The Princeton Journal of Interdisciplinary Research
ISSN 3069-8200

Comparative Analysis of AI-Generated Translations of Ancient Greek Texts

Jacob Wu

Del Norte High School, San Diego, California, United States

Abstract

This paper examines differences between AI-generated and human translations of Ancient Greek texts to assess the reliability of generative AI models. Focusing on verb semantics, it compares English translations produced by ChatGPT to those by established scholars, W.H.S. Jones (Pausanias) and Charles Darwin Adams (Hippocrates), analyzing two context-independent passages. Semantic fields of verbs are assigned and compared across three versions (original Greek, human translation, and AI output). Divergences reveal that human translators often employ interpretive or stylistic choices that can shift meanings, while AI models consistently produce literal translations, sometimes missing cultural context or subtle nuances—in one case, rendering "painting" as "inscription." The findings indicate that, although AI may provide swift translations, it frequently fails to capture interpretive depth and historical context needed for classical studies. As a result, AI translations are less suitable for scholarly and

pedagogical use compared to authoritative human versions. The paper concludes that generative AI tools for classical languages should include clear warnings about their limitations, ensuring that researchers and educators use them critic ally and with caution. This study highlights the ongoing need for human expertise in interpreting ancient texts and suggests guidelines for responsible use of AI in the digital humanities and classics fields.

Keywords: AI, translation, ancient Greek

Introduction

The translation of classical texts from Ancient Greek is central to the study of classical civilization, philosophy, and medicine. Accurately conveying their meaning in modern languages requires more than fluency in Greek; it demands a deep understanding of semantic nuance and translation theory (Karimullah, 2020). Over the past century, translators such as W.H.S. Jones and Charles Darwin Adams have produced widely respected English editions of these texts, balancing literal accuracy with idiomaticity in English and the preservation of the author's tone and intent (Adams, 1927; Burton, 1938-1939; Database of Classical Scholars. (n.d.-a). JONES, William Henry Samuel; Database of Classical Scholars. (n.d.-b). ADAMS, Charles Darwin).

With recent developments in natural language processing, generative AI models such as ChatGPT now allow for immediate English translation of Ancient Greek passages without formal language training (Wannaz & Miyagawa, 2024; Rapacz & Smywiński-Pohl, 2025). While this capability is impressive, it creates new opportunities and risks for pedagogy and scholarship, it raises an important question: Can AI-generated translations match the semantic precision and interpretive depth of human translators? Comparative research indicates that AI-generated

translations often provide grammatically correct and literal outputs but may neglect cultural or idiomatic context (Elkins, 2024; Lu et al., 2025).

This study seeks to answer the question by comparing translations produced by AI models to those written by established scholars. We focus on the consistency of verbs in translation, and specifically, the preservation of semantic fields, a concept that is central to understanding the intent of both the author and translator. Semantic fields are clusters of words associated with a specific conceptual domain, such as verbs of communication ("say," "speak," or "acknowledge"), perception ("see", "appear", "seem"), or motivation ("seek", "wish", or "desire"). Categorizing verbs into these fields could help identify when and how AI and human translations diverge from the source text.

This study centers on two research questions:

(1) Do AI translations preserve the meaning of verbs in the original Greek?

(2) Are AI translations valuable tools for classical studies or pedagogy?

To test this, we selected two passages to be used as case studies. The first passage was from the treatise On Ancient Medicine from the Hippocratic Corpus, a medical treatise associated with 5th century BCE physician Hippocrates. The second passage was from Pausanias' Description of Greece, a 2nd century CE work that provides a detailed firsthand account of Ancient Greece. Both works could pose challenges for modern translation as they rely on precise terminology and require understanding of cultural or domain-specific context.

We selected human translations of Hippocrates and Pausanias made by Charles Darwin Adams, a late 19th century American classicist, and W.H.S. Jones, an early 20th century British classicist. These two authoritative human translations attempt to accurately render the source text into English, while also considering idiomaticity and author's intent. To generate comparable AI outputs, we used careful prompt engineering to ensure translations stylistically resembled the human translations.

The methodology section of this paper will explain the process of selecting the Ancient Greek passages and the AI and human translations, and the process of assigning and comparing verbs and their semantic fields between source text and target texts. Analysis includes tables of results which highlight meaningful discrepancies between the original Greek and its translations. Interpretation explains the human and AI's translation choices that resulted in those discrepancies, offering a conclusion as to the translation abilities of generative AI and a solution for possible use in classical language learning.

