6] Mohamed Elghazali. (2020, October). Getting Started with Custom Translator. [Video]. YouTube. <https://www.youtube.com/watch?v=TykB6WDTkRc>
- [7] Ogundare, O. & Araya, G. Q. (2023). Comparative Analysis of CHATGPT and the Evolution of Language Models. *Arxiv*. Retrieved from <https://arxiv.org/pdf/2304.02468>
- [8] Opaluwah, A. (2024). Cultural Discourse in African Poetry: Output by Human Translators and Machine Translation Systems. *International Journal of Language, Literacy and Translation*. <https://doi.org/10.36777/ijollt2024.7.2.112>
- [9] Opaluwah, A. (2025). The Expertise of the European Translators of African Poetry. *Multidisciplinary Journal of Media and Intercultural Communication*. Retrieved from [https://www.micjournal.org/article\\_213961.html](https://www.micjournal.org/article_213961.html)
- [10] Peng, K. et al. (2023). Towards Making the Most of ChatGPT for Machine Translation. Retrieved from <https://arxiv.org/pdf/2303.13780>
- [11] Ray, P. P. (2023). ChatGPT: A comprehensive review of background, applications, key challenges, bias, ethics, limitations and future scope. *ScienceDirect*, 3, 121-154. Retrieved from <https://www.sciencedirect.com/science/article/pii/S266734522300024X>
- [12] Rodrigue, J. (2012, January 17). Muhammad Ali at 70: How the 'Louisville Lip' became 'The Greatest'. *The Guardian*. Retrieved from <https://www.theguardian.com/theguardian/from-the-archive-blog/2012/jan/17/muhammad-ali-70-greatest>
- [13] Scheub, H. (2000). *A Dictionary of African Mythology: The Mythmaker as Storyteller*. New York, USA: Oxford University Press. Kindle Edition.
- [14] Si, S. et al. (2023). Evaluating the Capability of ChatGPT on Ancient Chinese. *Arxiv*. Retrieved from <https://arxiv.org/html/2312.15304v1>
- [15] Snover, M. et al. (2009). Fluency, Adequacy, or HTER? Exploring Different Human Judgments with a Tunable MT Metric. *Aclanthology*. Retrieved from <https://aclanthology.org/W09-0441.pdf>
- [16] Soyinka, W. (1982). *Idanre: poème*. (A. Bordeaux, Trans.). Dakar: Nouvelles éditions Africaines. (Original work published 1967).
- [17] Wang, S. et al (2024). What is the Best way for ChatGPT to Translate Poetry? *Arxiv*. Retrieved from <https://arxiv.org/abs/2406.03450>
- [18] Wei, Y. (2023). A Comparative Study Between Manual and ChatGPT Translations on Literary Texts Taking Kung I-chi as an Example. *Asia-Pacific Journal of Humanities and Social Sciences*, 3(4), 57-66. Retrieved from <http://www.appublishing.com.au/public/upload/file/20231219/1702971274160762.pdf>
- [19] Zan, C. et al (2024). Building Accurate Translation-Tailored LLMs with Language Aware Instruction Tuning. Retrieved from <https://arxiv.org/pdf/2403.14399>