

Übersetzen von Latein und Altgriechisch mit ChatGPT – Reloaded: Zur Struktur interpretativer Entscheidungen

Sylvia Kurowsky

Abstract: Building on an earlier exploratory study of AI-assisted translation from Latin and Ancient Greek, this article examines how current versions of ChatGPT reshape translation as a guided, dialogical workflow. Focusing on the role of instruction and interaction settings, the paper revisits the same test passages used in the initial study and compares a text-close version with a fluent, reader-oriented translation across different model modes. The case studies – drawn from ancient texts (Cicero, Plutarch) as well as modern scholarly texts written in Latin and Ancient Greek – reveal a stable structural pattern: while text-close translations preserve semantic openness, reader-oriented versions enhance coherence and readability but frequently resolve ambiguity and introduce implicit interpretative commitments. Variations between model modes affect the degree of semantic specification, yet do not alter this underlying functional distinction. AI-assisted translation thus emerges as a heuristic instrument that does not replace philological control, but shifts and reconfigures interpretative decisions already at the stage of translation itself.

1. Einleitung

Übersetzungen aus dem Lateinischen und Altgriechischen mit Hilfe von Large Language Models (LLM) sind inzwischen zu einem festen Bestandteil des wissenschaftlichen Alltags geworden. Während frühere Beiträge vor allem erkundeten, ob LLMs wie ChatGPT grundsätzlich in der Lage sind, antike Texte in moderne Sprachen zu übertragen, hat sich die Fragestellung inzwischen verschoben.¹ Nicht mehr das ‚Ob‘, sondern das ‚Wie‘ steht im Zentrum: Wie arbeitet ein dialogfähiges LLM bei Übersetzungsaufgaben, wie lassen sich seine Ergebnisse steuern und welche Formen der philologischen Kontrolle bleiben dabei unverzichtbar?²

- 1 Zu qualitativen, textanalytischen Bewertungen der Übersetzungsleistung von ChatGPT bei altgriechischen und lateinischen Texten s. Kurowsky (2025). Zu Studien, die Bewertungsverfahren und Kriterien literarischer Übersetzung mit LLMs systematisch reflektieren, s. Zhang et al. (2025). Quantitative Evaluationsstudien, die Übersetzungsqualität anhand numerischer Vergleichswerte oder standardisierter Benchmarks messen, bieten Jiao et al. (2023) für eine benchmark-basierte Bewertung von ChatGPT als Übersetzungssystem, Wannaz / Miyagawa (2024) für eine Bewertung der Übersetzungsleistung antiker Sprachen (Altgriechisch und Koptisch) anhand von Referenzübersetzungen und berechneten Übereinstimmungswerten sowie Roşu et al. (2025) für eine technisch ausgerichtete Pipeline-Studie zur textnahen Übersetzung lateinischer Texte (auf Basis von BLEU/BLEURT-Scores). Vgl. ferner Volk et al. (2024) zu Latein und historischen Sprachstufen.
- 2 Die gegenwärtige Diskussion um LLMs ist dabei nicht als singuläres Ereignis zu verstehen, sondern als Teil einer längerfristigen digitalen Transformation geisteswissenschaftlicher Arbeitsweisen. Bereits vor 20 Jahren wurde darauf hingewiesen, dass digitale Repräsentationsformen nicht nur Werkzeuge darstellen, sondern die epistemischen Bedingungen von Textarbeit verändern, s. Crane (2006): Er betont im Kontext digitaler Bibliotheken, dass computergestützte Analyseformen den Maßstab und damit auch die Art geisteswissenschaftlicher Fragestellungen verschieben. So kann auch die KI-gestützte Übersetzung nicht lediglich als technisches Hilfsmittel, sondern als Modifikation wissenschaftlicher Praxis begriffen werden.

Denn große Sprachmodelle wie ChatGPT erzeugen Übersetzungen nicht auf der Grundlage expliziter Bedeutungsrepräsentationen, sondern als statistisch wahrscheinliche sprachliche Fortsetzungen. Ihre Ausgaben gewinnen Plausibilität primär aus formaler Kohärenz und Mustertreue, nicht aus überprüfbaren Wahrheitsbedingungen. Da diese Modelle probabilistisch operieren und keine eigenen Kriterien zur Kennzeichnung epistemischer Unsicherheit implementieren, werden Mehrdeutigkeiten, Ellipsen oder editorische Probleme in der Regel nicht als solche markiert, sondern in kohärente Zieltextstrukturen überführt. Gerade hierin liegt die methodische Relevanz für die Philologie: Die durch Glättung entstehende Kohärenz darf nicht mit semantischer Eindeutigkeit verwechselt werden. Diese grundlegende Modelllogik ist früh dargelegt worden³ und bildet einen geeigneten Rahmen für die folgende Untersuchung.

Der vorliegende Beitrag knüpft an einen früheren Aufsatz der Verfasserin an,⁴ in dem die Übersetzungsleistungen von ChatGPT in einer damaligen Modellversion⁵ erstmals erprobt und anhand ausgewählter Beispiele diskutiert sowie auch bewertet wurden.⁶ Diese Untersuchung war bewusst explorativ angelegt und zielte darauf, Möglichkeiten und Grenzen des Systems im schulischen und akademischen Kontext auszuloten. Seitdem haben sich jedoch sowohl die technischen Grundlagen der Modelle als auch ihre praktische Nutzung spürbar verändert, so dass eine erneute Befragung des Gegenstands notwendig erscheint.⁷

Rückblickend zeigt sich zudem, dass die damalige Vorgehensweise methodische Grenzen aufwies, die für die vorliegende Neubearbeitung zentral sind. Zum einen wurde ChatGPT stellenweise nicht nur als Übersetzungswerkzeug, sondern auch als Auskunftsource über seine eigenen Funktionsweisen, Trainingsgrundlagen und zeitliche Abdeckung herangezogen. Solche Selbstauskünfte eines Sprachmodells lassen sich jedoch nicht als belastbare Metainformationen verstehen, sondern sind als Teil seines generierten Antwortverhaltens selbst Gegenstand der Analyse. Ihre Einbeziehung in der früheren Untersuchung diente primär der heuristischen Annäherung an ein neuartiges System. Annahmen über klar bestimmbare Trainingszeiträume oder eindeutig abgrenzbare Datengrundlagen⁸ sind auf dieser Grundlage methodisch nicht absicherbar. Zum anderen wurden die damaligen Übersetzungen überwiegend mit

3 Schon Floridi / Chiriatti (2020) zeigen, dass GPT-3 Texte als statistisch plausible sprachliche Artefakte erzeugt, ohne semantisches Verstehen oder epistemische Autorität – und dass genau diese Entkopplung von Textqualität und Bedeutung methodisch ernst genommen werden muss. ChatGPT basiert – wie andere LLMs – auf tiefen, transformer-basierten Modellen, deren Entwicklung und Evaluierung in umfassenden Untersuchungen dargelegt worden ist wie z. B. in Zhao et al. (2023).

4 Kurowsky (2025).

5 GPT 4.o.

6 Beitrag ausführlich besprochen in Schubert (2026), 65–75. Für ihre Einschätzung von ChatGPT und LLMs allg. vgl. Schubert (2026), 182–183.

7 OpenAI hat GPT-5.2 am 11. Dezember 2025 als Nachfolger von GPT-5.1 veröffentlicht: <https://openai.com/de-DE/index/introducing-gpt-5-2/> (alle Links abgerufen am 25.02.2026). GPT-5.2 ist das Modell für diese aktualisierte Übersetzungsuntersuchung.

8 Im Gegensatz zu älteren Modellen weist GPT-5.2 einen Knowledge Cutoff (‘Wissensstand’ bzw. genauer ‘Trainingsdatenstand’) vom August 2025 auf: <https://academy.openai.com/public/resources/latest-model>, hierbei handelt es sich wahrscheinlich um eine dynamische URL, daher sind eine PDF-Version der Website ebenso wie alle Chatprotokolle zur dauerhaften Nachvollziehbarkeit im Datenrepositorium (Link: <https://doi.org/10.11588/propylaeumdok.00006995>) abgelegt. Das bedeutet: Soweit das Modell ohne Websuche operiert, ist es mit ‘Wissen’ bis zu diesem Datum trainiert worden. Das ist ein klar dokumentierter Fortschritt gegenüber früheren Versionen, deren Trainingsdatenstand (z. B. häufig als ‘bis 2021’ angegeben, s. Kurowsky [2025], 3, 13) auf einem deutlich früheren Stichtag beruhte und damit einen entsprechend begrenzteren Datenhorizont aufwies. Vgl. hierzu auch: <https://www.allmo.ai/articles/list-of-large-language-model-cut-off-dates>.

minimaler Instruktion⁹ erzeugt, häufig beschränkt auf die schlichte Aufforderung, einen gegebenen Text zu übersetzen.

Der Unterschied liegt nicht primär in einer linearen ‚Verbesserung‘ der Übersetzungsleistung, sondern in der gesteigerten und konsistenteren Instruktionssensitivität aktueller Modellversionen, die eine differenziertere methodische Steuerung ermöglicht: Gegenwärtige LLMs reagieren in hohem Maße auf differenzierte Instruktionen und erlauben eine gezielte Steuerung des Übersetzungsprozesses, etwa im Hinblick auf Textnähe, Lesbarkeit oder interpretative Zurückhaltung. Übersetzungen sind damit nicht mehr als singuläre Ergebnisse zu verstehen, sondern als Produkte eines instruktionssensitiven Arbeitsprozesses,¹⁰ in dem Zielvorgaben und Modellreaktionen eng ineinandergreifen. Die Übersetzungsleistung hängt dabei in hohem Maße von der expliziten Formulierung der jeweiligen Aufgabenstellung ab.

Ziel dieses Beitrags ist es daher nicht, eine bloße Aktualisierung früherer Testergebnisse vorzunehmen oder die Übersetzungsqualität pauschal zu bewerten.¹¹ Vielmehr soll gezeigt werden, dass sich das Übersetzen mit ChatGPT heute als gesteuerter Arbeitsprozess beschreiben lässt, in dem Übersetzungsziele, Instruktionen und Modellreaktionen wechselseitig aufeinander abgestimmt werden.¹² Der Erkenntnisgewinn liegt dabei weniger in der Beurteilung einzelner Ergebnisse als in der Beobachtung, wie und an welchen Stellen Übersetzungs- und Interpretationsentscheidungen überhaupt einsetzen.¹³ KI-gestützte Übersetzungen strukturieren damit nicht nur den Zieltext, sondern verschieben den Zeitpunkt und den Ort philologischer Entscheidung – eine Verschiebung, die methodisch reflektiert werden muss.

Der Beitrag versteht sich als methodischer Erfahrungsbericht aus der Praxis, nicht als technologische Bewertung. Übersetzen mit KI erscheint hier weder als Ersatz philologischer Kompetenz noch als bloßes Hilfsmittel, sondern als ein Instrument, das die Verantwortung der Übersetzenden neu akzentuiert und eine bewusste, reflektierte Arbeitsweise erfordert.

9 Für ChatGPT wirken Instruktionen als externe Entscheidungsparameter, s. hierfür grundlegend Ouyang et al. (2022). Auch Wei et al. (2022) zeigen, dass Instruktionen den internen Verarbeitungspfad verändern. Prompts steuern nicht nur den Output, sondern die Entscheidungsstruktur. Zum Prompting als eigenes Paradigma s. Liu et al. (2023) und Lee (2024), 2354. Überdies betont z. B. Pym (2023), 121–123, dass Übersetzungsentscheidungen von mehr als textinternen Faktoren abhängen: Ausgangstexte legen ihre Übersetzungen nicht eindeutig fest. Das bedeutet jedoch keine Beliebigkeit, sondern macht Übersetzen zu einem Prozess interpretativer Entscheidungen, der auf Beobachtung, Gewichtung und Kontextualisierung beruht. Toury (1995), 53–69 versteht Übersetzen als normengeleitete Praxis, bei der Entscheidungen nicht regelhaft, sondern durch implizite, kulturabhängige Orientierungsmuster gesteuert werden, die sich erst im Übersetzungsprozess selbst manifestieren. Dass Verstehen nicht voraussetzungslos erfolgt, sondern stets durch vorgängige Orientierungsmuster strukturiert ist, arbeitet Gadamer (1960 = 1990), 270–281 grundlegend heraus. Auch im Übersetzungsprozess wirken solche Vorstrukturierungen bereits vor der eigentlichen Analyse, etwa in Form expliziter Instruktionen. Zur systematischen Variation zwischen Ausgangs- und Zieltexten in antiken Sprachen, bedingt durch linguistische, textuelle und kulturelle Faktoren, s. Palladino / Yousef (2023).

