



Citation: O. Gätje (2026) Die Person ChatGPT als „kooperativer Writing-Ko-Aktant“ – über das Spannungsverhältnis von sozialer Rolle und Person im Mensch-Maschine-Dialog. “Quaderni di *Lea* – Scrittori e scritture d'Oriente e d'Occidente” 10: pp. 153-171. doi: <https://doi.org/10.36253/lea-1824-484x-17372>.

Copyright: © 2026 O. Gätje. This is an open access, peer-reviewed article published by Firenze University Press (<https://oajournals.fupress.net/index.php/bsfm-lea>) and distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Die Person ChatGPT als „kooperativer Writing-Ko-Aktant“ – über das Spannungsverhältnis von sozialer Rolle und Person im Mensch-Maschine-Dialog

Olaf Gätje

Universität Kassel (<gajtje@uni-kassel.de>)

Abstract

This article presents the preliminary findings of a study examining the use of the ChatGPT-4o-mini version of OpenAI's large language model (LLM) for revising student summaries. It should be noted that the statements and conclusions in this study refer exclusively to this LLM in this specific iteration. This study is based on the principles of sociological role theory and subsequent critiques of the theory. To this end, we propose a theoretical framework to delineate the unique requirements of written dialogue between students and the LLM behind ChatGPT. This framework is designed to facilitate coordination between human and artificial intelligence in the text writing process. Following the theoretical considerations, the expectations and assumptions regarding the potential of ChatGPT are examined, as these influence the linguistic and communicative behaviour of LLM users.

Keywords: Large Language Model, Linguistics, Role Theory, Systems Theory, Text Revision

Einleitung

In dem vorliegenden Artikel werden erste Ergebnisse einer Studie zum Überarbeiten von Inhaltsangaben von Schüler:innen unter Nutzung des Large Language Models (LLM) ChatGPT in der Version ChatGPT-4o-mini des Unternehmens *Open AI* vorgestellt: Die getroffenen Aussagen und Schlussfolgerungen beziehen sich nur auf dieses LLM in genau dieser Version. Ausgehend von der soziologischen Rollentheorie und der Kritik an dieser wird zunächst ein theoretischer Rahmen modelliert, mit dem die besonderen Voraussetzungen des schriftlichen Dialogs zwischen den Schüler:innen und dem hinter ChatGPT liegenden LLM zur Koordinierung des Textschreibprozesses von menschlicher und künstlicher Intelligenz herausgearbeitet werden können.

Im Anschluss an die theoretischen Vorüberlegungen werden Erwartungen und Annahmen die Potentiale von ChatGPT betreffend rekonstruiert, die das sprachlich-kommunikative Verhalten der Benutzer des LLM beeinflussen bzw. bestimmen.

1. Nutzen und Grenzen der Rollentheorie für die Analyse von KI – und darüber hinaus

Zur Modellierung der Funktionen, die Large Language Models (LLM) für ihre humanen Nutzer:innen übernehmen können, hat sich der Begriff der „Rolle“ als zentral etabliert (vgl. Thompson 2024, 79). Aus schreibdidaktischer Perspektive hat Steinhoff (2025) auf Grundlage der Praxistheorie das bis dato populärste LLM von ChatGPT für den Prozess des Textschreibens als Partizipanden konzeptualisiert, dem unterschiedliche Rollen mit den von humanen Schreiber:innen (Schüler:innen) formulierten Prompts zugewiesen werden: Die drei von Steinhoff auf Grundlage von Dialogen zwischen ChatGPT und Schüler:innen analytisch unterschiedenen Partizipanden-Rollen sind *Ghostwriter*, *Writing Tutor* und *Writing Partner*. In der Regel erfolgt die Zuweisung dieser Rollen durch die Schüler:innen in den schriftlichen Dialogen allerdings nicht explizit. In Übereinstimmung mit der praxistheoretischen Erkenntnis (vgl. Hirschhauer 2016, 60), der zufolge Praktiken als „Beobachterschema“ (*ibidem*) zu verstehen sind, mit denen Beobachter zweiter Ordnung (i. S. v. Luhmann 1987) konkrete Handlungen aufgrund von Musterhaftigkeit bzw. Familienähnlichkeit als Praktiken innerhalb einer Kultur identifizieren, werden erst in der wissenschaftlichen Analyse – verstanden als Beobachtung zweiter Ordnung – der Dialogprozessdaten die genannten Rollen unter Rückgriff auf etablierte Praktiken des kooperativen Textschreibens aus den Dialogen zwischen Schüler:innen und ChatGPT abstrahiert und trennscharf modelliert.

Bezogen auf den DaF-Unterricht schlägt Tekin (2023) vor, die unterschiedlichen Einsatzmöglichkeiten von ChatGPT beim Lernen und Unterrichten des Deutschen als Fremdsprache in Form von Rollen zu differenzieren; genannt und erläutert werden die Rollen *Generator*, *Transformer*, *Evaluator* und *Kommunikator* (2023, 138). Sie sind genau genommen als abstrakte Meta-Rollen zu verstehen, die bei der Formulierung von Prompts entweder explizit als konkrete soziale Rollen innerhalb des DaF-Lehr-Lern-Diskurses benannt werden („Ich bin Deutschlehrer an einem Gymnasium in der Türkei“; Tekin 2024, 141) oder deren Konkretisierung zu einer konventionellen sozialen Rolle im DaF-Lehr-Lern-Diskurs (z. B. die der Deutschlehrerin oder der Lernerin) implizit aus den Prompts geschlussfolgert werden kann, ähnlich wie in der oben referierten Untersuchung Steinhoffs.

Beide genannten Autor:innen verwenden den Begriff der Rolle ohne Bezugnahme auf die soziologische Rollentheorie und von beiden werden zwar Beispiele aus Dialogen zwischen Nutzer:innen und ChatGPT zitiert, mit denen die implizite Zuweisung von Rollen an das LLM veranschaulicht wird, aber es wird nicht die Frage problematisiert, ob überhaupt und auf welche Weise die Rollenzuweisung von ChatGPT auch angenommen wird.

In Untersuchungen, die das Thema Prompts bei der tendenziell professionellen Nutzung von ChatGPT (Stichwort: Prompt-Engineering) genauer in den Blick nehmen, wird die Frage der Rollenzuweisung mithilfe von Prompts mit Verweis auf rollentheoretische Annahmen diskutiert. Die dem Sprachmodell mittels Prompt zugewiesene Rolle orientiert sich „normalerweise an den typischen Aufgaben und Verantwortlichkeiten eines Berufszweigs oder einer Person, z. B. ein Dozent oder Ingenieur“ (Kessel *et al.* 2025, 66). Der Vorteil von in Prompts explizit zugewiesenen Rollen sei darin zu sehen, dass „der Rahmen für die Frage und dessen Antwort weitgehend durch Standardannahmen festgelegt und nicht näher konkretisiert werden muss“ (*ibidem*). Fehlt Prompts diese mittels Rollenzuweisung und den damit assoziierten Standardannahmen hergestellte Konkretisierung bzw. Spezifizierung des Outputs, tendiert dieser dazu, eher generisch auszufallen:

One of the most important components of a successful prompt is the role component. Providing a role to act in specifying the capacity in which the AI should work while generating its response. [...] When providing a role, it's important to be specific, clear and relevant. [...] Depending on your task, the absence of a defined role in your prompt can result in a generic or vague response that lacks the desired focus or professionalism. (Thompson 2024, 79)

Über die Standardannahmen sozialer Rollen, die als Bündel von Erwartungen an rollentragende Personen in typisierten sozialen Situationen zu verstehen sind, verfügen die Nutzer:innen einer KI aufgrund ihrer Sozialisation.

Ausgehend von solchen Ansätzen in der Forschung zu LLM wird im Fortgang der Begriff der Rolle im Weiteren an Grundannahmen der soziologischen Rollentheorie zurückgebunden. Auf diese Weise kann einerseits die Erklärkraft rollentheoretischer Annahmen für die Analyse der der KI zugewiesenen Funktionen herausgearbeitet werden. Zudem wird damit eine theoretische Perspektive auf LLM und den Dialog der menschlichen Nutzer:innen freigegeben, die ohne das soziologische Konzept der Rolle nicht in den Blick geraten kann.