Method

Text Selection: Authors, Passages, and Constraints

We selected short, idiom-free passages from two Ancient Greek authors: Pausanias (2nd century CE) and Hippocrates (5th century BCE). Each passage was approximately 300 words in length. These authors offer structurally and stylistically distinct kinds of prose, which allows us to examine how generative AI performs across different textual domains. Pausanias' Description of Greece is a historical and geographical text that describes temples, monuments, and local customs. It tends to use straightforward grammar and detailed factual descriptions. Hippocrates'

medical treatises are characterized by concise language, specific terminology, and a critical distinction between agents and patients.

The temporal distance of the authors is relevant for translation analysis. Both texts pose challenges for modern translation. The medical terminology in Hippocrates, for instance, is not always translatable with standard lexica, and Pausanias, writing well after the Classical period, reflects later Greek prose conventions which may be hard to translate given a standard dictionary. Neither author can be translated with full confidence using only standard dictionaries. These challenges exemplify the kinds of linguistic and interpretive challenges that persist in classical translation.

We deliberately avoided passages that rely heavily on context, intertextual references, or poetic devices. Instead, we chose syntactically simple sentences with minimal ambiguity and idiomatic phrasing. This restriction separates out mistranslations due to errors in parsing or interpreting grammatical roles, as the focus lies in contextual misinterpretations.

Human Translations: W.H.S. Jones and Charles Darwin Adams

For comparison, we used English translations by W.H.S. Jones and Charles Darwin Adams. Jones, who translated the works of Hippocrates, combined philological rigor with a broader concern for how Greek medicine should be interpreted in cultural and philosophical contexts [4]. Adams, who also translated the works of Hippocrates, emphasized fidelity to the structural and stylistic features of the original Greek [5]. These translators were chosen not only because of their scholarly reputation but also because their work typifies early 19th or 20th century academic standards: formal, concise, and professional [2-3]. These translations remain

readable while not influenced by the same tools, language, and expectations that shape modern 21st century practice.

AI Translation: Prompt Engineering

We generated AI translations using the widely-used generative model OpenAI's ChatGPT. Recognizing that these models are highly responsive to input phrasing, we developed tailored prompts that instructed the AI to emulate the formal style of the human translators. The goal was to reduce stylistic variability and produce outputs that were as comparable as possible to those of Jones and Adams.

The two specific prompts read:

"Translate the following passage. Aim for a translation that gives the reader a true version of the language. Preserve original word order and morphology as far as intelligibility allows. Prefer literal accuracy over fluent English style even if resulting sentences seem terse, archaic, or obscure. Maintain technical and medical vocabulary rigorously, resisting paraphrase or interpretive smoothing. Use sentence division and punctuation that mirrors the Greek syntax."

"Translate the following passage. Your goal is to render a faithful, literal, and scholarly translation of the source text, adhering closely to the original Greek syntax, diction, and structure. Use formal, slightly archaic English rather than modern paraphrase. Preserve Greek word order and sentence structure where possible, even if the result sounds stilted or indirect in English. Translate prepositions and particles with precision, maintaining ambiguity if the Greek itself is ambiguous."

These prompts were meant to produce translations comparable to those of Adams and Jones, respectively. ChatGPT's temporary chat feature was turned on so that translations did not use or update the model's memory of past translations. Multiple outputs were generated for each passage, and we selected those that best matched the structure and tone of the reference translations.

Analysis

The verb is the root of a sentence, shaping the overall semantic structure of a text. Verbs carry agency and causality; they express actions and events, and they establish the relationships between subjects and objects. They require nuanced interpretation, as changes in the meaning of verbs may affect how actions and relationships are conveyed in translation. Accurate verb translation is thus essential to preserve the integrity of the source text. By focusing on verbs, this study identifies whether AI-generated translations can successfully capture the essence of the original Greek's meaning.

In order to identify these subtle differences, verbs are categorized into semantic fields. Semantic fields are inductively-defined categorizations of words that have related meanings or are associated with a specific concept. Categorizing verbs from the original Greek text, the human translation, and the AI translation into their respective semantic fields enables a comparison of the translation choices made by both translators. This can help assess whether the translations adequately preserve the semantic meaning present in the original Greek, or whether they introduce stylistic decisions that may detract from the meaning of the source. This approach provides a more holistic understanding of translation fidelity, highlighting not only literal word-

for-word mismatches, but also more subtle divergences. This is especially valuable in examining whether a translator tends to prioritize idiomaticity or word-for-word fidelity.