10 Nämlich als Produkt für diesen Moment, in dem Übersetzung stattfindet. Grundsätzlich ist jede Neuübersetzung als Teil eines Verständnis-Prozesses anzusehen, wie im Falle von Editionen, vgl. Sahle (2013a), 244.

11 Vgl. oben Anm. 1.

12 Die zunehmende Einbettung von Übersetzen in technologiegestützte Umgebungen und die damit einhergehende Verschiebung der Übersetzerrolle werden in der Translatologie seit Längerem diskutiert. Übersetzen ist dabei nicht mehr als isolierter Akt zu verstehen, sondern als Teil einer komplexen Mensch-Maschine-Konstellation, in der Steuerung, Bewertung und Entscheidungsfindung zentrale Rollen spielen, s. O’Hagan (2020).

13 Zu Übersetzungsentscheidungen, allerdings noch für GPT-3, vgl. Brown et al. (2020).

2. Methodisches Vorgehen

Die folgenden Beobachtungen beruhen auf erneuten Übersetzungstests mit ChatGPT-5.2. Dabei wurden die initialen Übersetzungstests im standardmäßig aktivierten Nutzungsmodus durchgeführt (Modus: Auto; Personalisierung und gespeicherte Erinnerungen aktiviert). Das Setting entspricht der üblichen dialogischen Verwendung des Systems im Arbeitskontext der Verfasserin und wurde daher als eigenständiger Untersuchungsmodus berücksichtigt. Neben standardisierten Testumgebungen erlaubt gerade die Dokumentation realer Nutzungssituationen eine differenzierte Einschätzung der praktischen Einbindung von LLMs in philologische Arbeitsprozesse. Später wurden Vergleichstests unter deaktivierter Personalisierung und Erinnerung sowie mit expliziter Moduswahl ergänzt, die der methodischen Kontrolle dienen und eine differenzierte Betrachtung möglicher Kontext- und Moduseffekte erlauben.

Im Unterschied zu früheren Versuchen der Verfasserin, bei denen die KI lediglich mit der Aufforderung „Bitte übersetze den folgenden Text“¹⁴ konfrontiert worden war, wurde hier ein stärker gesteuertes Vorgehen gewählt. Ausgangspunkt war die Überlegung, dass Übersetzen, insbesondere aus klassischen Sprachen, kein einheitlicher Vorgang ist, sondern je nach Zielsetzung unterschiedliche Anforderungen an Textnähe, Lesbarkeit und Interpretation¹⁵ stellt.

Dementsprechend wurden die Übersetzungen jeweils in zwei Varianten angefertigt. In einem ersten Durchgang wurde ChatGPT zu einer möglichst textnahen Übersetzung angehalten.¹⁶ Sie zielte darauf, potenzielle Mehrdeutigkeiten und strukturelle Alternativen sichtbar zu machen. Der dadurch entstehende Variantenraum ist daher nicht stets mit genuiner textimmanenter Ambiguität identisch, sondern reflektiert teilweise auch die Offenlegung möglicher Anschlussstellen durch das Modell. Diese Variante entspricht dem, was in der philologischen Praxis häufig als Arbeitsgrundlage für die detaillierte Textanalyse dient.

In einem zweiten Durchgang wurde im selben Chat¹⁷ dieselbe Textstelle als stärker interpretierende, d.h. leserorientierte Übersetzung,¹⁸ angefordert. Hier lag der Schwerpunkt auf Verständlichkeit und Lesbarkeit, ohne jedoch zusätzliche historische, biographische oder sachliche Informationen zu ergänzen.

14 Wie z. B. in Kurowsky (2025), 18.

15 In der Übersetzungswissenschaft gilt der Übersetzungsprozess als interpretativer Akt, der notwendigerweise subjektive Deutungsanteile umfasst, vgl. exemplarisch Sahle (2013b), 147, 153.

16 Der Begriff ‚textnah‘ bezeichnet hier keine strikt wortwörtliche oder lexikographisch minimalistische Übersetzung. Gemeint ist vielmehr eine Übersetzungsweise, die syntaktische Strukturen, semantische Mehrdeutigkeiten und valenzbedingte Offenheiten des Ausgangstextes möglichst sichtbar hält und interpretative Festlegungen nicht vorwegnimmt.

17 Dadurch sind die beiden Versionen nicht unabhängig voneinander, sondern stehen in einem sequenziellen Verhältnis. Genau diese Praxis entspricht jedoch einem typischen Arbeitsablauf, bei dem ein erster Übersetzungsschritt, zumindest gedanklich, als Ausgangspunkt für weitere Zielvorgaben dient.

18 In den zugrunde liegenden Prompts wurde diese zweite Variante als „flüssige deutsche Arbeitsübersetzung“ bezeichnet. Der Begriff knüpft terminologisch an Kurowsky (2025) an, wo ChatGPT-Übersetzungen insgesamt als „brauchbare Arbeitsübersetzungen“ charakterisiert wurden (ebd., 30). Im vorliegenden Beitrag wird diese Bezeichnung präzisiert und als ‚leserorientierte Übersetzung‘ gefasst, um sie klar von der textnahen Variante zu unterscheiden. Im Kontext KI-gestützter Übersetzung bezeichnet ‚leserorientiert‘ eine Fassung, deren Formulierungen so gestaltet sind, dass sie für einen generischen menschlichen Rezipienten unmittelbar verständlich und kohärent erscheinen. Gemeint ist damit keine individuelle Anpassung an konkrete Leserprofile, sondern die Orientierung an wahrscheinlichen Mustern menschlicher Sprachrezeption. Anschlussfähigkeit fungiert hier nicht als ästhetische Kategorie, sondern als kommunikative Funktionsbedingung. Empirische Untersuchungen zu LLMs zeigen, dass solche Systeme systematisch zu kohärenter, dialogisch anschlussfähiger Sprachproduktion tendieren und entlang menschenzentrierter Dimensionen wie ‚Interaction Naturalness‘ und ‚Intent Alignment‘ evaluiert werden, s. Wang et al. (2024), Mehta (2026). Die in der HCI- und NLP-Forschung beschriebenen user-zentrierten Designparadigmen unterstreichen, dass Large Language Models zunehmend als interaktive, auf menschliche Verständlichkeit ausgerichtete Systeme konzipiert und untersucht werden, s. Belosevic (2025). Der Ausdruck ‚leserorientiert‘ bezeichnet darüber hinaus folglich die funktionale Ausrichtung der Übersetzung, nicht eine geschlechtsspezifische Adressierung.

zen, die im Ausgangstext nicht explizit enthalten sind.¹⁹ Diese zwei Übersetzungsvarianten wurden im selben Chat erzeugt. Die leserorientierte Übersetzung entstand damit sequenziell auf Grundlage derselben Interaktionsumgebung.

Ziel der Untersuchung ist jedoch kein experimentell kontrollierter Vergleich unabhängiger Modellläufe, sondern die Analyse der Verschiebungen, die sich unter variiertem Zielvorgabe innerhalb eines typischen Arbeitskontextes zeigen. Auf diese Weise lassen sich Unterschiede zwischen textnaher und stärker glättender Übersetzung sichtbar machen, wie sie auch bei menschlichen Übersetzungen regelmäßig auftreten. Die Unterscheidung zwischen textnaher und leserorientierter Übersetzung lässt sich bereits klassisch bei Schleiermacher verorten.²⁰ Neu ist hier jedoch, dass diese Differenz nicht mehr eine einmalige methodische Grundentscheidung darstellt, sondern durch Prompt-Variation gezielt erzeugbar und modulierbar gemacht wird.

Die Verwendung unterschiedlicher Instruktionen²¹ dient dabei nicht der Optimierung im Sinne eines ‚bestmöglichen‘ Ergebnisses, sondern der analytischen Vergleichbarkeit. Sie ist als methodische Variable zu verstehen: In der aktuellen LLM-Forschung gilt die Promptgestaltung als zentraler Einflussfaktor auf Struktur und Ausrichtung der Antworten.²² Die unterschiedlichen Übersetzungsvarianten sind daher nicht als Eigenschaften des Modells im abstrakten Sinne zu lesen, sondern als Reaktionen auf explizit gesetzte Zielvorgaben. Sie erlauben es, typische Stärken und Schwächen des Modells präziser zu beobachten, etwa im Umgang mit komplexer Syntax, Mehrdeutigkeiten, impliziten Bezügen oder metaphorischer Sprache. Gleichzeitig wird deutlich, in welchem Maße die Qualität der Übersetzung von der expliziten Zielsetzung abhängt.

Die Auswertung der Übersetzungen erfolgt aus philologischer Perspektive. Maßstab sind dabei keine technischen Kennzahlen, sondern etablierte Kriterien der Textarbeit: Genauigkeit im Umgang mit Morphologie und Syntax, Sensibilität für semantische Nuancen, Transparenz von Strukturentscheidungen sowie der Grad interpretativer Eingriffe. Im Zentrum steht jedoch nicht die Frage nach der ‚Qualität‘ einzelner Lösungen im Sinne einer normativen Richtig-oder-Falsch-Bewertung,²³ sondern die Analyse der Entscheidungsstruktur, die sich im jeweiligen Übersetzungsmodus manifestiert.

Wo es analytisch sinnvoll erscheint, werden die KI-Übersetzungen mit bestehenden menschlichen Übersetzungen verglichen, ohne letztere als normative Referenzinstanz anzusetzen. Dabei wird nicht die Hierarchisierung von Ergebnissen, sondern die Sichtbarmachung unterschiedlicher Strategien im Umgang mit sprachlicher Offenheit, Ellipsen, Mehrdeutigkeiten und syntaktischen Spannungen ange-

19 Aufgrund der Kürze der hier untersuchten Texte und Interaktionen zeigten sich keine Anzeichen kontextabhängiger Performanzverschiebungen. In längeren Dialogsequenzen kann es jedoch zu graduellen Veränderungen der terminologischen Präzision oder zu impliziten Selbstangleichungen an zuvor generierte Formulierungen kommen – Effekte, die in der Forschung zur Kontextnutzung großer Sprachmodelle beschrieben werden, wie z. B. in Liu et al. (2024). Solche Phänomene betreffen nicht die grundsätzliche instruktionsgebundene Struktur der Differenz zwischen textnaher und leserorientierter Übersetzung, können jedoch die Stabilität einzelner Ausformulierungen beeinflussen. Eine methodisch kontrollierte Nutzung KI-gestützter Übersetzung erfordert daher neben klaren Zieldefinitionen auch eine bewusste Steuerung und gegebenenfalls Begrenzung des Interaktionskontexts.

20 Vgl. Schleiermacher (1813 / 1963), 47, der zwischen einer Übersetzung unterscheidet, die den Leser zum Autor hinführt, und einer, die den Autor dem Leser annähert.

21 Die Variation der Prompts kann als gezielte Steuerung des Inputs verstanden werden, ähnlich der Idee der ‚kontrollierten Sprache‘, also einer standardisierten und in Wortschatz sowie Syntax bewusst eingeschränkten Varietät, die zur Reduktion von Mehrdeutigkeiten und zur Verbesserung maschineller Übersetzung entwickelt wurde, s. Bowker / Buitrago Ciro (2019), 37–54.