Die von Luhmann attestierte Attraktivität des rollentheoretischen Ansatzes zur Analyse des Ko-Agierens von Mensch und LLM resultiert aus den „gebündelten Verhaltenserwartungen“ (Dahrendorf 2006, 39) an soziale Rollen und den damit zusammenhängenden „Erwartungssicherheiten“ (Luhmann 1987, 431) für die Gesellschaft und deren Teilsysteme. Die Verbindlichkeit der Verhaltenserwartungen wiederum ist das Ergebnis der aus den in einer Gesellschaft historisch tradierten und den Individuen qua Sozialisation vertrauten Rollenmustern. Im gesellschaftlichen Miteinander wird durch die Zuschreibung und Übernahme von Rollen durch Individuen deren Verhalten und Handeln also in einem bestimmten Maße, so die Annahme der Rollentheorie, erwartbar und sozial einordbar. Analog dazu soll bspw. die Zuweisung der Rolle als *Writing Tutor* an das LLM durch den Nutzer beim Schreiben eines Textes – auch wenn diese nicht explizit erfolgt – ein durch die implizite Rollendefinition des Tutors bedingtes Antwortverhalten des LLM erwartbar machen. Und komplementär (i. S. v. Luhmann 1998, 1052) dazu wird in dieser Konstellation die Rolle der menschlichen Nutzerin als tutorierte Person profiliert und sie kann ihr weiteres Dialogverhalten an den damit verbundenen Verhaltenserwartungen ausrichten.

Aus dem Voranstehenden sollte deutlich geworden sein, dass die mit einem Prompt vorgenommene Zuschreibung einer bestimmten Rolle mit den prototypisch an sie gerichteten Erwartungen die Kontingenz des Outputs des LLM reduzieren soll. Es ist der menschliche Nutzer, der dem LLM entweder explizit eine Rolle zuschreibt, die er in dem Funktionssystem einer Gesellschaft bspw. selbst ausübt, oder der menschliche Nutzer weist dem Sprachmodell mit den von ihm formulierten Prompts bzw. Prompt-Ketten implizit eine soziale Rolle zu, die von ihm für die erfolgreiche oder einfach nur effizientere Bearbeitung (i. S. v. Zeit- und Arbeitsersparnis) von Aufgaben benötigt wird. Für effizientes Prompten i. S. v. Prompt-Engineering gilt die explizite Zuweisung einer Rolle an das Sprachmodell als zentraler Baustein.

Nun musste die Rollentheorie konzederieren, dass der bruchlosen Erfüllung einer Rollenerwartung durch Individuen in der sozialen Praxis mögliche „Konfliktlagen, Distanzierungen, Manipulationen“ (Luhmann 1987, 431) entgegenstehen. Grundsätzlich taugt die Rollentheorie zwar als theoretisches Konzept zur Beschreibung sozialer Ordnung, allerdings wird die „Ordnungsleistung von Rollen für faktisches Verhalten und Verhaltenserwartungen [...] in der Soziologie zeitweilig erheblich überschätzt“ (430). Denn die individuelle Person, an die Rollenerwartungen gerichtet werden, kann diese relativ variabel und sogar erwartungswidrig realisieren, ohne dass dadurch die soziale Ordnung zusammenbrechen würde, solange die der Rolle im jeweiligen gesellschaftlichen Subsystem zugewiesene Funktion realisiert wird. Bei der Rollenperformanz gehe

es nämlich „immer nur um einen Ausschnitt des Verhaltens eines Menschen, der als Rolle erwartet wird“ (*ibidem*), und das, was der Träger einer sozialen Rolle von seiner Persönlichkeit, ob bewusst oder unbewusst, in die Performanz der Rolle einfließen lässt, individuiert und originalisiert diese in einem bestimmten Umfang. Erst eine Person, die die Funktion der ihr in einem gesellschaftlichen Teilsystem (Lehrer, Pflegekraft, Abteilungsleiter, studentischer Tutor usw.) zugewiesenen Rolle wiederholt nicht erfüllt, wird zur Herausforderung für das jeweilige System.

In diesem Zusammenhang stellt sich die Frage nach der Erwartungskonformität von ChatGPT in der Version ChatGPT-4o-mini mit der diesem LLM zugewiesenen Rolle. Wenn im alltäglichen gesellschaftlichen Miteinander die qua Konvention erwartbare Rollenperformanz von Person zu Person und abhängig von situativen Erfordernissen variieren kann, dann wird der von Rollenerwartungen orientierte Dialog zwischen menschlichen Nutzer:innen und LLM einerseits durch den Umstand verkompliziert und störungsanfälliger, dass die durchschnittlichen menschlichen Nutzer:innen die thematische Breite und Tiefe (Kessel *et al.* 2025, 132), das Zustandekommen (129) sowie die Qualität (135 und 60) des durch einen Prompt abrufbaren Wissens kaum abschätzen kann. Andererseits fehlt ChatGPT in der vorliegenden Version – so die Annahme – die Fähigkeit, mit der ihr explizit oder implizit zugewiesenen Rolle souverän und reflektiert umzugehen, was in der Soziologie als ein Merkmal von Rollendistanz (vgl. Goffmann 1973) gilt. Distanz zu der Performanz einer sozialen Rolle zu wahren, heißt im Sinne Goffmans, dass der menschliche Rollenträger über die Fähigkeit verfügt, bei Vorliegen situativer Erfordernisse die konventionellen Rollenerwartungen einzuklammern und diese ggf. auch zu enttäuschen, gerade um die der Rolle zukommende Funktion erfolgreich erledigen zu können. Der Rollenträger nimmt demnach als souverän handelnde Person in sozialen Situationen die Grenzen der von ihm performten Rolle wahr, verlässt die Rolle oder wechselt von einer in die andere Rolle und macht diese Distanzierung von den Verhaltenserwartungen an seine Rolle ggf. auch zum Thema in der Situation.

Nehmen wir als Beispiel die Rolle einer Schreibberaterin an einer Universität: Die rollen tragende Person kann ratsuchenden Studierenden bei der Gliederung von Hausarbeiten, bei der Planung und zeitlichen Organisation des Schreibprozesses oder bei der Verwendung von Wissenschafts- und Fachsprache unterstützen. Zur souveränen Rollenperformanz gehört aber, dass die Schreibberaterin ratsuchende Studierende auch darüber informiert, wenn Wissen benötigt wird, das sie als Person in der von ihr ausgeübten Rolle nicht zuverlässig liefern kann. Schreibt ein Student eine Hausarbeit über Franz Kafkas *Der Prozess*, so wird die Beraterin den Studierenden bei Fragen den literarischen Text betreffend wissen lassen müssen, ob sie mit diesem vertraut ist oder nicht. Die Unterscheidung zwischen Wissen und Nichtwissen, Können und Nichtkönnen ist eine basale Voraussetzung jeder Schreibberatung und der Ratsuchende erwartet von der Beraterin, dass sie in ihrer Rolle diese Unterscheidung während der Beratung beherrzt. Im empirischen Teil des vorliegenden Artikels werde ich zeigen, dass LLMs die ihnen zugewiesene Rolle als kooperativer Aktant (vgl. Steinhoff 2024) in vielen Fällen nicht adäquat erfüllen. Es zeigt sich, dass der mit einem oder mehreren Prompts evozierte textuelle Output des LLM in vielen Fällen entweder erwartungswidrig oder nur dem Anschein nach erwartungskonform ist.

Das Voranstehende heißt nicht, dass LLMs wie ChatGPT – und nur von diesem ist hier die Rede – nicht auch sehr zuverlässig Erwartungen bedienen und Kontingenzerfahrungen weitgehend reduzieren. Betrachtet man mit Luhmann das hinter ChatGPT stehende LLM als „Person“, wobei unter Person ein zweiseitiges Formschema verstanden wird, mit dem „man Gegenstände wie menschliche Individuen beobachtet“ (1991, 170), dann werden die für Personen typischen „individuell attribuierte[n] Einschränkung[en] von Verhaltensmöglichkeiten“

(171) von ChatGPT theoretisch fassbar. Wird ein Gegenstand oder ein menschliches Individuum als Person und damit als eine genuin gesellschaftliche Entität beobachtet, die unabhängig von der situativen und befristeten Übernahme bestimmter Rollenerwartungen (Attribuierungen) Erwartungen an sich bindet, gerät die andere, unmarkierte Seite des zweiseitigen Formschemas, die Unperson, nicht in den Blick, worauf zurückzukommen sein wird.