Tables 1 and 2 display a comparative analysis of verbs from Hippocrates' On Ancient Medicine Part 5 and Pausanias' Description of Greece 1.1.1-2, respectively. Specifically, they show only the particular instances where the assigned semantic field of at least one of the translations differs from the assigned semantic field of the original Greek. These discrepancies are highlighted in red.

Table 1

Overview of Discrepancies in Semantic Fields in Hippocrates.

Greek	Gloss	Field (Greek)	Jones	Field (Jones)	GPT	Field (GPT)
σκεψώμεθα	consider	Thought	inquire	Investigation	consider	Thought
εὑρημένην	found	Investigation	invented	Creation	found	Investigation
πρὸς*	toward	Motivation	take	Acquisition	according to*	Motivation
ζητήσαντες	sought	Investigation	cultivated	Development	sought	Investigation
εὑρόντες	found	Investigation	invented	Creation	found	Investigation
ὑφεῖλον	take away	Removal	diminished	Decrease	took away	Removal
εὗρον	found	Investigation	invented	Creation	discovered	Investigation
ἀφίκοντο	arrived at	Arrival	had recourse to	Assistance	arrived at	Arrival
προσφερόμενοι'	offered	Giving	was administered	Application	offering	Giving

Table 2

Overview of Discrepancies in Semantic Fields in Pausanias.

Greek	Gloss	Field (Greek)	Jones	Field (Jones)	GPT	Field (GPT)

πρόκειται	lies	Position	stands out	Perception	lies	Position
ἐστὶ	is	Existence	see	Perception	is	Existence
ἦν	were	Existence	had	Ownership	had	Ownership
ὑπέπλει	sailed	Sailing	in command*	Leadership	sailing	Sailing
κατεσκευάσατο	constructed	Construction	made	Establishment	constructed	Construction
ἀναθέντες	dedicated	Worship	set up	Establishment	have dedicated	Worship
γεγραμμένος	painted	Art	portrait*	Art	written	Writing

Tables 3-4 display a similar comparative analysis of verbs from the same respective passages, focusing on particular instances where the correspondence between verbs in the original Greek text and at least one of their translations is disrupted. Specifically, it refers to a verb omission, where a verb present in the Greek source text is not rendered as a verb in at least one of the target texts, or a verb addition, where at least one translator introduces a verb where the source text did not have one. In these cases, the corresponding meaning of the verb may be absent entirely, or it may be replaced by a noun, adjective, or other part of speech. These non-verbs are noted with an asterisk. Entries where the part of speech of the translated concept differs from that of the original Greek are highlighted in yellow.

Table 3

Overview of Discrepancies between Parts-of-speech in Hippocrates

Greek	Gloss	Field (Greek)	Adams	Field (Adams)	GPT	Field (GPT)
ὁμολογεομένως	acknowledged	Communication	admitted	Communication	admittedly*	Communication
κάμνοντας	suffering	Negative Experience	the sick*	Negative Experience	are ailing	Negative Experience
κάμνουσι	are ill	Negative Experience	in sickness*	Negative Experience	ailing	Negative Experience
ὑγιαίνουσιν	are healthy	Positive Experience	in good health*	Positive Experience	healthy*	Positive Experience

ὑγιαίνοντες	are healthy	Positive Experience	in health*	Positive Experience	are healthy	Positive Experience
πρὸς*	toward	Motivation	take	Acquisition	according to*	Motivation
ἐποίησαν	made little	Decrease			made very few	Decrease
			used	Application		
ἔστι	is	Existence				
φανερός ἐγένετο	became visible	Perception	manifestly*	Perception	became clear	Perception
ἰσχυρὸν*	strong	Capability	was strong	Capability	strength*	Capability
ἐδύναντο	be able	Capability			able	Capability
ἔχοι	be moderately	Regulation			be moderately composed	Regulation

Table 4

Overview of Discrepancies Between Parts-of-speech in Pausanias

Greek	Gloss	Field (Greek)	Jones	Field (Jones)	GPT	Field (GPT)
πλέοντι	sails	Sailing			sails	Sailing
			was	Description	being	Description
ὑπέπλει	sailed	Sailing	in command*	Leadership	sailing	Sailing
			went	Travel	had departed	Travel
πλέουσιν	sailing	Sailing	mariners*	Sailing	sailing	Sailing
εἶναι	be	Establishment			be	Establishment
			is	Description	was	Description
γεγραμμένος	painted	Art	portrait*	Art	written	Writing

The entries highlighted in blue are those that have been discounted. The first entry from Hippocrates, κάμνοντας, refers to a participle, which is derived from a verb but does not really function as one. Since English does not have a participial system like that of Ancient Greek, it is perfectly reasonable for the translator to diverge in their translation, and thus these instances are not meaningful discrepancies. The other entry, ἔστι, is part of an idiomatic phrase ἔστιν ὅτε, meaning “sometimes”, so this entry, while technically a verb, is also discounted. The two

highlighted entries in Pausanias result from the usage of predicate nominatives in the source text, so it is therefore necessary for translators to insert copulas in the target text. Those entries are discarded from consideration as well.