22 Vgl. z. B. Wei et al. (2022) sowie Liu et al. (2023).

23 Vgl. zu solchen Kategorisierungen in ‚richtig‘ und ‚falsch‘ Pym (2023), 124 und noch deutlicher in seiner vorigen Edition desselben Werkes Pym (2014), 102: Übersetzungsentscheidungen sind nach ihm nicht einfach richtig oder falsch, sondern beruhen auf bewussten Abwägungen, da Übersetzende nie über vollständige Informationen hinsichtlich Kontext, Intention oder Wirkung verfügen.

strebt. Die Untersuchung erhebt dabei keinen Anspruch auf vollständige historisch-philologische Durchdringung jeder einzelnen Detailfrage. Stattdessen erfolgt eine selektive Betrachtung exemplarischer Passagen, an denen sich die Struktur von Offenhaltung, Kohärenzbildung und semantischer Fixierung besonders deutlich beobachten lässt. Die Auswahl folgt somit nicht dem Prinzip maximaler Vollständigkeit, sondern analytischer Relevanz.

Digitale Methoden spielen in diesem Beitrag insofern eine Rolle, als ChatGPT als Werkzeug in einem digitalen Arbeitsprozess eingesetzt wird. Eine weitergehende Modellierung, Quantifizierung oder algorithmische Analyse der Ergebnisse ist hingegen nicht intendiert. Ziel ist es vielmehr, aus der praktischen Arbeit mit dem System methodische Einsichten zu gewinnen, die für die philologische und alt-historische Textarbeit anschlussfähig sind.

3. Übersetzungen mit ChatGPT

Im Folgenden werden ausgewählte Übersetzungen mit ChatGPT vorgestellt. Als Testmaterial dienten sowohl antike Texte als auch moderne wissenschaftliche Texte in lateinischer und altgriechischer Sprache. Während die antiken Beispiele einen Rückbezug von ChatGPT auf etablierte Übersetzungen erlauben, wurden mit den modernen Texten bewusst Stellen herangezogen, die für das Modell weniger wahrscheinlich in verbreiteten Übersetzungsressourcen aufzufinden sind. Auf diese Weise soll überprüft werden, wie sich ChatGPT bei Übersetzungsaufgaben verhält, wenn eine Wiedererkennung des Textes oder ein Rückgriff auf vorhandene Übersetzungen unwahrscheinlich ist.

3.1 Übersetzungen antiker Texte

Für den ersten Teil der Analyse wurden lateinische und altgriechische Texte herangezogen, die inhaltlich eng aufeinander bezogen sind. Konkret handelt es sich um eine Äußerung Catilinas, die sowohl bei Cicero als auch in der Darstellung Plutarchs überliefert ist. Da Plutarch an dieser Stelle ausdrücklich auf Ciceros Bericht Bezug nimmt und dessen Aussage interpretiert, bietet sich hier ein direkter Vergleich an.

Um den jeweiligen Kontext angemessen zu berücksichtigen, wurden nicht nur die unmittelbar zitierte Passage, sondern auch jeweils ein bis zwei vorangehende und folgende Sätze in die Übersetzungsaufgabe miteinbezogen. Die Textstellen wurden jeweils unter zwei unterschiedlichen Zielvorgaben mithilfe von ChatGPT übersetzt:²⁴ einer möglichst textnahen und einer leserorientierten Übersetzung (s. Tab. 1). Der Vergleich dieser beiden Varianten ermöglicht es, Unterschiede im Umgang mit Syntax, Mehrdeutigkeit und impliziten Bezügen sichtbar zu machen sowie typische Verschiebungen zwischen textnaher und glättender Übersetzung herauszuarbeiten.

24 In dem früheren Beitrag (Kurowsky [2025]) wurden die KI-Übersetzungen mit einer eigenen textnahen Übersetzung sowie mit bestehenden humanen Übersetzungen verglichen. In der vorliegenden Untersuchung wird darauf bewusst verzichtet. Der Fokus liegt hier nicht auf dem Vergleich zwischen menschlicher und KI-basierter Übersetzung, sondern auf der Analyse von KI-Übersetzungen unter unterschiedlichen Zielsetzungen (textnah vs. leserorientiert). Eine zusätzliche Referenzübersetzung würde diesen methodischen Schwerpunkt eher überlagern als schärfen. Je nach individuellem Bedarf können jedoch zum Vergleich die menschlichen Übersetzungen aus Kurowsky (2025), 15ff. herangezogen werden.

Cic. Mur. 51 (Text: Clark [1905])	ÜS ChatGPT (textnah) ²⁵	ÜS ChatGPT (leserorientiert) ²⁶
<i>Itaque postridie frequenti senatu Catilinam excitavi atque eum de his rebus iussi, si quid vellet, quae ad me adlatae essent dicere. Atque ille, ut semper fuit apertissimus, non se purgavit sed indicavit atque induit. Tum enim dixit duo corpora esse rei publicae, unum debile infirmo capite, alterum firmum sine capite; huic, si ita de se meritum esset, caput se vivo non defuturum. Congemuit senatus frequens neque tamen satis severe pro rei indignitate decrevit; nam partim ideo fortes in decernendo non erant, quia nihil timebant, partim, quia <omnia>.</i>	Daher ließ ich am folgenden Tag, bei zahlreich versammeltem Senat, Catilina herbeirufen und befahl ihm in Bezug auf diese Dinge, falls er irgendetwas wolle, das vorzutragen, was mir zugetragen worden war. Und er, wie er stets von größter Offenheit war, verteidigte sich nicht, sondern legte offen dar und belastete [sich selbst / andere]. Denn damals sagte er, es gebe zwei Körper des Gemeinwesens: den einen schwach, mit krankem Haupt, den anderen fest, ohne Haupt; diesem werde, wenn man sich ihm gegenüber so verdient gemacht habe, ein Haupt, solange er lebe, nicht fehlen. Der zahlreich versammelte Senat stöhnte auf und beschloss dennoch nicht hinreichend streng entsprechend der Ungeheuerlichkeit der Sache; denn zum Teil waren sie deshalb nicht entschlossen im Beschließen, weil sie nichts fürchteten, zum Teil, weil [alles / alle Dinge / alle Umstände].	Daher ließ ich am folgenden Tag den Senat zahlreich zusammentreten, rief Catilina herbei und forderte ihn auf, zu diesen Angelegenheiten, falls er es wolle, zu dem Stellung zu nehmen, was mir zugetragen worden war. Er aber, wie er es immer war, von größter Offenheit, versuchte nicht, sich zu rechtfertigen, sondern legte alles offen und belastete sich selbst. Denn damals sagte er, der Staat habe zwei Körper: den einen schwach, mit einem kranken Haupt, den anderen kräftig, aber ohne Haupt; diesem werde, wenn man sich ihm gegenüber so verdient gemacht habe, solange er lebe, ein Haupt nicht fehlen. Der zahlreich versammelte Senat stöhnte auf, fasste jedoch keinen Beschluss, der der Ungeheuerlichkeit der Sache wirklich entsprochen hätte; denn teils fehlte es ihnen an Entschlossenheit beim Beschließen, weil sie nichts fürchteten, teils, weil ihnen alles gleichgültig war.

Tab. 1: Übersetzungsvergleich Cic. Mur. 51 (Catilinas Worte sind fett markiert).

Der Vergleich zwischen textnaher und leserorientierter Übersetzung macht deutlich, dass sich die Unterschiede nicht auf einzelne Wortentscheidungen beschränken, sondern eine grundsätzliche Verschiebung im Umgang mit interpretativer Offenheit anzeigen. Während die textnahe Übersetzung Ambiguitäten und Ellipsen sichtbar hält und damit die Entscheidungsnotwendigkeit an den Leser zurückgibt, verlagert die leserorientierte Übersetzung diese Entscheidungen in den Übersetzungsprozess selbst.

Dies zeigt sich exemplarisch an der reflexiven Festlegung bei *indicavit atque induit*. In der textnahen Übersetzung wird bei *induit* allerdings neben der naheliegenden reflexiven Lesart (‚belastete sich selbst‘) auch die Möglichkeit einer transitiven Deutung (‚belastete andere‘) angedeutet. Diese zweite Option ist jedoch nicht textlich ausgewiesen: Weder wird ein Objekt genannt noch ein Wechsel der Re-

25 Link zu Chat 1: <https://chatgpt.com/share/6978b5d2-18f4-800f-a694-47078d12ae93>. Prompt: „Übersetze den folgenden lateinischen Text möglichst textnah ins Deutsche. Vorgaben: Halte die syntaktische Struktur so transparent wie möglich. Löse Partizipialkonstruktionen nur dort auf, wo es im Deutschen zwingend nötig ist. Ergänze keine Informationen, die nicht im Text stehen. Markiere mehrdeutige oder syntaktisch schwierige Stellen in eckigen Klammern. Gib bei solchen Stellen höchstens zwei Übersetzungsalternativen an. Keine stilistische Glättung auf Kosten der Textstruktur.“

26 Link zu Chat 1: <https://chatgpt.com/share/6978b5d2-18f4-800f-a694-47078d12ae93>. Die beiden Übersetzungen wurden im selben Dialog mit ChatGPT erstellt, um die Vergleichbarkeit der Ergebnisse zu erhöhen. Ziel war es, die Übersetzungszielsetzung gezielt zu variieren, ohne weitere Einflussfaktoren wie Sitzungswechsel oder unterschiedliche Ausgangskontexte einzuführen. Prompt: „Übersetze den folgenden lateinischen Text nun als flüssige deutsche Arbeitsübersetzung. Vorgaben: Der Text soll gut lesbar und zusammenhängend sein. Triff notwendige Übersetzungsentscheidungen, ohne sie zu kommentieren. Ergänze keine historischen, biographischen oder sachlichen Informationen, die nicht im Text stehen. Verzichte auf Markierungen, Klammern oder Alternativen. Orientiere dich am Sinn des Textes, nicht an einer möglichst engen Abbildung der lateinischen Syntax.“

ferenz markiert. Sie ergibt sich allein aus der abstrakten Valenz des Verbs, nicht aus der konkreten syntaktischen oder rhetorischen Struktur der Passage. Die KI macht hier eine theoretische Möglichkeit sichtbar, die im gegebenen Kontext weder syntaktisch noch semantisch markiert ist. So wird der Unterschied zwischen textnaher und leserorientierter Übersetzung hier sehr deutlich: Denn während die textnahe Übersetzung neben der naheliegenden reflexiven Lesart auch eine transitive Deutung andeutet, hält die leserorientierte Übersetzung diese Offenheit nicht aufrecht, sondern entscheidet sich eindeutig für die reflexive Lesart (‚belastete sich selbst‘). Diese Festlegung ist kontextuell gut begründet: Sie fügt sich in Ciceros Darstellung Catilinas als bewusst provokant auftretenden Redners ein, der sich nicht verteidigt, sondern seine Haltung demonstrativ zur Schau stellt. Zugleich zeigt der Vergleich, dass die textnahe Übersetzung hier eine Möglichkeit sichtbar macht, die zwar aus der abstrakten Valenz des Verbs *induere* ableitbar ist, im konkreten Kontext der Passage jedoch keine textliche Grundlage besitzt. Der Kontrast verdeutlicht damit zweierlei: Einerseits macht die textnahe Übersetzung formale Möglichkeiten sichtbar, ohne sie zu gewichten, andererseits übernimmt die leserorientierte Übersetzung genau diese Gewichtung und bindet die Stelle stärker an den Kontext. Die interpretative Entscheidung wird so aus der Analyse in den Übersetzungsschritt selbst verlagert. Gerade an solchen Stellen zeigt sich, dass KI-Übersetzungen nicht nur Hilfsmittel, sondern eigenständige Interpretationsakte darstellen, deren Voraussetzungen und Konsequenzen sichtbar gemacht werden müssen.

Eine weitere, noch grundsätzlichere Verschiebung zeigt sich bei der Wendung *quia <omnia>*. Die spitzen Klammern markieren hier bereits eine editorische Unsicherheit: *omnia* ist nicht sicher überliefert, sondern stellt eine Ergänzung dar, mit der eine syntaktisch unvollständige Textstelle provisorisch gefüllt wird. Die textnahe Übersetzung trägt diesem Befund Rechnung, indem sie die Ellipse als solche sichtbar hält und auf eine inhaltliche Ausdeutung verzichtet. Auffällig ist, dass hier mehr als zwei mögliche Ergänzungen angeboten werden, obwohl die Instruktion eine Begrenzung vorsah.²⁷ Dies verweist auf die besondere Dichte der Unsicherheit an dieser Stelle, die offenbar nicht ohne weiteres auf wenige Alternativen reduzierbar ist. Der Befund legt nahe, dass formale Instruktionen dort an Bindungskraft verlieren, wo der Text selbst eine hohe semantische Offenheit aufweist.