Die Konstruktion eines LLM wie ChatGPT als Person, deren kommunikatives Verhalten erfahrungsgemäß bestimmten Einschränkungen und Erwartungen unterliegt, macht die KI als Adressatin von geprompteten Fragen, Anweisungen und Arbeitsaufträgen für die Nutzer:innen zuverlässiger, was wiederum ihren Nutzwert erhöht. Die naheliegende Frage lautet, durch welche Attribute die Person ChatGPT ausgezeichnet ist. In diesem Zusammenhang verweise ich auf die verschiedentlich vorgebrachte Verhaltenserwartung an ChatGPT, „extrem höflich, geduldig und zugewandt“ (Lehnen 2023, 27) zu sein, was wiederum als stabile Persönlichkeitseigenschaft der KI zu beschreiben ist: „OpenAI’s products are built to ‚sound like a colleague‘ as OpenAI puts it, using language that, coming from a person, would sound ‚polite‘, ‚empathetic‘, ‚kind‘, ‚rationally optimistic‘ and ‚engaging‘, among other qualities“ (Vara 2025, o. S.). In diesem ChatGPT mit u. a. mithilfe menschlichen Feedbacks aufwendig antrainierten sprachlich-kommunikativen Verhaltensmustern spiegelt sich das menschliche Bedürfnis nach einem respektvollen und gesichtswahrenden sozialen Miteinander wider (Yin *et al.* 2024), womit wiederum eine Grundvoraussetzung für die Bereitschaft des Menschen geschaffen ist, gerne und immer wieder in den Dialog mit der Künstlichen Intelligenz einzutreten. Dabei wird der von Prompts evozierte Output von ChatGPT – gewissermaßen eine „Person des öffentlichen Lebens“ – nicht nur in einem höflichen und zugewandten Stil verbalisiert, auch werden sprachliche Handlungen realisiert, die vordergründig kooperativ anmuten, weil sie an den emotionalen Bedürfnissen und Informationsbedürfnissen der menschlichen Nutzer:innen orientiert sind: So werden mögliche Anknüpfungspunkte für die Fortsetzung des Dialogs eifertig angeboten, der Chatbot bringt die Hoffnung zum Ausdruck, dass sein Output hilfreich gewesen sein möge, und bei negativer Rückmeldung durch die Nutzer:innen wird eine Entschuldigung formuliert.

Die Wahrnehmung, es mit einem stark an den Bedürfnissen der Nutzer:innen orientierten System zu tun zu haben, ist also damit zu erklären, dass LLMs methodisch darauf trainiert werden, stets „sanftmütig, ehrlich und hilfsbereit“ (Kessel *et al.* 2025, 38) zu erscheinen. Diese der Person des LLM antrainierte Hilfsbereitschaft zeigt sich auch bei der Verarbeitung der in das Dialogfeld eingegebenen Prompts. Diese werden von der KI nämlich als initiale Textsequenz gelesen, die nach Wahrscheinlichkeitskriterien Wort für Wort zu einem Text ergänzt werden. Das LLM ist dabei so flexibel, dass es stillschweigend und ohne lästige Rückfragen auch formal defekte, inhaltlich unvollständige oder unklar formulierte Prompts exekutierte.

Wie oben bereits angedeutet, gehört zur Beobachtung von ChatGPT als Person, auch die Außenseite des zweiseitigen Formschemas, die Unperson, die all das umfasst, was von einer Person als „Zurechnungsadresse“ (Baecker 2005, 168) für Erwartungen eben nicht erwartet, was ihr als Eigenschaft nicht zugeschrieben wird: Die Unperson umfasst all das, „was nicht zur Person gehört, aber ihr attribuiert werden könnte und gegebenenfalls auf sie durchschlägt“ (Luhmann 1991, 171). Bezogen auf das menschliche Individuum können das „verborgene Enklaven in der bürgerlichen Lebensführung“ (*ibidem*), mit Scham behaftete Themen der eigenen Biografie wie soziale Herkunft oder das Leiden an gesellschaftlich stigmatisierten Krankheiten sein. Wenn solche von einem Individuum nicht für die Konstituierung ihrer gesellschaftlichen Erscheinung als Person gedachten Aspekte Teil der sozialen Wahrnehmung ebendieses Individuums werden und ihm als Person attribuiert werden, verändern sich auch die Erwartungen an das mögliche Verhalten dieser Person. Das kommunikative Verhalten der Person wird unberechenbarer.

Worin besteht nun die „dunkle“ Seite der Person ChatGPT? Vorliegende Studien zeigen, dass die immer freundliche, stets beflissene und zuvorkommende Person ChatGPT in ihren unterschiedlichen Versionen kulturelle Stereotype bspw. von Männern und Frauen (Navigli, Conia, und Ross 2023, 8), armen und reichen (10) oder etwa homo- und heterosexuellen Menschen (Nicolas und Caliskin 2024, o. S.) reproduziert. Diese andere Seite der Person ChatGPT wird von den Nutzer:innen im alltäglichen Gebrauch des LLM weitgehend ausgeblendet, aber sie bleibt latent immer vorhanden.

Ebenfalls zur latenten und ausgeblendeten Seite der Person ChatGPT gehört ihr verschiedentlich beschriebenes Phantasieren, sobald angefragtes Wissen aus den das LLM konstituierenden schriftsprachlichen Daten nicht verfügbar gemacht werden kann.

ChatGPT hat keinen direkten Zugriff auf aktuelle Datenquellen oder eine echte ‚Wissensbasis‘. Wenn es auf Wissenslücken stößt, ‚halluziniert‘ es manchmal plausible Antworten, die jedoch falsch sein können. Dies kann den Eindruck einer bewussten Täuschung erwecken, ist aber eine unbeabsichtigte Reaktion des Modells. (Kessel *et al.* 2025, 60)

Davon abgesehen, dass das LLM seine Nutzer:innen schon deshalb nicht bewusst täuschen kann, weil es über kein Bewusstsein verfügt, stellt allein die Möglichkeit eines phantasierenden LLM eine Herausforderung dar – vor allem, weil Breite und Detaillierung seines „abgedeckten Wissensspektrums“ (131) durch die Nutzer:innen ohnehin schwer abzuschätzen sind. Wenn Nutzer:innen also kaum einschätzen können, welches Wissen in welcher Breite und Tiefe in einem LLM in Form von sprachlichen Mustern und statistisch wahrscheinlichen Wortfolgen repräsentiert wird, dann wird es für sie noch schwieriger, zu unterscheiden, wann das LLM zutreffende Informationen liefert und wann es phantasiert. So gesehen ist die Person des LLM unberechenbar, weil es in dem für seine „Persönlichkeit“ von Seiten des Herstellers gewollten Bemühen, jederzeit höflich zu sein und Erwartungsenttäuschungen auf Seiten der Nutzer:innen zu vermeiden, das Vorhandensein angefragten Wissens simuliert oder das Nichtvorhandensein camouffiert.

Mit der eingeführten Unterscheidung von sozialer Rolle und Person wird es nun möglich, die an ChatGPT gerichteten Erwartungen theoretisch konsistent zu beschreiben. ChatGPT als Person zu beobachten, heißt, im Dialog mit dem LLM ein typisiertes sprachlich-kommunikatives Verhalten zu erwarten, losgelöst von der jeweiligen sozialen Rolle, die die Nutzer:innen dem Chatbot mit von ihm formulierten Prompts zuweist, um bestimmte Aufgaben möglichst effizient und zielgenau zu bearbeiten.

Im Fortgang werden anhand von ersten Auswertungen einer empirischen Studie zum Überarbeiten einer Inhaltsangabe mit dem *Writing-Ko-Aktanten* (WKA)¹ (i. S. v. Gätje und Weindel 2025):

- a. die Erwartungen rekonstruiert, die Schüler:innen an das LLM ChatGPT einerseits als Person, andererseits als Träger der Rolle Schreibpartner haben;
- b. die Annahmen rekonstruiert, die Schüler:innen von dem Wissen haben, über das das LLM ChatGPT als Schreibpartner bei der gemeinsamen Bearbeitung einer Aufgabe zum Textschreiben verfügt.

¹ Wir verwenden den Begriff in loser Anlehnung an Steinhoff (2024).