The rest of the entries in Tables 1-4 are considered meaningful divergences that are deserving of closer analysis and interpretation.

Interpretation

The comparative analysis of semantic fields and verb omissions and additions demonstrates that AI translations are rather consistent in preserving verbs and their semantic fields in both of the stylistically distinct Ancient Greek passages. While human translators such as Jones and Adams frequently make stylistic and interpretive adjustments that alter the verb's semantic field or part of speech, these changes often reflect an awareness of idiomaticity in English, cultural context, and the author's intended meaning. For example, Jones translates εὑρημένην as "invented" despite the more literal translation of "found" or "discovered," underscoring medicine as a deliberate, cultivated art as Hippocrates would have described it. In contrast, ChatGPT's translations tend toward hyper-literalism, providing word-for-word correspondence and top dictionary definitions, which can lead to awkward semantic distortions that sometimes ignore relevant cultural context and detract from the author's intended meaning. While this could offer more accuracy to the meaning of the stand-alone word, there was one instance of an unequivocal inaccuracy in translation: Pausanias uses the participle γεγραμμένος to refer to a painting, while ChatGPT wrongfully translate this as "written," ignoring context and simply following the dictionary gloss.

The human translations show that divergence from literal meaning is not necessarily a flaw. These choices reframe a verb's meaning to capture its idiomatic sense or to smooth the reading experience in English. These stylistic interpretations are often guided by the translator's understanding of historical context, authorial style, and genre conventions. In contrast, GPT's more literal approach loses interpretive nuance, and when its lexical choice is incorrect, it is typically wrong in a way that signals a lack of access to contextual knowledge rather than an alternative but still accurate reading (Yuan & Gan, 2024). Overall, the intended meanings of sentences are generally better communicated in the human translations. Nonetheless, the meanings of verbs on their own are preserved in AI translations, while they are frequently altered in human translations.

From a pedagogical perspective, this distinction matters. AI translations are attractive to early-stage learners, because they are immediate and require no training in the source language (Ross, 2023). However, this creates a sort of paradox: the learners most likely to depend on AI are also the least equipped to identify or correct its errors. Without an understanding of the source language, a student might accept an inaccurate translation as authoritative, particularly if the text "sounds" plausible, which it often does. This uncritical acceptance of AI stems from a broader cultural attitude towards AI systems, in which they are treated as objective and encyclopedic. More advanced learners such as graduate students or scholars are less at risk to this, since they will most likely have access to a broader range of resources and thus can cross-check AI output against dictionaries, commentaries, and other translations. In these contexts, AI can serve as a useful auxiliary tool for quickly generating literal translations, exploring lexical options, or testing one's own reading abilities.

Given these findings, AI translations should not be used exclusively as stand-alone scholarly resources. AI systems already provide warnings to sensitive or regulated subjects such as financial or medical advice. In a similar vein, translation tools could thus display a targeted warning such as: “This translation may contain inaccuracies and should not replace work by a qualified translator or consultation of authoritative editions.” Such messaging would be especially important in educational settings where students may lack access to multiple resources.

In summary, while AI translation models can produce fluent and generally accurate renderings of Ancient Greek texts that preserve the semantic meanings of verbs in a rather literal manner, they lack the interpretive depth and context that a human translator would have. In classical language learning contexts, they cannot replace human translations, but could function as one of many tools to help support language learning or translation for users who are able to identify and correct their shortcomings. Institutions, educators, and platform designers should consider both technical improvements and practical safeguards such as usage disclaimers to ensure that AI translation is used responsibly and effectively.

Summary

This study compared ChatGPT-generated translations of Ancient Greek texts with established human translations in order to evaluate whether AI models can preserve the semantic meaning of verbs. Passages from Hippocrates’ On Ancient Medicine and Pausanias’ Description of Greece were analyzed alongside respective translations by W.H.S. Jones and Charles Darwin Adams. The analysis focused on the semantic fields of verbs to capture the broader concept conveyed, such as communication, perception, or motivation. Results indicate that ChatGPT’s

translations tend towards hyper-literalism, generally preserving surface-level verb meanings, while lacking the nuance to adjust for idiomaticity and cultural context. On the other hand, human translators make interpretive or stylistic decisions that often differ in semantic field in order to better reflect authorial intent. These implications are significant for pedagogy: students with little background in Ancient Greek may accept AI translations without the ability to verify accuracy. This may function as a useful tool for more advanced readers, however, functioning more as a literal aid instead of replacing human scholarship. Overall, AI translations of Ancient Greek demonstrate promise in producing readable approximations, but fall short in interpretive depth.