Dagegen löst die leserorientierte Übersetzung die Stelle mit ‚weil ihnen alles gleichgültig war‘ auf und behandelt die hypothetische Ergänzung faktisch wie gesicherten Text. Denn diese Lesart ist zwar plausibel, jedoch im lateinischen Text nicht explizit angelegt. Der Parallelismus zu *quia nihil timebant* legt vielmehr nahe, dass die Ellipse bewusst offen gehalten ist und unterschiedliche Ergänzungen zulässt, so wie diejenige der Verfasserin, welche *quia <omnia>* mit ‚weil sie vor allem Angst hatten‘ übersetzt, da sie *timebant* elliptisch dazudenkt.²⁸ Die KI-Übersetzung entscheidet sich hier für eine psychologisierende Interpretation, die den Sinnzusammenhang klar konturiert, zugleich aber eine Deutungsmöglichkeit festschreibt, die im Ausgangstext nicht zwingend vorgegeben ist. Der Vergleich macht deutlich, dass die leserorientierte Übersetzung hier eine doppelte Entscheidung trifft: Sie ergänzt nicht nur den Sinn, sondern stabilisiert zugleich eine unsichere Textgrundlage. Die textnahe Übersetzung bewahrt demgegenüber sowohl die syntaktische als auch die textkritische Offenheit der Passage und überlässt deren Bewertung der weiteren Analyse.

Beide Entscheidungen der leserorientierten Übersetzung erhöhen die Kohärenz und Lesbarkeit des deutschen Textes, reduzieren jedoch zugleich den Spielraum für alternative Deutungen. Die Übersetzung fungiert damit nicht nur als Verständnishilfe, sondern übernimmt bereits einen Teil der interpretativen Arbeit, die in der philologischen Praxis üblicherweise erst im Anschluss an eine Arbeitsübersetzung geleistet wird und dort ausdrücklich der kritischen Prüfung und Einordnung bedarf.

Der Vergleich mit der textnahen Übersetzung macht zugleich sichtbar, warum es an diesen Stellen zu einer solchen Verschiebung kommt. Textnahe KI-Übersetzungen tendieren dazu, aus der Valenz des

27 Vgl. Prompt zum Chat oben in Anm. 25.

28 Vgl. Übersetzung in Kurowsky (2025), 15.

Verbs ableitbare Möglichkeiten sichtbar zu machen, ohne zwischen abstrakter Anschlussfähigkeit und textlich realisierter Struktur zu unterscheiden. Gerade bei ausgelassenen Objekten oder elliptischen Konstruktionen entsteht so eine über das Textsignal hinausgehende Offenheit, die nicht aus dem Text selbst resultiert, sondern aus der systematischen Berücksichtigung potenzieller Valenzen. Dabei ist zu berücksichtigen, dass die textnahe Variante durch die entsprechende Instruktion gezielt darauf angelegt ist, mögliche Deutungsoptionen sichtbar zu machen. Der entstehende Variantenraum ist daher nicht ohne Weiteres mit textimmanenter Ambiguität gleichzusetzen, sondern kann teilweise auch als Effekt der gewählten Promptstrategie verstanden werden. Diese Offenheit muss erst durch kontextuelle Gewichtung philologisch eingehegt werden.

Für die historisch-philologische Textarbeit bedeutet dies, dass der Einsatz KI-gestützter Übersetzungen – sofern sie nicht im Prompting auf Textnähe begrenzt wurden – stets mit der Frage verbunden ist, an welcher Stelle Interpretationsentscheidungen getroffen werden sollen: im Übersetzungsschritt selbst oder in der anschließenden Analyse. Gerade dieser Vergleich macht sichtbar, dass die Übersetzungszielsetzung entscheidend darüber mitbestimmt, wie offen oder geschlossen ein antiker Text der lesenden Person entgegentritt.

Die bislang beobachteten Verschiebungen betreffen einen lateinischen Text, dessen syntaktische und semantische Offenheiten klar lokalisierbar sind. Im nächsten Schritt wird derselbe Übersetzungsansatz auf eine griechische Passage angewandt (s. Tab. 2), in der Plutarch Catilinas Äußerung aufgreift und deutend verarbeitet.

So zeigen die folgenden Beobachtungen, dass sich mit dem Wechsel vom lateinischen zum griechischen Text nicht der grundlegende Mechanismus der Übersetzung ändert, sondern die Form der sprachlichen Offenheit, auf die die Übersetzung reagiert. Während sich diese bei dem Textausschnitt von Cicero vor allem in valenzbedingten Unbestimmtheiten und Ellipsen manifestiert, entsteht sie bei der Plutarch-Stelle aus semantischer Verdichtung, partizipialer Strukturierung und bewusst andeuten-der Redeweise.²⁹

Gerade dieser doppelte Vermittlungsschritt – lateinischer Ausgangstext, Deutung durch einen griechischen Autor, moderne Übersetzung – erweist sich als besonders aufschlussreich für die Frage, wie KI-Übersetzungen mit impliziten Sinnschichten, semantischer Verdichtung und interpretativer Vorprägung umgehen.

29 Vgl. zu den Besonderheiten von Altgriechisch und Latein in Bezug auf digitale Textanalyse Wittig / Kurowsky (2026) Kap. 3.

Plut. Cic. 14,6–7 (Text: Ziegler [1964])	ÜS ChatGPT (textnah) ³⁰	ÜS ChatGPT (leserorientiert) ³¹
ὁ δὲ πολλοὺς οἰόμενος εἶναι τοὺς πραγμάτων καινῶν ἐφιεμένους ἐν τῇ βουλῇ, καὶ ἅμα τοῖς συνωμόταις ἐνδεικνύμενος, ἀπεκρίνατο τῷ Κικέρωνι μανικὴν ἀπόκρισιν. „τί γάρ“ ἔφη „πράττω δεινόν, εἰ δυοῖν σωμάτων ὄντων, τοῦ μὲν ἰσχυροῦ καὶ κατεφθινηκότος, ἔχοντος δὲ κεφαλὴν, τοῦ δ' ἀκεφάλου μὲν, ἰσχυροῦ δὲ καὶ μεγάλου, τούτῳ κεφαλὴν αὐτὸς ἐπιτίθημι;“ τούτων εἰς τὴν βουλὴν καὶ τὸν δῆμον ἡγιγμένων ὑπ' αὐτοῦ, μᾶλλον ὁ Κικέρων ἔδεισε, καὶ τεθωρακισμένον αὐτὸν οἷ τε δυνατοὶ πάντες ἀπὸ τῆς οἰκίας καὶ τῶν νέων πολλοὶ κατήγον εἰς τὸ πεδῖον.	Er aber, der meinte, es gebe im Senat viele, die nach Neuerungen in den Staatsangelegenheiten verlangten, und der sich zugleich gegenüber den Mitverschwörern kenntlich machte, antwortete dem Cicero mit einer wahnsinnigen Antwort: „Was nämlich“, sagte er, „tue ich Schreckliches, wenn es zwei Körper gibt, von denen der eine schwächlich und zugrunde gegangen ist, aber einen Kopf hat, der andere aber ohne Kopf ist, dafür stark und groß, und ich diesem selbst einen Kopf aufsetze?“ [Nachdem er mit diesen Worten gegenüber Senat und Volk Andeutungen gemacht hatte / Während er mit diesen Worten gegenüber Senat und Volk Andeutungen machte], geriet Cicero in größere Furcht, und ihn, [bewaffnet / gepanzert], geleiteten sowohl alle einflussreichen Männer von seinem Haus aus als auch viele der jungen Männer hinab auf das Feld.	Er aber, der meinte, im Senat gebe es viele, die nach politischen Neuerungen verlangten, und der sich zugleich den Mitverschwörern zu erkennen gab, erwiderte Cicero mit einer geradezu wahnsinnigen Antwort: „Was“, sagte er, „tue ich denn Schreckliches, wenn es zwei Körper gibt, von denen der eine schwächlich und verfallen ist, aber einen Kopf hat, der andere dagegen ohne Kopf, dafür aber stark und groß, und ich diesem selbst einen Kopf aufsetze?“ Nachdem er damit sowohl gegenüber dem Senat als auch gegenüber dem Volk Andeutungen gemacht hatte, geriet Cicero in noch größere Furcht, und alle einflussreichen Männer sowie viele der jungen Leute begleiteten ihn, bewaffnet, von seinem Haus hinab auf das Marsfeld.

Tab. 2: Übersetzungsvergleich Plut. Cic. 14,6–7 (Catilinas Worte sind fett markiert).

Für die altgriechische Passage bleiben die Unterschiede zwischen textnaher und leserorientierter Übersetzung insgesamt begrenzt. Die beiden Partizipien οἰόμενος und ἐνδεικνύμενος werden in beiden Fassungen attributiv und koordiniert wiedergegeben: Eine funktionale Differenzierung (etwa kausal oder final) wird nicht vorgenommen. Somit bleibt die syntaktische Gleichrangigkeit der Partizipien gewahrt. Abweichungen beschränken sich hierbei auf lexikalische Varianten („kenntlich machen“ / „zu erkennen geben“), ohne dass sich die semantische Struktur im Deutschen verändert.

Auch die εἰ-Periode innerhalb der Rede wird in beiden Übersetzungen formal adäquat abgebildet. Die rhetorische Frage (τί γάρ πρᾶττω δεινόν, εἰ ...) bleibt als Konditionalgefüge erhalten: Protasis und Apodosis werden klar unterschieden. Die Genitivi absoluti werden in beiden Versionen parataktisch zum Prädikat entfaltet, ohne zusätzliche Explikation der metaphorischen Referenz. Eine Umstrukturierung der Bedingungslogik ist nicht festzustellen. Unterschiede betreffen lediglich idiomatische Glättungen.

30 Link zu Chat 2: <https://chatgpt.com/share/697b9b8e-acc0-800f-b243-0322b83284a9>. Prompt: „Übersetze den folgenden altgriechischen Text möglichst textnah ins Deutsche. Vorgaben: Halte die syntaktische Struktur so transparent wie möglich. Löse Partizipialkonstruktionen nur dort auf, wo es im Deutschen zwingend nötig ist. Ergänze keine Informationen, die nicht im Text stehen. Markiere mehrdeutige oder syntaktisch schwierige Stellen in eckigen Klammern. Gib bei solchen Stellen höchstens zwei Übersetzungsalternativen an. Keine stilistische Glättung auf Kosten der Textstruktur.“

31 Link zum Chat 2: <https://chatgpt.com/share/697b9b8e-acc0-800f-b243-0322b83284a9>. Prompt: „Übersetze den folgenden altgriechischen Text nun als flüssige deutsche Arbeitsübersetzung. Vorgaben: Der Text soll gut lesbar und zusammenhängend sein. Triff notwendige Übersetzungsentscheidungen, ohne sie zu kommentieren. Ergänze keine historischen, biographischen oder sachlichen Informationen, die nicht im Text stehen. Verzichte auf Markierungen, Klammern oder Alternativen. Orientiere dich am Sinn des Textes, nicht an einer möglichst engen Abbildung der altgriechischen Syntax.“

Ein deutlicherer Unterschied zeigt sich bei πεδίων: Während die textnahe Übersetzung den lexikalischen Befund mit ‚Feld‘ beibehält, entscheidet sich die leserorientierte Fassung für die historisch konventionalisierte Konkretisierung ‚Marsfeld‘. Diese Spezifizierung ist im griechischen Wortlaut nicht explizit markiert, sondern setzt historisches Kontextwissen voraus. Hier tritt eine kontextuelle Festlegung an die Stelle einer offeneren Wiedergabe, wodurch der Text eine konkrete topographische Verankerung erhält. Damit verschiebt sich der Schwerpunkt von der politischen Andeutung eines offenen Handlungsraums zur szenischen Ausgestaltung: Was bei Plutarch als politisch aufgeladener Raum zwischen institutionellen Akteuren angelegt ist, erscheint nun als klar geführte Abfolge von Ort, Handlung und beteiligten Gruppen. Die Übersetzung schließt so einen Deutungsraum, der im Original bewusst offen gehalten ist.³²

Insgesamt zeigt der Vergleich in dieser Passage nur geringe Differenzen zwischen beiden Fassungen. Die Abweichungen beschränken sich auf punktuelle idiomatische Glättungen, eine temporale Festlegung des Genitivus absolutus (τούτων ... ἡνιγμένων) sowie eine kontextuelle Konkretisierung bei πεδίων. Strukturelle oder semantische Verschiebungen sind nicht festzustellen.