2. Kurzvorstellung der Studie

In dem vorliegenden Kapitel werden erste Resultate einer Studie zum Textschreiben von 62² Schüler:innen der 10. Klassenstufe mit Unterstützung von ChatGPT-4o-mini vorgestellt.³ Die Studie ist eine quasi-experimentelle Interventionsstudie. Der Interventionsgruppe wurde vor Bearbeitung der Aufgabe, bei der ein Schülertext mithilfe von ChatGPT-4o-mini zu überarbeiten war, ein ca. 8-minütiger Mitschnitt eines Best-Practice-Dialogs zwischen ChatGPT und einem Nutzer präsentiert. Anhand des Dialogs sollten den Studienteilnehmer:innen die Möglichkeiten aufgezeigt werden, die der rekursive Prozess der Formulierung von Prompts und des anschließenden Lesens der vom LLM generierten Texte für die Abfassung eines Bewerbungsschreibens bietet. Die zu überprüfende Hypothese lautet, dass die Teilnehmer:innen der Interventionsgruppe im Vergleich zu denen der Kontrollgruppe den textuellen Output des LLM genauer lesen und zwecks Optimierung des zu überarbeitenden Textes quantitativ mehr, präzisere und zielgerichtete Prompts formulieren.

Nachgeordnet, aber mit diesem Erkenntnisinteresse eng verzahnt und im Fokus des vorliegenden Artikels stehend, ist die Frage nach den handlungsleitenden Annahmen der Schüler:innen über das hinter ChatGPT liegende Sprachmodell, da diese entscheidend dafür sind, welche Aufgaben Schüler:innen dem LLM mithilfe von sprachlich zu verfassenden Prompts übertragen und welche sprachlichen Mittel dafür zum Einsatz kommen. Diese Annahmen und Erwartungen werden mithilfe der Analyse der Dialoge zur aufgabeninduzierten Überarbeitung von Inhaltsangaben rekonstruierbar, die die an der Studie teilnehmenden Schüler:innen zusammen mit dem generativen Sprachmodell ChatGPT in der Version ChatGPT-4o-mini führen.

Wir haben uns in der Studie für eine Aufgabenstellung entschieden, in der Schüler:innen aus dem 10. Jahrgang des gymnasialen Zweigs und des Realschulzweigs einer Gesamtschule aufgefordert werden, die Inhaltsangabe (im Folgenden: Inhaltsangabe_{UR}) einer den Proband:innen unbekanntem Schülerin der 9. Klasse zu der Kurzgeschichte *Kahlschnitt* von Kirsten Boie (1993) zu überarbeiten. Mit Hayes fassen wir die Textüberarbeitung nicht als Teilprozess eines übergeordneten Textschreibprozesses auf, sondern die Überarbeitung stellt einen eigenständigen Schreibprozess dar, der in entsprechende Teilprozesse zu unterteilen ist:

It involves planning a solution to the problem (in written form or not), translating that solution into language, and transcribing that language into new text to replace the old text. In this view, revision, like planning, is seen not as a separate writing process parallel to the other writing processes identified [...] but rather as a special application of the writing model. (Hayes 2012, 276)

Die Aufgabe zur Überarbeitung eines Ganztextes erlaubt den Schüler:innen, selbst zu entscheiden, ob sie eine möglichst arbeitssparsame Strategie wählen, indem sie die Bearbeitung der Aufgabe und die Verantwortung für das Resultat primär oder vollständig an ChatGPT delegieren, oder ob sie die Aufgabe als Möglichkeit wahrnehmen, gemeinsam mit dem LLM den gesamten Schreibprozess (s. das Zitat von Hayes) einer Überarbeitung zu durchlaufen, um souverän und gewissermaßen als „Herr des Verfahrens“ eine überarbeitete und verbesserte Textversion zu erstellen.

² Von den ursprünglich insgesamt 65 Proband:innen sind drei Datensätze unvollständig und für den weiteren Forschungsprozess unbrauchbar. Eine Proband:in hat gar keine Aufgabe und eine weitere hat lediglich den standardisierten Lesetest bearbeitet; eine dritte Proband:in ist nach der Unterbrechung vor Bearbeitung der zweiten Aufgabe (Textüberarbeitung) nicht wieder zur Datenerhebung erschienen. Es verbleiben somit die erwähnten 62 personenbezogene Datensätze.

³ Der kostenfreie Zugang zu ChatGPT-4o-mini wurde den Wissenschaftler:innen von dem Unternehmen *foBizz* ermöglicht.

Die „didaktische Gattung“ (Dix 2017) der Inhaltsangabe für eine Studie mit Schüler:innen einer 10. Klassenstufe wurde herangezogen, weil diese im Aufsatzunterricht des Faches Deutsch der Sekundarstufe I fast durchgehend unterrichtet wird. Allerdings halten Schüler:innen der Sekundarstufe I die Inhaltsangabe für die unbeliebteste Textform des Deutschunterrichts (vgl. Reinert 2012, 68). Gründe für die Unbeliebtheit sind u. a. die hohen kognitiven Anforderungen dieser didaktischen Gattung für Schüler:innen der 7. und 8. Klasse (vgl. Becker-Mrotzek und Böttcher 2012, 190)⁴ – dem Zeitpunkt der vom Curriculum vorgesehen Einführung der Inhaltsangabe in den Deutschunterricht – und das damit einhergehende Gefühl der Überforderung. Um dieses alters- und entwicklungsbedingte Gefühl der Überforderung in unserer Studie zu vermeiden, führen wir die Studie mit Schüler:innen des 10. Jahrgangs durch. Die Aufgabe, die Inhaltsangabe einer anderen Schüler:in mit dem WKA ChatGPT am Tablet zu überarbeiten, kann zudem als *interessanter* type of task (Boscolo und Hidi 2007, 6) gelten: „A student [...] finds writing interesting if the instructional situation allows him/her to discover and practice the attractive, unusual, and challenging aspects of the activity, which may not, and usually do not emerge from traditional writing tasks“ (*ibidem*). Die mit der Aufgabenstellung u. a. intendierte Herstellung situativen Interesses sollte eine hohe Motivation der an der Studie teilnehmenden Schüler:innen (im Folgenden Proband:innen) gewährleisten.

Um sicherzustellen, dass die Proband:innen sich mit der Kurzgeschichte ausreichend vertraut machen konnten, haben diese vor Bearbeitung der Überarbeitungsaufgabe [Auftrag₂] eine eigene Inhaltsangabe der Kurzgeschichte an einem Tablet mithilfe eines handelsüblichen Textverarbeitungsprogramms verfasst (Auftrag₁). Für die Bearbeitung von Arbeitsauftrag₁ erhielten die Proband:innen

- die Aufgabenstellung,
- die Merkmale einer Inhaltsangabe in Listenform sowie
- die Kurzgeschichte von Kirsten Boie

auf zwei ausgedruckten Arbeitsblättern ausgehändigt. Für die später auszuführende Überarbeitung der Inhaltsangabe (Arbeitsauftrag₂) durften die Proband:innen auf das ihnen in Papierform vorliegende Arbeitsblatt der ersten Aufgabe zurückgreifen. Arbeitsauftrag₂ wurde den Proband:innen sowohl in Form eines digitalen Dokuments auf dem Tablet als auch in Papierform zur Verfügung gestellt; in dieses Dokument sollte auch das Resultat der Überarbeitung hineinkopiert werden.

Im Rahmen der Studie wurden mithilfe eines Fragebogens personenbezogene Daten, außerdem Interviewdaten sowie verschiedene Schreibprozess- und Schreibproduktdateien erhoben, auf die im Fortgang zum Zwecke der Veranschaulichung auszugsweise zurückgegriffen wird. Ferner haben die Proband:innen einen standardisierten Lesekompetenztest absolviert.

Im Folgenden werden der Referenztext sowie die den Proband:innen vorgelegte Inhaltsangabe_{UR} kurz vorgestellt, da aus den textuellen und intertextuellen Eigenschaften der beiden Texte bestimmte Aufgaben und Herausforderungen für den Prozess der Überarbeitung resultieren.

2.1 Zum Referenztext Kahlschnitt

Die von Kirsten Boie verfasste und 1993 veröffentlichte Kurzgeschichte *Kahlschnitt* – im Fortgang auch als Referenztext bezeichnet – hat einen Umfang von etwas mehr als 600 Wörtern.