Conclusion

Comparative analysis demonstrates that AI translations do preserve the meanings of Ancient Greek verbs, albeit in a narrowly literal way that lacks interpretive nuance. AI often defaults to the top dictionary definition, which generally does produce functional translation. However, there are some instances where AI translations are inaccurate, and these instances highlight risks for pedagogy. For new learners, over-reliance on AI can develop misplaced trust on AI-generated translations. Without the ability to verify mistakes made by AI translations, learners may accept distortions without critical analysis. To prevent this, AI translation platforms should display targeted warning messages such as “This translation may contain inaccuracies and should not replace work by a qualified translator or consultation of authoritative editions.” Similar to what ChatGPT outputs when asked for financial advice, such safeguards would help emphasize the current limits of AI. Used responsibly, AI translation may act as a useful tool to aid higher level learners who will have access to a broader range of resources and be able to cross-check AI output against dictionaries, commentaries, and other translations. In these

contexts, AI may be used for quickly generating literal translations, exploring lexical options, or testing one's own reading abilities.

Outlook

This analysis centers around the preservation of verbs across translation. Future research should broaden beyond verbs, examining the semantic discrepancies among nouns, adjectives, and other parts of speech that also shape the meaning of sentences in more subtle ways. Other models beyond ChatGPT should be tested as well, including Gemini and Claude. Likewise, other methods of comparing semantic meaning such as the semantic role labeling could add further precision in capturing differences in translation. In addition, the corpus could be expanded to include a wider variety of Ancient Greek texts as well as more modern translation. Given the increasing prevalence of AI translation, more systematic evaluation is needed to properly assess their reliability and what the appropriate use should be in scholarship and pedagogy.

References

Adams, C. D. (1927). Demosthenes and His Influence. Longmans, Green and Co.

Burton, H. E. (1938–1939). Charles Darwin Adams. Classical Journal, 34, 59–60.

Database of Classical Scholars. (n.d.-a). JONES, William Henry Samuel. https://dbcs.rutgers.edu/all-scholars/8684-jones-william-henry-samuel

Database of Classical Scholars. (n.d.-b). ADAMS, Charles Darwin. https://dbcs.rutgers.edu/all-scholars/8495-adams-charles-darwin

Elkins K (2024) In search of a translator: using AI to evaluate what's lost in translation. Front. Comput. Sci. 6:1444021. doi: 10.3389/fcomp.2024.1444021

Karimullah, K.I. Hippocrates transformed: crafting a Hippocratic discourse of medical semiotics in English, 1850–1930. Humanit Soc Sci Commun 7, 27 (2020).

Lu SC, Xu C, Kaur M, Edelen MO, Pusic A, Gibbons C. Can machine translation match human expertise? Quantifying the performance of large language models in the translation of patient-reported outcome measures (PROMs). J Patient Rep Outcomes. 2025 Jul 25;9(1):94. doi: 10.1186/s41687-025-00926-w.

Rapacz, M., Smywiński-Pohl, A. (2025). Morphology-enhanced neural models for Ancient Greek. In H. Hettiarachchi, T. Ranasinghe, P. Rayson, R. Mitkov, M. Gaber, D. Premasiri, F. A. Tan, & L. Uyangodage (Eds.), Proceedings of the First Workshop on Language Models for Low-Resource Languages (pp. 145–154). Association for Computational Linguistics.

Ross EAS. A New Frontier: AI and Ancient Language Pedagogy. Journal of Classics Teaching. 2023;24(48):143-161. doi:10.1017/S2058631023000430

Wannaz, A.-C., Miyagawa, S. (2024). Assessing large language models in translating Coptic and Ancient Greek ostraca. In M. Hämäläinen, E. Öhman, S. Miyagawa, K. Alnajjar, & Y. Bizzoni (Eds.), Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities (pp. 463–470). Association for Computational Linguistics.

Yuan, X. Gan, L. (2024). Classic Translation and Humanistic Spirit in the Age of AI. Journal of Literature and Art Studies, 14(10), 831–838. doi: 10.17265/2159-5836/2024.10.001