Die Analyse der Passage bei Plutarch zeigt, dass sich die bei Cicero beobachteten Effekte nicht einfach wiederholen, sondern unter veränderten sprachlichen Voraussetzungen verschieben. Während im lateinischen Text vor allem valenzbedingte Offenheiten und Ellipsen den Interpretationsspielraum bestimmen, entsteht dieser im Griechischen zusätzlich durch partizipiale Verdichtung, hypothetische Rahmung und bewusst andeutende Redeweise. In beiden Fällen reagiert die leserorientierte Übersetzung darauf mit einer stärkeren strukturellen und semantischen Fixierung, die Kohärenz und Lesbarkeit erhöht, zugleich aber implizite Spannungen reduziert.

Übersetzen erweist sich damit auch hier nicht als rein vermittelnder Akt, sondern als Eingriff in das Verhältnis von Gesagtem und Gemeintem, von Andeutung und Explikation.³³ Die leserorientierte Übersetzung macht diese Verschiebung im Vergleich zur textnahen besonders deutlich, weil sie dazu tendiert, semantische Leerstellen zu schließen und narrative Zusammenhänge explizit zu ordnen. Gerade darin liegt ihr analytischer Wert: Nicht weil sie entscheidet, wie der Text zu verstehen ist, sondern weil sie sichtbar macht, wo Entscheidung überhaupt ansetzt.

3.2 Übersetzungen von Latein und Altgriechisch moderner Autoren

Die bisherigen Beobachtungen haben gezeigt, dass ChatGPT-Übersetzungen antiker Texte nicht als fertige Interpretationen, wohl aber als generierte Zieltexte in einen philologischen Prüfprozess integriert werden können. Gerade bei vielfach edierten, kommentierten und übersetzten Autoren wie Cicero und Plutarch, zu denen umfangreiche Überlieferungs- und Forschungstraditionen vorliegen, liefern sie stabile Ausgangspunkte für weiterführende philologische Arbeit. Zugleich ist jedoch zu berücksichtigen, dass für solche Texte aufgrund ihrer breiten Überlieferung und Rezeption von einem besonders reichhaltigen Trainingsumfeld auszugehen ist.

Um die Übersetzungsleistung der KI unter Bedingungen zu prüfen, bei denen ein Rückgriff auf etablierte Vorlagen weitgehend ausgeschlossen ist, werden im Folgenden Texte herangezogen, die nicht zur antiken Texttradition gehören und für die das Training mit existierenden Übersetzungen kaum in vergleichbarer Form möglich gewesen sein dürfte. Denn gerade weil sich bei antiken Texten nicht immer trennen lässt, ob Übersetzungsentscheidungen auf textinterner Analyse oder auf implizitem Vorwissen beruhen, ist der zugrunde liegende Wissensstand des verwendeten KI-Systems mitzudenken.

32 Wie Pelling (2011 = 2002), 268 gezeigt hat, operiert Plutarchs Erzählen häufig vor dem Hintergrund möglicher Alternativen und Deutungsvarianten, wodurch narrative Autorität gerade nicht durch Eindeutigkeit, sondern durch kontrollierte Mehrstimmigkeit erzeugt wird.

33 Zum Übersetzungsbegriff s. Kurowsky (2025), 1–2.

Untersucht werden die Vorworte zweier Dissertationen moderner Autoren – zur Vergleichbarkeit dieselben Texte wie in der früheren Publikation der Autorin –,³⁴ die ihre Arbeiten Ende des 19. bzw. Anfang des 20. Jahrhunderts bewusst in Latein bzw. Altgriechisch verfasst haben.³⁵ Die Auswahl von Vorworten als Untersuchungsgrundlage erfolgt nicht zufällig. Vorworte zeichnen sich durch einen hohen Grad an autorseitiger Kontrolle und argumentative Klarheit aus, da sie die Zielsetzung und methodische Anlage einer Arbeit explizit formulieren. Zugleich weisen sie eine komplexe, aber vergleichsweise kontrollierte Syntax auf und verzichten weitgehend auf stark stilisierte oder literarisch verdichtete Ausdrucksformen.

Die beiden Texte waren zum Zeitpunkt der Durchführung der ersten Übersetzungstests nicht frei online zugänglich und wurden daher unter kontrollierten Bedingungen für die Untersuchung aufbereitet.³⁶ Im Anschluss an das bisherige Vorgehen wird zunächst die lateinische Dissertation betrachtet (s. Tab. 3).

Die lateinische Dissertation von Kolfhaus stellt an die KI andere Anforderungen als die zuvor untersuchten antiken Texte. Der Text ist syntaktisch klar strukturiert, argumentativ linear geführt und weitgehend frei von rhetorischer Verdichtung. Zugleich weist er eine hohe begriffliche Präzision auf, insbesondere in der Verwendung abstrakter Substantive und verbaler Nominalkonstruktionen. Übersetzungsentscheidungen betreffen hier daher weniger die Auflösung semantischer Mehrdeutigkeiten als die Gewichtung funktional naher Ausdrucksmöglichkeiten.

34 Vgl. Kurowsky (2025), 24–26.

35 Klassische Anleihen im Ausdruck sind bei solchen Texten grundsätzlich möglich. Sie sind jedoch als moderne wissenschaftliche Texte zu verstehen und nicht als Teil der antiken Überlieferungstradition.

36 Die Texte sind inzwischen durch die erste Version dieses Beitrages online zugänglich. Zum Zeitpunkt der durchgeführten Übersetzungstests waren sie jedoch weder frei verfügbar noch Teil gängiger digitaler Textkorpora. Unabhängig davon zielt die Analyse nicht auf den Nachweis vollständiger Trainingsabstinenz, sondern auf die Übersetzungsleistung unter Bedingungen, bei denen ein unmittelbarer Rückgriff auf etablierte Übersetzungs- oder Kommentierungstraditionen weitgehend ausgeschlossen ist.

Kolfhaus Vorwort	ÜS ChatGPT (textnah) ³⁷	ÜS ChatGPT (leserorientiert) ³⁸
<i>Huius disputationis erit quaerere, utrum liber inter Plutarchi scripta quae dicuntur moralia traditus, qui inscribitur ‚de communibus notitiis‘, iure Plutarchi nomen prae se ferat an subditus sit. Quae quaestio cum magna ex parte ad illam, quo studio hiatus evitandi Plutarchus usus sit, redeat, primo capite de hoc studio iis quae ab aliis de hac re inventa sunt examinatis disserendum est.</i>	Ziel dieser Erörterung wird es sein zu untersuchen, ob das unter den sogenannten moralischen Schriften Plutarchs überlieferte Buch, das den Titel <i>De communibus notitiis</i> trägt, mit Recht den Namen Plutarchs vor sich her trägt oder [untergeschoben ist / einem anderen untergeschoben worden ist]. Da diese Frage zu einem großen Teil auf jene zurückführt, [mit welchem Bestreben / aus welchem methodischen Bemühen] Plutarch beim Vermeiden von Hiaten verfahren ist, muss im ersten Kapitel über dieses Bestreben gehandelt werden, nachdem [das, was von anderen zu dieser Sache ermittelt worden ist / die von anderen zu dieser Sache gewonnenen Ergebnisse] geprüft worden sind.	Ziel dieser Untersuchung ist es zu klären, ob das unter den sogenannten moralischen Schriften Plutarchs überlieferte Buch mit dem Titel <i>De communibus notitiis</i> zu Recht den Namen Plutarchs trägt oder ob es untergeschoben ist. Da diese Frage zu einem großen Teil auf das Bemühen zurückführt, mit dem Plutarch die Vermeidung von Hiaten betrieben hat, ist im ersten Kapitel dieses Bemühen zu erörtern, nachdem die hierzu von anderen gewonnenen Ergebnisse geprüft worden sind.

Tab. 3: Übersetzungsvergleich eines modernen wissenschaftlichen Textes in lateinischer Sprache.

In der textnahen Übersetzung bleibt dieser Entscheidungsspielraum sichtbar. So werden bei *subditus sit* sowohl eine passive (‚untergeschoben ist‘) als auch eine stärker akteurbezogene Lesart (‚einem anderen untergeschoben worden ist‘) offengehalten, ebenso bei der Bestimmung von *studio* im Zusammenhang mit der Hiatermeidung. Ähnlich wie bei der textnahen Übersetzung für Cic. Mur. 51 für *induit* (s.o.) ist die zweite Variante ‚einem anderen ...‘ nicht inhaltlich festzumachen: Das Objekt zu *subditus sit* kann nur Plutarch selbst sein. Möglicherweise hatte die Anweisung zum Markieren von Mehrdeutigkeiten an diesen Stellen für ChatGPT eine höhere Priorität als die Herstellung des Sinnzusammenhangs.³⁹

Demgegenüber entscheidet sich die leserorientierte Übersetzung konsequent für eine funktionale Glättung: Die fraglichen Stellen werden eindeutig festgelegt und in einen flüssigen argumentativen Zusammenhang überführt. Diese Festlegungen sind für das Textverständnis unproblematisch, betreffen jedoch genau jene begrifflichen Nuancen, die für eine philologische Einordnung des methodischen Selbstverständnisses des Autors relevant sind. Damit zeigt sich, dass auch bei syntaktisch kontrollierten und argumentativ linearen Texten Übersetzungen nicht neutral bleiben, sondern durch begriffliche Festlegung interpretative Akzente setzen. Dadurch gewinnt der Text an Lesbarkeit, verliert jedoch die

37 Link zu Chat 3: <https://chatgpt.com/share/697ba7e3-ffac-800f-9e59-cb8778e89934>. Prompt: „Übersetze den folgenden lateinischen Text möglichst textnah ins Deutsche. Vorgaben: Halte die syntaktische Struktur so transparent wie möglich. Löse Partizipialkonstruktionen nur dort auf, wo es im Deutschen zwingend nötig ist. Ergänze keine Informationen, die nicht im Text stehen. Markiere mehrdeutige oder syntaktisch schwierige Stellen in eckigen Klammern. Gib bei solchen Stellen höchstens zwei Übersetzungsalternativen an. Keine stilistische Glättung auf Kosten der Textstruktur.“

38 Link zu Chat 3: <https://chatgpt.com/share/697ba7e3-ffac-800f-9e59-cb8778e89934>. Prompt: „Übersetze den folgenden lateinischen Text nun als flüssige deutsche Arbeitsübersetzung. Vorgaben: Der Text soll gut lesbar und zusammenhängend sein. Triff notwendige Übersetzungsentscheidungen, ohne sie zu kommentieren. Ergänze keine historischen, biographischen oder sachlichen Informationen, die nicht im Text stehen. Verzichte auf Markierungen, Klammern oder Alternativen. Orientiere dich am Sinn des Textes, nicht an einer möglichst engen Abbildung der lateinischen Syntax.“

39 Die beobachtete Tendenz, formal plausible, semantisch nicht kontextgestützte Alternativen zu generieren, korrespondiert mit Befunden zur sensitivitätsabhängigen Outputvariation großer Sprachmodelle im Umgang mit unterschiedlichen Instruktionszielen, wie sie etwa von Liu et al. (2023) beschrieben werden.

feinen Differenzierungen, die für eine philologische Analyse des methodischen Selbstverständnisses des Autors von Bedeutung sein können.