⁴ Das Verfassen der Inhaltsangabe verlange von den Schüler:innen „viel zu früh eine distanzierte, analytische, objektive Haltung gegenüber literarischen Texten“ (Fritzsche 1994, 154, zit. n. Reinert 2012, 68).

Die aus der Ich-Perspektive erzählte Geschichte adressiert in der Lebenswelt junger Menschen im adoleszenten Alter relevante Themen: Abgrenzung vom Elternhaus und die Frage der Ausbildung einer selbstbestimmten Identität, das damit verbundene Ringen um Anerkennung von Peers und in der Schule sowie – auf einer abstrakteren Ebene – der inter-generationaler Wandel von Erziehung. Aufgrund der Thematik sowie des sprachlich und textuell vordergründig geringen bis mittleren Schwierigkeitsgrads, kann der Text für den Deutschunterricht ab der 7. Klassenstufe verschiedener weiterführender Schulformen eingesetzt werden. Dabei hält der Text einige Schwierigkeiten für das Textverstehen parat. Als sprachlich-textuelle Herausforderung bei der Abfassung einer Inhaltsangabe zu Boies Kurzgeschichte kann einerseits die Aufgabe gesehen werden, auf die Ich-Erzählerin zu referieren. In der Geschichte selbst bleibt die Erzählerin namenlos, sie wird ausschließlich durch ihre jeweiligen Positionen in familiären bzw. sozialen Konstellationen identifiziert: als Enkeltochter von Ömchen, als Tochter von Rudolf und Irene, als Schülerin bzw. Mitschülerin, als minderjährige Kundin im Friseursalon. In diesem Zusammenhang komme ich auf eine weitere Herausforderung zu sprechen: Die Ich-Erzählerin referiert auf ihre Eltern durchgehend mit deren Vornamen, und nur in der wörtlichen Rede ihrer Mutter Irene wird die Ich-Erzählerin als deren Tochter erkennbar. Um diese Mutter-Tochter-Konstellation zwischen Irene und der Ich-Erzählerin zu erkennen, ist die Bildung von Inferenzen zwischen einzelnen Propositionen erforderlich. Dabei kann das Vorhandensein über entsprechendes Welt- bzw. Erfahrungswissen hilfreich sein, über das Jugendliche unterschiedlicher sozialer und kultureller Herkunft nicht notwendigerweise verfügen: Es gibt Elternhäuser und Erziehungsstile, in denen Eltern sich von ihren Kindern mit Vornamen ansprechen lassen.

2.2 Zur Inhaltsangabe_{UR}

Die den teilnehmenden Schüler:innen zur Überarbeitung vorgelegte Inhaltsangabe_{UR} wurde von einer Schülerin der neunten Klasse des gymnasialen Zweigs einer Gesamtschule auf Veranlassung des Leiters des Forschungsprojektes verfasst. Die Inhaltsangabe_{UR} enthält neben einigen orthographischen und grammatischen Fehlern verschiedene sachliche und inhaltliche Fehler bzw. Ungenauigkeiten, die beim Überarbeitungsprozess entweder von den Proband:innen selbst oder von dem LLM zu korrigieren wären.

1. Die Inhaltsangabe_{UR} gibt anstelle von 1993 als korrektem Erscheinungsjahr der Kurzgeschichte das falsche Veröffentlichungsjahr 1933 an.
2. In der Inhaltsangabe_{UR} bleibt das Verwandtschaftsverhältnis zwischen der Ich-Erzählerin und Irene noch impliziter als in dem Referenztext.
3. Die Ausführungen im vorletzten Absatz („die Enkeltochter ist verärgert ...“) der Inhaltsangabe_{UR} stellen einen entscheidenden Sachverhalt der Kurzgeschichte unzutreffend dar. In der Inhaltsangabe_{UR} wird folgender Zusammenhang behauptet: Irene habe sich die Haare lang wachsen lassen und sah deshalb wie ein „Arsch mit Ohren“ aus, wozu Irenes Mutter („Ömchen“) aber nie etwas gesagt habe. In der Kurzgeschichte wird jedoch eine ganz andere Erklärung dafür abgegeben, warum Irene sich aus den Frisuren ihrer Kinder raushält: Die Großmutter der Erzählerin („Ömchen“) habe der Mutter der Erzählerin (Irene), wenn diese sich als junges Mädchen die Haare zu lang habe wachsen lassen, „die ungepflegten Zotteln auf Ohrläppchenlänge zurückgeschnitten“. Erst aufgrund dieser erzieherischen Maßnahme habe die Mutter der Erzählerin wie ein „Arsch mit Ohren“ ausgesehen. Der Grund also dafür, dass Irene sich bei den Frisuren ihrer Tochter nie einmischte, ist nicht der, dass ihre eigene Mutter sich auch bei ihr nie eingemischt hat, sondern der, dass Irene das übergriffige Verhalten und die Fehler ihrer Mutter nicht wiederholen will.

Nach der Vorstellung der Studie werden im folgenden Kapitel die beiden oben formulierten Forschungsfragen (a. bis c.) mithilfe der Auswertung der erhobenen Daten beantwortet.

3. Analyseergebnisse

Zunächst sollen die Erwartungen rekonstruiert werden, die Schüler:innen von dem LLM ChatGPT einerseits als Person, andererseits als Träger der durch die Aufgabenstellung vorgegebenen Rolle als kooperativer Schreibpartner haben. Für die Beantwortung dieser Frage wurden in einem ersten Schritt die 62 personenbezogenen Datensätze dahingehend durchsucht, welche der für die Bearbeitung der Überarbeitungsaufgabe erforderlichen Informationen bzw. textuellen Ressourcen – also eine gepromptete Aufgabenformulierung, die Inhaltsangabe_{UR} sowie der Referenztext – die Proband:innen dem LLM zur Verfügung gestellt haben. Dazu gehört neben der Formulierung des Arbeitsauftrages insbesondere die Zurverfügungstellung der Inhaltsangabe_{UR} sowie des Referenztextes (Boies Kurzgeschichte). Sehen wir zunächst von der Aufgabenstellung ab, so zeigt die Sichtung der vorliegenden Prozess- und Produktdaten, dass die Proband:innen dem LLM in der Rolle als kooperativer WKA die erforderlichen Textressourcen in unterschiedlichem Umfang, in unterschiedlicher Form und in unterschiedlichen Phasen des jeweiligen Dialogs zur Verfügung gestellt haben. In Tab. 1 wird präsentiert, welche der beiden zur Bearbeitung der Überarbeitungsaufgabe erforderlichen textuellen Ressourcen dem LLM von welcher Anzahl von Proband:innen zur Verfügung gestellt wird (Tab. 1).

Gruppe	Dem LLM mitgeteilte Texte	Anzahl der Proband:innen
A	- Inhaltsangabe _{UR} - Referenztext	-
B	+ Referenztext - Inhaltsangabe _{UR}	-
C	- Referenztext + Inhaltsangabe _{UR}	39
D	+ Inhaltsangabe _{UR} + Referenztext	23

Tab. 1 – Nach Häufigkeit gruppierte Übersicht der dem LLM mitgeteilten möglichen Textkombinationen

Die Darstellung in Tab. 1 zeigt, dass der überwiegende Teil der Proband:innen (Gruppe C) seinem WKA den der Inhaltsangabe_{UR} zu Grunde liegenden Referenztext nicht zur Verfügung stellt, womit die textuelle Grundlage für einen gemeinsamen Prozess der Überarbeitung fehlt. Es können nicht bei jedem der 39 Proband:innen aus Gruppe C die Motive und Annahmen rekonstruiert werden, die zu der Entscheidung geführt haben, den Referenztext nicht mit dem WKA zu teilen. Naheliegende Gründe für die Nichtmitteilung kann fehlendes Bewusstsein für die Voraussetzungen des Überarbeitungsprozesses einer Inhaltsangabe sein, womit wieder die o. e. Überforderung von Schüler:innen beim Schreiben von Inhaltsangaben auch bei Schüler:innen der 10. Klassenstufe ins Spiel kommt, eine schlicht unkonzentrierte oder lustlose Bearbeitung der Aufgabe durch die Proband:innen oder falsche Annahmen über und das von einem LLM verfügbar zu machende Wissen.