Ob sich diese Verschiebung auch bei einem wissenschaftlichen Text in altgriechischer Sprache beobachten lässt, wurde im nächsten Schritt am Vorwort einer in Altgriechisch verfassten Dissertation untersucht (s. Tab. 4). Dabei stellt sich erneut die Frage, wie ChatGPT-Übersetzungen mit begrifflicher Präzision und syntaktischer Kontrolle umgehen, wenn keine etablierte Übersetzungstradition vorliegt.

Dassaritis Vorwort	ÜS ChatGPT (textnah) ⁴⁰	ÜS ChatGPT (leserorientiert) ⁴¹
Πλούταρχος ὁ Χαιρωνεύς ἐστὶν ὁ συγγραφεὺς ἐκεῖνος, ὃς ὑπολείπεται μὲν κατὰ τὴν λέξιν τῶν δοκίμων καὶ ἀπαραμίλλων ἐκείνων τοῦ λόγου τεχνιτῶν, τοῦθ' ὅπερ καὶ αὐτὸς συναισθανόμενος ὁμολογεῖ μὲν καὶ κατὰ τὰς αἰθερίους καὶ δαιμονίους τοῦ πνεύματος πτήσεις, οὐ μὴν ἀλλὰ κέκτηται μεγάλην ἱστορικὴν, ἠθικὴν τε καὶ παιδαγωγικὴν, πολιτικὴν καὶ φιλολογικὴν σημασίαν, ἔτι δὲ καὶ φιλοσοφικὴν ἐν τοῖς μεταγενεστέροις τῆς ἀρχαιότητος χρόνοις.	Plutarch von Chaironeia ist jener Schriftsteller, der zwar sowohl in der Wortwahl gegenüber jenen bewährten und unvergleichlichen Handwerkern der Rede zurückbleibt, [dies eben, was / gerade das, was] er auch selbst, sich dessen bewusst, eingesteht, und zwar auch im Hinblick auf die ätherischen und daimonischen Flüge des Geistes; nichtsdestoweniger aber besitzt er große historische, ethische und pädagogische, politische und philologische Bedeutung, dazu noch auch eine philosophische in den späteren Zeiten der Antike.	Plutarch von Chaironeia ist jener Schriftsteller, der zwar in der sprachlichen Ausgestaltung hinter jenen bewährten und unvergleichlichen Meistern der Rede zurückbleibt – was er selbst, dessen sich bewusst, auch einräumt, ebenso im Hinblick auf die ätherischen und daimonischen Höhenflüge des Geistes –, der jedoch nichtsdestoweniger eine große historische, ethische und pädagogische, politische und philologische Bedeutung besitzt und darüber hinaus auch eine philosophische in den späteren Zeiten der Antike.

Tab. 4: Übersetzungsvergleich eines modernen wissenschaftlichen Textes in altgriechischer Sprache.

Im Vorwort der Dissertation von Dassaritis steht nicht die argumentative Linearität, sondern eine bewusst spannungsreiche Selbstcharakterisierung Plutarchs im Zentrum. Plutarch wird als jemand vorgestellt, der im Bereich der λέξις hinter den ‚bewährten und unvergleichlichen‘ Redekünstlern zurückbleibt, dieses Defizit jedoch selbst reflektiert und einräumt (συναισθανόμενος, ὁμολογεῖ). Zugleich wird diese Einschränkung durch die Wendung οὐ μὴν ἀλλὰ ausdrücklich relativiert und in eine umfassende positive Würdigung überführt, die Plutarchs Bedeutung in historischen, ethischen, pädagogischen, politischen und philosophischen Zusammenhängen betont.

Die textnahe Übersetzung lässt diese Spannung zwischen Selbstbeschränkung und Aufwertung bewusst bestehen und hält insbesondere die doppelte Perspektive aufrecht, die sich aus der Kombination von sprachlicher Zurücknahme und geistigen ‚Höhenflügen‘ (αἰθερίαὶ καὶ δαιμόνιοι πτήσεις) ergibt. In der leserorientierten Übersetzung werden diese Ebenen stärker aufeinander bezogen und in eine klarere argumentative Bewegung überführt: Die konzessive Struktur tritt deutlicher hervor, zugleich verliert die Selbstreflexion des Autors einen Teil ihrer Schwebelast. Sichtbar wird damit weniger ein Bedeutungsverlust als eine Verschiebung der Akzentsetzung – von einer ambivalenten Selbstverortung hin zu einer eindeutig geführten Bewertung.

40 Link zu Chat 4: <https://chatgpt.com/share/697bb2cb-f898-800f-8deb-02273080f4d2>. Prompt: „Übersetze den folgenden altgriechischen Text möglichst textnah ins Deutsche. Vorgaben: Halte die syntaktische Struktur so transparent wie möglich. Löse Partizipialkonstruktionen nur dort auf, wo es im Deutschen zwingend nötig ist. Ergänze keine Informationen, die nicht im Text stehen. Markiere mehrdeutige oder syntaktisch schwierige Stellen in eckigen Klammern. Gib bei solchen Stellen höchstens zwei Übersetzungsalternativen an. Keine stilistische Glättung auf Kosten der Textstruktur.“

41 Link zu Chat 4: <https://chatgpt.com/share/697bb2cb-f898-800f-8deb-02273080f4d2>. Prompt: „Übersetze den folgenden lateinischen Text nun als flüssige deutsche Arbeitsübersetzung. Vorgaben: Der Text soll gut lesbar und zusammenhängend sein. Triff notwendige Übersetzungsentscheidungen, ohne sie zu kommentieren. Ergänze keine historischen, biographischen oder sachlichen Informationen, die nicht im Text stehen. Verzichte auf Markierungen, Klammern oder Alternativen. Orientiere dich am Sinn des Textes, nicht an einer möglichst engen Abbildung der lateinischen Syntax.“

Somit bestätigt und präzisiert die Untersuchung der Dissertationstexte die zuvor gewonnenen Befunde. Auch unter Bedingungen, bei denen keine etablierte Übersetzungs- oder Kommentierungstradition vorausgesetzt werden kann, lassen sich ChatGPT-generierte Übersetzungen in einen philologischen Analyseprozess einbinden, sofern ihre impliziten Entscheidungen systematisch offengelegt werden. Zugleich können sie als erzeugte Zieltexte analysiert werden, deren interpretative Strukturierung selbst zum Gegenstand der Untersuchung wird. Überdies zeigt sich, dass sich der Charakter der Übersetzungsentscheidungen mit der Textsorte und der jeweiligen sprachlichen Ausgestaltung verschiebt: Während bei antiken Texten vor allem implizite Spannungen, rhetorische Verdichtung und narrative Offenheit betroffen sind, treten bei den modernen wissenschaftlichen Texten Fragen der begrifflichen Gewichtung, syntaktischen Hierarchisierung und argumentativen Akzentsetzung in den Vordergrund.

In beiden Fällen reagieren leserorientierte Übersetzungen mit einer Tendenz zur Explikation und strukturellen Ordnung, die Lesbarkeit und argumentative Klarheit erhöht, zugleich aber interpretative Vorentscheidungen trifft. Jede Übersetzung – auch eine KI-generierte – ist damit als eine perspektivische Repräsentation zu verstehen, die notwendigerweise Informationsverluste und -gewinne produziert.⁴² Deutlich wird in dieser Untersuchung weniger ein pauschaler Bedeutungsverlust als eine Verschiebung der Perspektive: Ambivalenzen, Selbstrelativierungen und bewusst offen gehaltene Spannungen des Ausgangstextes werden zugunsten einer klar geführten Bewertung zurückgenommen. Gerade im Vergleich der unterschiedlichen Texttypen wird so sichtbar, dass die analytische Relevanz von ChatGPT-Übersetzungen weniger aus ihrer Nähe zu kanonischem Vorwissen erwächst als aus der Art und Weise, wie sie sprachliche Offenheit verarbeiten.

Unter den Bedingungen moderner wissenschaftlicher Texte, für die weder eine umfangreiche Übersetzungs- und Kommentierungstradition noch ein naheliegender Rückgriff auf etabliertes Vorwissen vorausgesetzt werden kann, zeigen sich also dieselben Übersetzungsmuster wie zuvor. Die leserorientierten Übersetzungen reagieren erneut mit Explikation, struktureller Ordnung und begrifflicher Festlegung auf sprachliche Offenheit, wenn auch auf anderen Ebenen als in den antiken Texten. Somit ergibt sich die analytische Relevanz ChatGPT-gestützter Übersetzungen weniger aus der Überlieferungsdichte oder einer möglichen Trainingsrepräsentanz der Texte als aus der systematischen Art und Weise, in der sie unterschiedliche Formen sprachlicher Offenheit verarbeiten.⁴³

Die beobachteten Unterschiede lassen sich dabei weniger als Qualitätsurteile denn als unterschiedliche Weisen des Umgangs mit textueller Offenheit verstehen, die im Übersetzungsprozess entweder bewusst sichtbar gehalten oder zugunsten einer stabilen Sinnführung geschlossen wird.

3.3 Zur Modus- und Kontextabhängigkeit der Übersetzungsergebnisse

Im Anschluss an die bisherigen Analysen stellt sich die methodische Frage, inwieweit die beobachteten Effekte von der konkreten Interaktionskonfiguration abhängen. Die zuerst untersuchten Übersetzungen wurden im Modus ‚Auto‘ unter aktivierter Personalisierung und mit aktivierter Erinnerungsfunktion erzeugt.⁴⁴ Da große Sprachmodelle kontextsensitiv operieren, war zu prüfen, ob die diagnostizierte Verschiebung interpretativer Entscheidungen in den Übersetzungsschritt selbst ein Artefakt dieser spezifischen Einstellung darstellt oder als strukturelle Folge der jeweiligen Instruktionslogik zu verstehen ist.

Zu diesem Zweck wurden sämtliche Testpassagen unter neutralen Bedingungen erneut generiert, nämlich in den Modi ‚Instant‘ und ‚Thinking‘, jeweils ohne Personalisierung und ohne berücksichtigte Er-

42 Vgl. hierzu Sahle (2013a), 135–137.

43 Vgl. zu sprachlicher Offenheit und Interpretation bei Übersetzungen Pym (2023), 114–115.

44 Unter ‚Personalisierung‘ wird hier die systemseitige Berücksichtigung userspezifischer Präferenzen und Einstellungen verstanden. ‚Aktivierte Erinnerung‘ bezeichnet die Einbeziehung vorangegangener Interaktionen im selben Arbeitskontext.

innerung. Texte und Aufgabenformulierung blieben dabei unverändert. Die Differenzierung zwischen textnaher und leserorientierter Übersetzung wurde grundsätzlich beibehalten: In den Vergleichstests wurde die im Initialsetting verwendete Bezeichnung ‚Arbeitsübersetzung‘ terminologisch präzisiert als ‚leserorientierte Übersetzung‘ formuliert. Inhaltlich blieb der funktionale Auftrag unverändert (flüssige, kohärente, entscheidungsfixierende Übertragung ohne Markierung von Alternativen), so dass die Vergleichbarkeit der Ergebnisse gewahrt bleibt.

Bei diesem Variationstest traten modusspezifische Unterschiede zutage, die nicht auf die Instruktionsstruktur zurückgeführt werden können. Ein Beispiel bietet die griechische Passage mit dem Wort *πεδίων*.⁴⁵ Während der ‚Thinking‘-Modus diese Stelle konsistent als ‚Marsfeld‘ konkretisiert, verbleibt der ‚Instant‘-Modus bei der allgemeineren Wiedergabe ‚Feld‘. Bemerkenswert ist, dass diese Differenz innerhalb des jeweiligen Modus stabil bleibt: Sie zeigt sich sowohl in der textnahen als auch in der leserorientierten Variante. Es handelt sich somit nicht um eine Veränderung der grundlegenden Strukturierung von Offenheit und Kohärenz, sondern um eine Variation im Grad semantischer Spezifizierung. In dieser Testkonfiguration zeigt sich der modusspezifische Unterschied auf der Ebene der Wortwahl, nicht jedoch in der funktionalen Struktur der Übersetzung.

Ein zweiter, in methodischer Hinsicht besonders aufschlussreicher Fall betrifft die lateinische Ellipse *quia <omnia>*.⁴⁶ In den textnahen Varianten bleibt die Stelle erwartungsgemäß offen, da die Instruktion explizit vorsieht, problematische oder unvollständige Passagen nicht aufzulösen. In den leserorientierten Fassungen zeigt sich jedoch eine modusspezifische Divergenz: Während einzelne Konfigurationen die Ellipse semantisch auffüllen und etwa eine Formulierung wie ‚weil ihnen alles gleichgültig war‘ wählen, belässt der ‚Thinking‘-Modus die Stelle selbst in der leserorientierten Übersetzung als offenes ‚weil alles ...‘. Hier wird also trotz kohärenzorientierter Rahmung auf eine semantische Ergänzung verzichtet.

Auch bei den wissenschaftlichen Textausschnitten zeigt sich in den Variationstests zunächst die Stabilität der instruktionsgebundenen Differenz. Unabhängig vom gewählten Modus (‚Auto‘ mit Personalisierung, ‚Thinking‘ oder ‚Instant‘ ohne Personalisierung) führen textnahe Vorgaben dazu, dass grammatische und semantische Spielräume explizit markiert werden. So erscheint etwa bei *subditus sit* in sämtlichen textnahen Varianten die Alternative ‚untergeschoben ist / untergeschoben worden ist‘,⁴⁷ wodurch der grammatische Deutungsraum sichtbar bleibt. In den leserorientierten Fassungen hingegen wird diese Alternative in allen Modi stabilisiert (‚ob es unecht ist‘ bzw. ‚ob es ihm untergeschoben wurde‘). Die Differenz ist hier eindeutig instruktionsgebunden, nicht modusspezifisch.

Modusspezifische Unterschiede treten jedoch auch in diesen wissenschaftlichen Passagen zutage, wenn auch weniger auffällig als bei *πεδίων*. So zeigt sich etwa bei der Wendung *quo studio hiatus evitandi* eine leichte Verschiebung der terminologischen Akzentuierung: Während einzelne Konfigurationen (insbesondere der ‚Thinking‘-Modus) zu stärker konzeptuell gefassten Lösungen wie ‚mit welcher Konsequenz‘ tendieren, verbleiben andere näher an der wörtlichen Semantik (‚mit welchem Eifer‘ / ‚mit welchem Bemühen‘).⁴⁸ Diese Differenzen betreffen nicht die argumentative Struktur der Passage, sondern den Grad begrifflicher Präzisierung.

45 Vgl. zu den Chats für die Plutarch-Stelle mit Thinking, neutral: <https://chatgpt.com/share/698de504-da20-800f-8f3d-7706a2fa007> (Chat 6), mit Instant, neutral: <https://chatgpt.com/share/698e1c05-1518-800f-aa41-bec369ad5c45> (Chat 10).

46 Vgl. zu den Chats für die Cicero-Stelle mit Thinking, neutral: <https://chatgpt.com/share/698de693-e2c0-800f-af80-d10e0b66edf3> (Chat 5), mit Instant, neutral: <https://chatgpt.com/share/698e194d-e44c-800f-8af7-a7c1e35e7996> (Chat 9).

47 Vgl. zu den Chats für die Kolfhaus-Stelle mit Thinking, neutral: <https://chatgpt.com/share/698de3ff-6648-800f-9de4-9e7358170239> (Chat 7), mit Instant, neutral: <https://chatgpt.com/share/698e1d07-ef40-800f-bb98-12cdedf05056> (Chat 11).

48 Links zur Kolfhaus-Stelle wie in Anm. 47.

Ein vergleichbarer Effekt zeigt sich im griechischen Beispiel τοῦ λόγου τεχνιτῶν. Die textnahen Fassungen markieren ‚Künstler des Wortes / Meister der Rede‘ als Alternative, unabhängig vom Modus.⁴⁹ In den leserorientierten Varianten wird jeweils eine Entscheidung getroffen. Die konkrete Wahl variiert leicht zwischen den Modi, doch die Grundstruktur bleibt identisch. Auch hier ist die Offenhaltung oder Stabilisierung instruktionsabhängig, während der Modus lediglich die terminologische Nuance moduliert.

Im Vergleich zu den antiken narrativen Passagen – etwa bei πεδίων oder der Ellipse *quia <omnia>* – zeigen sich die modusspezifischen Unterschiede in den wissenschaftlichen Texten weniger in inhaltlichen Festlegungen als in der sprachlichen Akzentuierung. Während bei πεδίων eine topographische Konkretisierung (‚Marsfeld‘ vs. ‚Feld‘) erfolgt und bei *quia <omnia>* die Frage der elliptischen Schließung divergiert, betreffen die Unterschiede in den wissenschaftlichen Passagen primär die begriffliche Gewichtung und stilistische Präzisierung.

Zu berücksichtigen ist ferner der Einfluss von Personalisierung und gespeicherter Erinnerung im Initialsetting. Während die neutralen Tests ohne aktivierte Erinnerungsfunktion durchgeführt wurden, operierte der Ausgangsmodus ‚Auto‘ unter Einbeziehung früherer Interaktionen und individueller Präferenzen. Die Vergleichstests zeigen jedoch keine strukturelle Verschiebung der grundlegenden Differenz zwischen textnaher und leserorientierter Rahmung. Personalisierung und Erinnerung wirken demnach nicht als eigenständige Interpretationsinstanz, sondern allenfalls als Modulator der Ausprägung bereits instruktionsgebundener Tendenzen. Sie können die Intensität semantischer Konkretisierung beeinflussen, nicht jedoch die funktionale Organisation von Offenheit und Stabilisierung.

Der Variationstest ergibt somit kein Bild völliger Identität der Ergebnisse, wohl aber eine stabile zweistufige Struktur: Die funktionale Rahmung der Aufgabe legt fest, ob Mehrdeutigkeit explizit markiert oder in den Zieltext integriert wird. Der gewählte Modus beeinflusst hingegen, wie stark Bedeutungsräume konkretisiert, kontextualisiert oder geschlossen werden. Interpretative Entscheidungen entstehen damit nicht ausschließlich im Modell und auch nicht ausschließlich in der Instruktion, sondern im Zusammenspiel beider Ebenen. Das bedeutet, dass weder Moduswahl noch Personalisierung diese zweistufige Struktur aufheben, sondern lediglich deren Ausprägung modulieren.

Für die historisch-philologische Arbeit bedeutet dies zweierlei. Erstens ist die explizite Formulierung des Übersetzungsziels die zentrale epistemische Variable: Sie strukturiert, welche Formen von Offenheit im Zieltext sichtbar bleiben. Zweitens kann die Wahl des Modus die Art der semantischen Spezifizierung beeinflussen und damit – etwa im Umgang mit Ellipsen oder topographischen Bezügen – praktische Relevanz für die Nutzung der erzeugten Übersetzungen gewinnen. Somit relativiert der Variationstest weder die zuvor gewonnenen Befunde noch ersetzt sie diese durch eine Modustheorie, sondern differenziert die Ebenen, auf denen Interpretationsentscheidungen im Prozess ChatGPT-gestützter Übersetzung wirksam werden.

49 Vgl. zu den Chats für die Dassaritis-Stelle mit Thinking, neutral: <https://chatgpt.com/share/698de369-b750-800f-8371-3c4ca2105a41> (Chat 8), mit Instant, neutral: <https://chatgpt.com/share/698e1e2e-a6a8-800f-b19e-30fc8384d646> (Chat 12).

4. Fazit

Die vorliegenden Untersuchungen haben gezeigt, dass ChatGPT-Übersetzungen weder als bloße Hilfsmittel noch als eigenständige Interpretationsinstanzen zu begreifen sind. Ihr analytischer Wert liegt vielmehr in der spezifischen Art und Weise, wie sie mit sprachlicher Offenheit umgehen: indem sie implizite Relationen explizieren, syntaktische Spannungen ordnen und semantische Leerstellen häufig füllen oder – je nach Konfiguration – in unterschiedlichem Maß offenhalten. Gerade darin liegt ihr heuristischer Wert: ChatGPT-Übersetzungen machen nicht nur sichtbar, wo historisch-philologische Interpretation ansetzt, sondern verlagern diese Entscheidungen bereits in den Übersetzungsschritt selbst und fordern dadurch eine bewusste methodische Reflexion.

In dieser Perspektive lässt sich der Unterschied zwischen textnaher und leserorientierter Übersetzung auch als Umgang mit textueller Offenheit beschreiben: Während textnahe Übersetzungen Mehrdeutigkeiten, Ellipsen und implizite Relationen sichtbar halten, schließen leserorientierte Übersetzungen diese Offenheit zugunsten kohärenter Sinnaktualisierung – ein Vorgang, der sich vergleichsweise mit Ecos Begriff des offenen Kunstwerks fassen lässt.⁵⁰

Der Vergleich antiker und moderner Texte hat zugleich deutlich gemacht, dass diese Effekte nicht primär von der Überlieferungs- und Rezeptionsdichte eines Textes oder von vermuteter Trainingsnähe abhängen, sondern von textsortenspezifischen Strukturen und sprachlichen Anforderungen. ChatGPT-Übersetzungen reagieren auf diese Anforderungen – insbesondere unter kohärenzorientierter Rahmung – mit Strategien der Glättung und Fixierung.⁵¹ Leserorientierte Übersetzungen dieser Art erzeugen erhöhte Kohärenz, verlagern jedoch interpretative Entscheidungen in den Übersetzungsschritt selbst, die im weiteren Analyseprozess bewusst kenntlich zu machen sind. Daraus ergibt sich für die historisch-philologische Arbeit die Aufgabe, diese im Übersetzungsschritt getroffenen Entscheidungen der KI erneut sichtbar und überprüfbar zu machen.

Während der erste Beitrag der Verfasserin den Schwerpunkt auf die Funktionsweise des Modells, auf trainingsbedingte Asymmetrien zwischen Zielsprachen sowie auf die grundsätzliche Brauchbarkeit von ChatGPT-Übersetzungen legt, verschiebt sich der Fokus der vorliegenden Untersuchung. Im Zentrum steht nun weniger die Frage, ob ChatGPT-Übersetzungen funktionieren, als vielmehr, wie sie Übersetzungsentscheidungen strukturieren und an welchen Stellen sie interpretative Akzente setzen. Die Reloaded-Analyse stellt keine Korrektur, sondern eine Präzisierung des zuvor gewonnenen Befunds dar.

So erweist sich der Einsatz von ChatGPT im historisch-philologischen Übersetzungskontext weniger als Ersatz menschlicher Deutung denn als heuristisches Instrument: nicht um Bedeutung festzuschreiben, sondern um sichtbar zu machen, an welchen Stellen Übersetzungs- und Interpretationsentscheidungen bereits im Übersetzungsschritt einsetzen und wie sie den weiteren Analyseprozess strukturieren.

Abschließend ist zu konstatieren, dass der Untersuchungsprozess selbst gezeigt hat, wie entscheidend die bewusste Kontrolle der Interaktionsparameter ist. Weder Moduswahl noch Personalisierung sind bloße technische Einstellungen, sondern konstitutive Bedingungen der Textproduktion. Ebenso erwies sich die Präzisierung der Instruktionssprache – etwa die terminologische Klärung zwischen ‚Arbeitsübersetzung‘ und ‚leserorientierter Übersetzung‘ – nicht als formale Nebensache, sondern als strukturierende Variable des Ergebnisses.⁵² Die Arbeit mit großen Sprachmodellen im historisch-philologi-

50 Eco (1973 = 2012), 27–59.

51 Wie Venuti (1995) in Kap. 1 zur Unsichtbarkeit, bes. S. 5–8 zeigt, ist Lesbarkeit in Übersetzungen kein neutrales Qualitätsmerkmal, sondern geht häufig mit einer Unsichtbarmachung der translatorischen Eingriffe einher, insofern syntaktische Glättung und semantische Explikation den Eindruck unmittelbarer Selbstverständlichkeit erzeugen und die zugrundeliegenden Übersetzungsentscheidungen verdecken.