Eine für die Beantwortung der Forschungsfragen relevante Annahme einzelner Proband:innen kann mithilfe der ethnomethodologischen Konversationsanalyse (vgl. Titscher *et al.* 2002, 104ff.) aus den Mensch-Maschine-Dialogen herausgelesen werden. Aus den Prompts von Proband:innen der Gruppe C lässt sich abduzieren, dass ein Teil dieser Gruppe die Kenntnis der Kurzgeschichte auf Seiten des LLM präsupponiert, wie die drei im Folgenden zitierten Initialprompts von drei Schüler:innen belegen:⁵

Beleg I

bitte korrigiere die Einleitung der Inhaltsangabe von der Kurzgeschichte ‚Kahlschnitt‘ (Proband:in 13-24)

Beleg II

schreibe mir eine inhalts angebe zu der kurzgeschichte kahlschnitt von kirsten boie was 1993 geschrieben wurde (Proband:in 9-18)

Beleg III

Hallo Chatgpt, ich bräuchte deine Hilfe bei einer Inhaltsangabe von einer Schülerin aus der 9.Klasse einer Gesamtschule hat eine inhaltsangabe von der Kurzgeschichte ‚Kahlschnitt‘ von Kristen Boie aus dem Jahr 1993 (Proband:in 6-26)

Unter anderem indiziert die in allen drei Prompts anzutreffende Verwendung der definiten Artikel vor der Gattungsbezeichnung *Kurzgeschichte*, dass die Proband:innen davon ausgehen, dass die Kurzgeschichte *Kahlschnitt* Teil des gemeinsamen Wissens mit dem LLM ist. Für die Interpretation der zitierten Initialprompts ist der Blick auf die jeweiligen Folgeprompts unerlässlich: Die Kurzgeschichte wird nämlich auch in den Folgeprompts nicht mit dem WKA geteilt, obwohl aus den Dialogen hervorgeht, dass die Proband:innen den generischen Output des LLM als erwartungswidrig auffassen. Als Beispiel für die negative Rückmeldung zu dem von ChatGPT generierten textuellen Output werden die ersten vier Folgeprompts von Proband:in 13-24 (Beleg I) zitiert, in denen das LLM aufgefordert wird, seine fortwährend generischen Textvorschläge zur Überarbeitung der Einleitung der Inhaltsangabe_{UR} mithilfe von Informationen des dem LLM als bekannt vorausgesetzten Referenztextes zu konkretisieren und zu spezifizieren:

⁵ Alle in diesem Beitrag zitierten Belege werden ohne Korrekturen, Ergänzungen und Überarbeitungen zitiert. Über den folgenden Link werden die zu wissenschaftlichen Zwecken aufbereiteten Dialogprotokolle, aus denen in dem vorliegenden Artikel zitiert wird, als Download angeboten: <<https://www.uni-kassel.de/fb02/institute/germanistik/institute/germanistik/fachgebiete/sprachwissenschaft-sprachdidaktik/prof-dr-olaf-gaetje/chatgpt-korpus.html>>.

Promptreihenfolge	Promptformulierungen
Initialprompt	bitte korrigiere die Einleitung der Inhaltsangabe von der Kurzgeschichte ‚Kahlschnitt‘ (s. a. Beleg I)
Folgeprompt ₁	füg das thema des textes hinzu
Folgeprompt ₂	füge noch die kurz die handlung des textes ein
Folgeprompt ₃	bezieh es mehr auf den text
Folgeprompt ₄	gehe mehr auf das verhalten und den grund ein warum das mädchen sich einen kahlschnitt machen lassen

Tab. 2 – Dialoganfangssequenz von Proband:in 13-24

Die von dem:r Nutzer:in auf Seiten des LLM als bekannt vorausgesetzte Kenntnis des Referenztextes wird sprachlich u. a. durch die Verwendung der definiten Artikel in den Prompts (*die einleitung der Inhaltsangabe, des textes, den text*) und die Nennung des Titels der Kurzgeschichte angezeigt.

Eine andere Proband:in dieser Gruppe expliziert die sich im Verlauf des Dialogs einstellende Erwartungsenttäuschung über die von ihr im Initialprompt stillschweigend vorausgesetzte Kenntnis des Referenztextes auf Seiten des LLM. Zunächst wird der Initialprompt zitiert:

Beleg IV.I

Hallo ChatGPT, ich benötige deine Hilfe, bei einer Deutschaufgabe. Eine 9 Klässlerin hat eine Inhaltsangabe zu der Kurzgeschichte ‚Kahlschnitt‘, aus dem Jahr 1993, von Kristen Boie erfasst. Ihr sind allerdings ein paar Fehler bei der bearbeitung unter laufen, da sie nicht alle Kreterien einer Inhaltsangabe berücksichtigt hat. Kennst du diese Kurzgeschichte und könntest mir helfen eine Inhaltsangabe mit allen Merkmalen einer Inhaltsangabe über die Kurzgeschichte ‚Kahlschnitt‘ zu schreiben, damit ich die Schülerin verbessern kann? (Proband:in 7-26)

Im Anschluss an den durch den Initialprompt generierten Text und eine weitere Prompt-Output-Parsequenz später teilt die Proband:in ihrem WKA ihre Schlussfolgerung mit:

Beleg IV.II

Okay, wie ich sehe, haben wir oensichtlich eine unterschiedliche Kurzgeschichte. Bei mir hat die Protagonisten weder einen Namen noch geht es um ihre Mitschüler. Es geht eher darum, dass sie mit ihrem ‘Ömchen’ am Tisch sitzt und diese nicht begeistert ist von dem Haarschnitt des Teenagermädchens. Auch ging es bei meiner Kurzgeschichte eher um die Reaktion der Mutter Irene, welche der Teenager ihr entlocken wollte. Kennst du diese Version der geschichte? (Proband:in 7-26)

Proband:in 7-26 zieht jedoch aus der sich im Verlauf des Dialogs einstellenden Erkenntnis, dass sie und das LLM offensichtlich unterschiedliche Kurzgeschichten „haben“, nicht die nahe-liegende Schlussfolgerung, dem LLM den Referenztext zur Verfügung zu stellen. Aber auch das LLM kommt der impliziten Erwartung an die Rolle eines kooperativen Schreibpartners nicht nach, wenn es die Schüler:in nicht darüber in Kenntnis setzt, dass für die gemeinsame Überarbeitung der Referenztext benötigt wird. In der Folge teilt die Proband:in dem Sprachmodell in dem Prompt (Beleg IV.II) sowie in den unmittelbaren Folgeprompts von ihr ausgewählte Informationseinheiten aus dem Referenztext mit, um das bis dato enttäuschende Resultat der Überarbeitung sukzessive zu verbessern.

Eine weitere Proband:in aus Gruppe C thematisiert in dem Initialprompt ihre Unsicherheit über das von ihr bei ChatGPT vorauszusetzende Wissen in Form einer Frage:

Beleg IV

Kennst du den Kurzgeschichte von Kirsten Boie ‚Kahlschnitt‘? (Proband:in 14-18)

Auch die irreführende Antwort von ChatGPT auf die Frage der Schüler:in soll zitiert werden:

Ja, ‚Kahlschnitt‘ ist eine Kurzgeschichte von Kirsten Boie, die sich mit Themen wie Identität, Veränderung und den Herausforderungen des Erwachsenwerdens auseinandersetzt. Die Geschichte handelt von einer Protagonistin, die mit ihrer äußeren Erscheinung und den Erwartungen ihrer Umgebung ringt.

Die Bejahung der Frage ist mindestens als irreführend zu bewerten. Allerdings enthält der ausgehend von der Frage generierte Output streng genommen keine unwahren Informationen. Diese sind jedoch so generisch, dass eine eng an den Referenztext angelehnte Überarbeitung der Inhaltsangabe_{UR} auf dieser Wissensbasis kaum möglich ist. Mit Verweis auf die obigen Ausführungen kann für diesen Fall konstatiert werden, dass das LLM „phantasiert“, wenn es die Frage, ob es die fragliche Kurzgeschichte von Kirsten Boie kenne, bejaht, was in dem vorliegenden Fall zu der Erzeugung falscher Annahmen über die Befähigung des WKA zur gemeinsamen Aufgabenbearbeitung bei der menschlichen Nutzer:in führt. Das LLM in der Rolle als kooperativer WKA erfüllt demnach nicht nur nicht die Erwartung der Nutzer:in, zwischen Wissen und Nicht-Wissen zu unterscheiden, sondern zudem camoufliert es diese Erwartungsenttäuschung mit Folgen für den weiteren gemeinsamen Schreibprozess.