52 Vgl. hierzu Ouyang et al. (2022), Wei et al. (2022) und Liu et al. (2023) wie oben in Anm. 9.

schen Kontext verlangt daher nicht nur textbezogene Analyse, sondern auch eine explizite Reflexion der eigenen Interaktionsentscheidungen. Künftige Untersuchungen werden diese Parameter systematischer variieren und dokumentieren müssen, um die Wechselwirkungen zwischen Instruktion, Modus, Personalisierung und Textsorte noch genauer zu bestimmen. Außerdem ist festzuhalten, dass in den untersuchten Passagen keine gravierenden Halluzinationen⁵³ auftraten: Als strukturelles Risiko bleibt jedoch die Möglichkeit unbelegter semantischer Ergänzungen bestehen, insbesondere bei textuell offenen oder elliptischen Stellen.

Damit wird Übersetzen mit ChatGPT zu einem Verfahren, das philologische Verantwortung nicht ersetzt, sondern neu organisiert: als kontrollierte, dokumentierbare Steuerung interpretativer Entscheidungen.

53 S. zu Halluzinationen Ji et al. (2023).

Quellenverzeichnis

Textausgaben

Clark (1905): Clark, A. C. (Hrsg.), M. Tulli Ciceronis Orationes, Bd. 1, Oxford 1905.

Ziegler (1964): Ziegler, K. (Hrsg.), Plutarchi vitae parallelae, Bd. 1.2, Leipzig 1964.

Literaturverzeichnis

Belosevic (2025): M. Belosevic, User-Centric Design Paradigms for Trust and Control in Human-LLM-Interactions: A Survey, Proceedings of the Fourth Workshop on Bridging Human-Computer Interaction and Natural Language Processing (HCI+NLP), 9. November 2025, Association for Computational Linguistics, 17–32, <https://doi.org/10.18653/v1/2025.hcinlp-1.3>.

Bowker / J. Buitrago Ciro (2019): L. Bowker / J. Buitrago Ciro, Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community, Bingley 2019.

Brown et al. (2020): T. B. Brown / B. Mann / N. Ryder / M. Subbiah / J. Kaplan / P. Dhariwal / A. Neelakantan / P. Shyam / G. Sastry / A. Askell / S. Agarwal / A. Herbert-Voss / G. Krueger / T. Henighan / R. Child / A. Ramesh / D. M. Ziegler / J. Wu / C. Winter / C. Hesse / M. Chen / E. Sigler / M. Litwin / S. Gray / B. Chess / J. Clark / C. Berner / S. McCandlish / A. Radford / I. Sutskever / D. Amodei, Language Models are Few-Shot Learners, NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems, 1877–1901, <https://dl.acm.org/doi/10.5555/3495724.3495883>.

Crane (2006): G. Crane, What Do You Do with a Million Books?, D-Lib Magazine 12/3 (2006), <https://www.dlib.org/dlib/march06/crane/03crane.html>.

Dassaritis (1889): E. Dassaritis, Die Psychologie und Pädagogik des Plutarch, Gotha 1889.

Eco (1973 = 2012): U. Eco, Das offene Kunstwerk, übers. v. G. Memmert, Frankfurt am Main 2012 (= 1973). Titel der Originalausgabe: Opera aperta, Mailand 1962.

Floridi / Chiriatti (2020): L. Floridi / M. Chiriatti, GPT-3: Its Nature, Scope, Limits, and Consequences. Minds and Machines 30 (2020), 681–694, <https://doi.org/10.1007/s11023-020-09548-1>.

Gadamer (1960 = 1990): H.-G. Gadamer, Gesammelte Werke, Bd. 1, Hermeneutik I: Wahrheit und Methode – Grundzüge einer philosophischen Hermeneutik, Tübingen 1990 (= 1960).

Ji et al. (2023): Z. Ji / N. Lee / R. Frieske / T. Yu / D. Su / Y. Xu / E. Ishii / Y. J. Bang / A. Madotto / P. Fung, Survey of Hallucination in Natural Language Generation, ACM Computing Surveys 55/12 (2023), 1–38, <https://dl.acm.org/doi/10.1145/3571730>.

Jiao et al. (2023): Y. Jiao / J. Wang / X. Huang / L. Chen / Z. Wang / X. Tu / J. Zhu / M. Wang / Y. Wang, Is ChatGPT a Good Translator? Yes With GPT-4 As The Engine, arXiv:2301.08745, <https://doi.org/10.48550/arXiv.2301.08745>.

Kolfhaus (1907): O. Kolfhaus, Plutarchi de communibus notitiis librum genuinum esse demonstratur, Marburg 1907.

Kurowsky (2025): S. Kurowsky, Übersetzen von Latein und Altgriechisch mit ChatGPT, Digital Classics Online 11 (2025), 1–33, <https://doi.org/10.11588/dco.2025.11.111062>.

- Lee (2024): T. Lee, Artificial Intelligence and Posthumanist Translation: ChatGPT versus the Translator, *Applied Linguistics Review* 15/6 (2024), 2351–2372, <https://doi.org/10.1515/applirev-2023-0122>.
- Liu et al. (2023): P. Liu / W. Yuan / J. Fu / Z. Jiang / H. Hayashi / G. Neubig, Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing, *ACM Computing Surveys* 55/9 (2023), 1–35, <https://doi.org/10.1145/3560815>.
- Liu et al. (2024): N. F. Liu / K. Lin / J. Hewitt / A. Paranjape / M. Bevilacqua / A. Petroni / P. Liang, Lost in the Middle: How Language Models Use Long Contexts, *Transactions of the Association for Computational Linguistics* 12 (2024), 157–173, https://doi.org/10.1162/tac1_a_00638.
- Mehta (2026): S. Mehta, HELM: A Human-Centered Evaluation Framework for LLM-Powered Recommender Systems, *arXiv:2601.19197v1* (2026), <https://doi.org/10.48550/arXiv.2601.19197>.
- O'Hagan (2020): M. O'Hagan, Translation and technology: Disruptive Entanglement of Human and Machine, In: M. O'Hagan (Hrsg.), *The Routledge Handbook of Translation and Technology*, London / New York 2020, 1–19.
- Ouyang et al. (2022): L. Ouyang / J. Wu / X. Jiang / D. Almeida / C. Wainwright / P. Mishkin / C. Zhang / S. Agarwal / K. Slama / A. Ray / J. Schulman / J. Hilton / F. Kelton / L. Miller / M. Simens / A. Askell / P. Welinder / P. F. Christiano / J. Leike / R. Lowe, Training Language Models to Follow Instructions with Human Feedback, *Advances in Neural Information Processing Systems* 35, *NeurIPS2022*, 27730–27744, https://proceedings.neurips.cc/paper_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf.
- Pelling (2002 = 2011): C. Pelling, *Plutarch and History*, Eighteen Studies, Swansea 2011 (= 2002).
- Palladino / Yousef (2023): C. Palladino / T. Yousef, To Say almost the Same Thing? A Study on Cross-Linguistic Variation in Ancient Texts and their Translations, *Digital Scholarship in the Humanities* 38/3 (2023), 1200–1212, <https://doi.org/10.1093/llc/fqac086>.
- Pym (2014): A. Pym, *Exploring Translation Theories*, London / New York ²2014.
- Pym (2023): A. Pym, *Exploring Translation Theories*, London / New York ³2023.
- Rosu et al. (2025): A. Roşu / A. Bota / D. Ştefănescu / R. Mititelu / M. Dinu, LITERA: Ad Verbum Translation of Latin Using Large Language Models, *Findings of the Association for Computational Linguistics: NAACL 2025*, 7371–7388, <https://doi.org/10.18653/v1/2025.findings-naacl.434>.
- Sahle (2013a): P. Sahle, *Digitale Editionsformen – Teil 2: Befunde, Theorie und Methodik. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels*, Norderstedt 2013.
- Sahle (2013b): P. Sahle, *Digitale Editionsformen – Teil 3: Befunde, Theorie und Methodik. Zum Umgang mit der Überlieferung unter den Bedingungen des Medienwandels. Textbegriffe und Recodierung*, Norderstedt 2013.
- Schleiermacher (1813 / 1963): F. Schleiermacher, Über die verschiedenen Methoden des Übersetzens, in: H. J. Störig (Hrsg.), *F. Schleiermacher, Das Problem des Übersetzens*, Stuttgart 1963 (Abhandlung v. Schleiermacher am 24. Juni 1813 in der Königlichen Akademie der Wissenschaften verlesen), 38–70.
- Schubert (2026): C. Schubert, Digital Classics: Digitale Transformation in den Altertumswissenschaften, in: W. Riess / G. Weber (Hrsgg.), *Alte Geschichte Forschung*, Göttingen 2026.
- Toury (1995): G. Toury, *Descriptive Translation Studies – and Beyond*, Amsterdam / Philadelphia 1995.

- Venuti (1995): L. Venuti, *The Translator's Invisibility, A History of Translation*, London / New York 1995.
- Volk et al. (2024): M. Volk / D. P. Fischer / L. Fischer / P. Scheurer / P. B. Ströbel, LLM-based Machine Translation and Summarization for Latin, *Proceedings of the Workshop on Language Technologies for Historical and Ancient Languages (LT4HALA 2024)*, 122–128, <https://aclanthology.org/2024.lt4hala-1.15/>.
- Wang et al. (2024): Y. Wang / Y.-S. Lin / R. Huang / J. Wang / S. Liu, Enhancing User Experience in Large Language Models through Human-Centered Design: Integrating Theoretical Insights with an Experimental Study to Meet Diverse Software Learning Needs with a Single Document Knowledge Base, *Computing and Artificial Intelligence* 2/1 (2024), 535, <https://doi.org/10.59400/cai.v2i1.535>.
- Wannaz / Miyagawa (2024): A.-C. Wannaz / S. Miyagawa, Assessing Large Language Models in Translating Coptic and Ancient Greek Ostraca, *Proceedings of the 4th International Conference on Natural Language Processing for Digital Humanities (NLP4DH 2024)*, 463–471, <https://doi.org/10.18653/v1/2024.nlp4dh-1.44>.
- Wei et al. (2022): J. Wei / X. Wang / D. Schuurmans / M. Bosma / B. Ichter / F. Xia / E. H. Chi / Q. V. Le / D. Zhou, Chain-of-Thought Prompting Elicits Reasoning in Large Language Models, *Advances in Neural Information Processing Systems* 35, *NeurIPS2022*, 24824–24837, https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf.
- Wittig / Kurowsky (2026): J. Wittig / S. Kurowsky, *Digitale Korpusarithmetik mit antikem Text*, *Digital Classics Books* 8, Heidelberg 2026.
- Zhang et al. (2025): R. Zhang / W. Zhao / S. Eger, How Good Are LLMs for Literary Translation, Really? Literary Translation Evaluation with Humans and LLMs, *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2025)*, 10961–10988, <https://doi.org/10.18653/v1/2025.naacl-long.548>.
- Zhao et al. (2023): W. X. Zhao / K. Zhou / J. Li / T. Tang / X. Wang / Y. Hou / Y. Min / B. Zhang / J. Zhang / Z. Dong / Y. Du / C. Yang / Y. Chen / Z. Chen / J. Jiang / R. Ren / Y. Li / X. Tang / Z. Liu / P. Liu / J.-Y. Nie / J.-R. Wen, A Survey of Large Language Models, *arXiv:2303.18223 [cs.CL]*, <https://doi.org/10.48550/arXiv.2303.18223>.

Autoreninformation⁵⁴

Dr. Sylvia Kurowsky
Universität Leipzig
Alte Geschichte
Ritterstr. 12, R. 302
04109 Leipzig
E-Mail: sylvia.kurowsky@uni-leipzig.de

⁵⁴ Die Rechte für Inhalt, Texte, Graphiken und Abbildungen liegen, wenn nicht anders vermerkt, bei der Autorin. Alle Inhalte dieses Beitrages unterstehen, soweit nicht anders gekennzeichnet, der Lizenz CC BY SA 4.0.