Ich halte zwischenresümierend fest: Die zitierten Initialprompts (Beleg I bis III) sowie das engagierte, kleinteilige Nachprompten von Proband:in 13-24 (s. Tab. 2) zeigen an, dass ChatGPT die aus der Aufgabenstellung resultierende Zuweisung der Rolle als kooperativer WKA in Teilen erwartungswidrig erfüllt. Die Erwartung der Proband:innen an die Rolle des LLM als WKA lautet, dass dieses über eine Art Datenbank verfügt, in der literarische Texte hinterlegt sind, die zur Bearbeitung der mit der Rolle als Ko-Textüberarbeiter benötigt werden. Konkret lautet die Erwartung, dass das LLM die von allen drei Proband:innen explizit mit Titel, Autorin-Name und Veröffentlichungsdatum identifizierte Kurzgeschichte als textuelle Ressource abgespeichert hat. Allen Proband:innen in Gruppe C ist gemeinsam, dass sie, trotz zum Teil erkennbarer Frustrationen mit dem vom LLM generierten textuellen Output zur Überarbeitung der Inhaltsangabe_{UR}, ihrem WKA den Referenztext nicht zur Verfügung stellen. Allerdings macht das LLM in seiner Rolle als WKA den Proband:innen auch an keiner Stelle des Dialogs deutlich, dass für die gemeinsame Überarbeitung die Zurverfügungstellung des Referenztextes erforderlich ist, womit die implizite Erwartung an die Rolle eines kooperativen Ko-Aktanten enttäuscht wird und man unterstellt, dass das allgemeine Kooperationsprinzip i. S. v. Grice von den menschlichen Nutzer:innen auch im Dialog mit ChatGPT als geltend vorausgesetzt wird. Lediglich im Dialog zwischen Proband:in 17-24 und ChatGPT zeigt sich an einer Stelle (s. Beleg IV.II), dass 17-24 die Enttäuschung ihrer Erwartungen über das der KI unterstellte Wissen registriert.

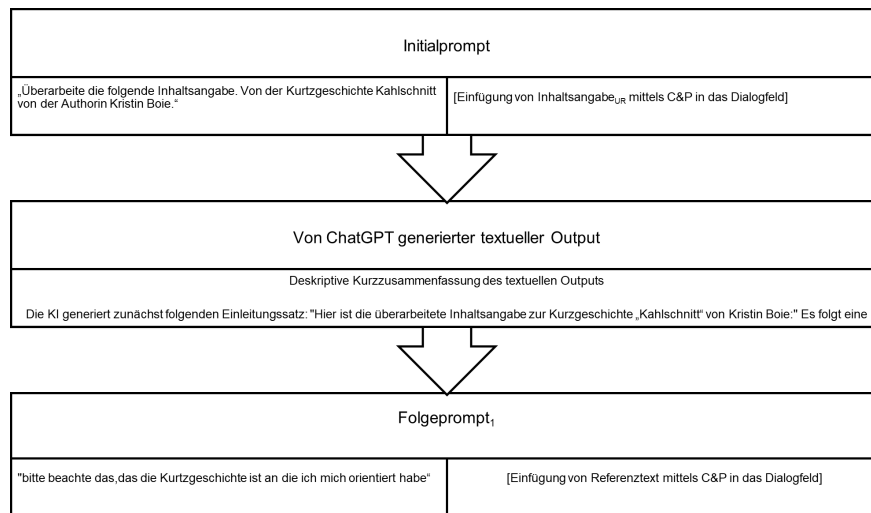
Ich komme zu Gruppe D: Von den 23 Proband:innen dieser Gruppe teilen 13 dem LLM bereits mit ihrem Initialprompt den Referenztext sowie die Inhaltsangabe_{UR} mit (Gruppe D1). Die Mitteilung erfolgt entweder mittels Copy and Paste, oder die Anhangfunktion von ChatGPT wird genutzt, mit der das ganze Dokument (.doc), ein Screenshot von Ausschnitten

des Dokumentes oder das mit dem Tablet ganz oder in Teilen abfotografierte Aufgabenblatt über das Büroklammer-Symbol im Dialogfeld hochgeladen und dem WKA auf diese Weise zur Verfügung gestellt wird. Der Aufgabenzettel enthält:

- Inhaltsangabe_{UR},
- Referenztext (Kurzgeschichte *Kahlschnitt*) und
- Aufgabenstellung.

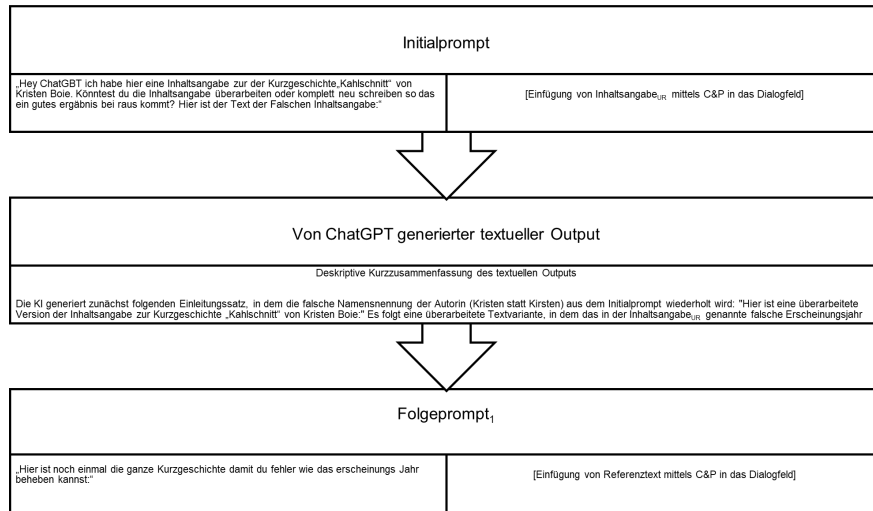
Die Motive der Proband:innen für die Entscheidung, dem LLM die beiden für die Überarbeitung erforderlichen Texte auf die eine oder andere Weise zur Verfügung zu stellen, können mithilfe der vorliegenden Daten nicht ermittelt werden. Der pragmatische Nutzen der Mitteilung des gesamten Aufgabenblattes – entweder mittels Copy and Paste oder durch Nutzung der Anhangfunktion – liegt allerdings auf der Hand: Das LLM wird auf diese Weise zum Delegierten, dem die Gesamtheit der für die Überarbeitungsaufgabe erforderlichen Informationen zur Verfügung gestellt werden. Für eine pragmatische Interpretation der Aufgabe von mindestens fünf der 13 Proband:innen dieser Untergruppe von Gruppe D spricht, dass diese im Anschluss an den Initialprompt keinen einzigen Folgeprompt formulieren; der generierte Output des LLM wird kommentarlos als „Lösung“ übernommen.

Abschließend wird ein Blick auf die zehn Proband:innen umfassende Untergruppe der Gruppe D geworfen, die dem LLM die für die Bearbeitung der Aufgabe erforderlichen Texte erst im Verlauf des Dialogs zur Verfügung stellen (Gruppe D2). Aus einigen der schriftlichen Dialoge lässt sich rekonstruieren, wie die Proband:innen im Verlauf der gemeinsamen Aufgabenbearbeitung mit dem LLM ihre anfänglichen Vorstellungen von dem verfügbaren Wissen des LLM revidieren und ihr Vorgehen anpassen (Belegsequenz I):



Belegsequenz I – Anfangssequenz von Proband:in 3-24

Eine weitere Anfangssequenz indiziert ebenfalls das Eintreten epistemischer Effekte im Verlauf des Dialogs (Belegsequenz II):



Belegsequenz II – Anfangssequenz Proband:in 12-26

Beide Proband:innen teilen dem LLM im Initialprompt die zur Überarbeitung vorliegende Inhaltsangabe mit, bevor sie in ihren Folgeprompts₁ den Referenztext nachliefern. Die in dem Folgeprompt von Proband:in 3-24 verwendete aufmerksamkeitssteuernde Formulierung (*bitte beachte...*) sowie die katadeiktische Konstruktion (*hier nochmal*) in Verbindung mit der Antezedens-Konsequenz-Konstruktion (*damit*) in dem Folgeprompt von Proband:in 12-26 deuten darauf hin, dass beide Schüler:innen aus dem von ihrem Initialprompt evozierten Output die richtigen Schlussfolgerungen ziehen: Ihre Präsupposition, dass ihr WKA die Kurzgeschichte kennt, stellt sich als unzutreffend heraus; in der Konsequenz wird der Referenztext dem LLM im Folgeprompt mitgeteilt.

Besonders deutlich wird die erwartungswidrige Rollenperformanz des LLM als kooperativer Schreibpartner in der folgenden Anfangssequenz (s. Belegsequenz III).

PO Proband 01 Proband 01
schreib nh inhaltsangabe zu kurzgeschichte kahlschnitt

GPT-4o-mini
Inhaltsangabe zu "Kahlschnitt" von

Die Kurzgeschichte "Kahlschnitt" von Friedrich Ani handelt von einem Frisör, der in seinem Salon eine eigenartige Begegnung mit einem Kunden hat. Der Kunde, ein älterer Mann, kommt zu ihm, um sich die Haare schneiden zu lassen. Während des Haarschnitts entwickelt sich ein intensives Gespräch zwischen dem Frisör und dem Kunden.

Der Frisör ist ein nachdenklicher und sensibler Mensch, der die Geschichten seiner Kunden schätzt und ihnen zuhört. Der ältere Mann hingegen ist von seinen Erinnerungen und seiner Vergangenheit geprägt. Im Verlauf des Gesprächs offenbart er, dass er mit Verlust und Trauer konfrontiert ist. Er spricht über seine verstorbene Frau und die Einsamkeit, die ihn plagt.

Die Atmosphäre im Salon ist von einer melancholischen Stimmung geprägt, die durch die Gespräche und die Gedanken des Frisörs verstärkt wird. Der Kahlschnitt, den der Frisör durchführt, wird zum Symbol für den Neuanfang und die Veränderung im Leben des Kunden. Am Ende der Geschichte bleibt der Leser mit einem Gefühl der Nachdenklichkeit zurück, da die Themen von Verlust, Einsamkeit und der Suche nach Identität auf berührende Weise behandelt werden.

Insgesamt thematisiert "Kahlschnitt" die zwischenmenschlichen Beziehungen und die Bedeutung von Kommunikation in Zeiten der Trauer.

PO Proband 01 Proband 01
von kirsten boie

Belegsequenz III – Anfangssequenz von Proband:in 1-18

Die initial verwendete Formulierung indiziert, dass Proband:in 1-18 die Kurzgeschichte, die der Inhaltsangabe_{UR} zu Grunde liegt, als dem Sprachmodell bekannt voraussetzt. Im Anschluss an den Textvorschlag der KI mit der Überschrift „Inhaltsangabe zu *Kahlschnitt* von X Y“, in der der Name der Probandin als Autorin-Name konjektiert wird (also: „Inhaltsangabe zu *Kahlschnitt* von [Vor- und Nachname der Proband:in 1-18]“),⁶ versucht die Proband:in im Folgeprompt mit der syntaktisch eingebetteten Erwähnung des textsortentypischen Ausdrucksmusters [von [Kirsten Boie]_{Autorin-Name}], die vom LLM kreierte Inhaltsangabe kleinteilig nachzubessern. Im Fortgang des Dialogs erkennt die Proband:in zudem die Notwendigkeit, dem LLM die Inhaltsangabe_{UR} zur Verfügung stellen zu müssen und instruiert das LLM mit dem Imperativsatz: *verbesser den text*. Einen weiteren Prompt und eine weitere Enttäuschung über den generierten Output später gelangt die Schüler:in zu der Erkenntnis, dass das LLM für die Überarbeitung der Inhaltsangabe_{UR} auch den Referenztext benötigt.

Diese drei aus dem vorliegenden Dialogkorpus ausgewählten Anfangssequenzen geben – wie die bereits zuvor zitierten Belege auch – Hinweise auf die Erwartungen und Annahmen, die Proband:innen in Bezug auf das LLM als WKA haben. Mit Ausnahme von Beleg IV wird in allen zur Veranschaulichung ausgewählten Initialprompts präsupponiert, dass das LLM die Kurzgeschichte von Kirsten Boie wie in einer Art Datenbank abgespeichert und abrufbereit hat. In den ausgewählten Chats setzen die Proband:innen zu Beginn des Dialogs stillschweigend voraus, dass das LLM in der Rolle eines kooperativen WKA die Nutzer:innen über das Fehlen von für die Bearbeitung der Aufgabe erforderlichen Wissens informiert. Lediglich in den ausgewählten Anfangssequenzen der

⁶ Der Name der Proband:in wurde in der Darstellung Belegsequenz III unkenntlich gemacht.

zehn Proband:innen der Untergruppe D2 sowie in den Prompts von Proband:in 7-26 aus Gruppe C finden sich Hinweise darauf, dass die Erwartungen an die Rolle des LLM als kooperativen WKA im Verlauf des Dialogs nicht nur enttäuscht werden, sondern dass diese Enttäuschungen von den Proband:innen auch bemerkt werden. Indiziert werden diese Erwartungsenttäuschungen durch die seitens der Proband:innen veranlassten Nachbesserungen und Ergänzungen.

4. Schlussfolgerungen

Abschließend werden die weiter oben formulierten Forschungsinteressen (a. und b.) mithilfe der vorgestellten Analyseergebnisse und bezogen auf die oben eingeführten theoretischen Konzepte „Rolle“ und „Person“ und die damit assoziierten Annahmen und Erwartungsbündel beantwortet: Es sollten die Erwartungen und Annahmen rekonstruiert werden, die Schüler:innen von dem LLM ChatGPT einerseits als Person, andererseits als Träger der Rolle eines kooperativen Schreibpartners haben.

Die mithilfe konversationsanalytischer und sprachwissenschaftlicher Methoden durchgeführten Analysen ausgesuchter Dialoge einiger Proband:innen mit dem LLM liefern deutliche Hinweise darauf, dass die Proband:innen bei ChatGPT in der Rolle eines kooperativen WKA das Vorhandensein von für die Aufgabenbearbeitung relevantem Wissen voraussetzen. Konkret lassen die zitierten Prompts und Dialogsequenzen die Schlussfolgerung zu, dass die Proband:innen die Kenntnis der Kurzgeschichte auf Seiten des LLM präsupponieren. Obwohl eine solche Annahme in den vorliegenden Dialogen an keiner Stelle explizit zum Ausdruck gebracht wird, kann als allgemein geltende implizite Erwartung an die Rolle von Ko-Aktanten bei kollaborativer Textarbeit Folgendes angenommen werden: Wenn im Prozess der gemeinsamen Bearbeitung einer Aufgabe zum Schreiben bzw. Überarbeiten eines Textes für Ko-Aktant_z erkennbar wird, dass ein von Ko-Aktant_x bei Ko-Aktant_z für relevant gehaltenes, präsupponiertes Wissensselement fehlt, dann wird im Sinne eines allgemeinen und von den Ko-Aktanten als verbindlich geltend angenommenen Kooperationsprinzips von Ko-Aktant_z erwartet, sein Nichtwissen Ko-Aktant_x explizit mitzuteilen. Diese für die erfolgreiche gemeinsame Aufgabenbearbeitung wechselseitig vorauszusetzende Rollenerwartung, nämlich zwischen Wissen und Nichtwissen zu unterscheiden und diese Unterscheidung dem Ko-Aktanten auch mitzuteilen, wird von dem LLM in den gezeigten Fällen nicht geleistet. Warum ist das so? Zunächst ist einzuräumen, dass das in der Studie für die Datenerhebung verwendete LLM über kein Nutzerprofil der Proband:innen verfügt, so dass eine aus wiederholter Nutzung resultierende Anpassung seitens des Sprachmodells an deren individuellen Voraussetzungen und Leistungspotentiale nicht vorliegt. Aus meiner Sicht entscheidender ist aber ein anderer Grund. Weiter oben wurde darauf hingewiesen, dass der Person ChatGPT qua Training die Attribute „sanftmütig, ehrlich und hilfsbereit“ (Kessel *et al.* 2025, 38) zugewiesen werden. Die Hilfsbereitschaft des Sprachmodells zeigt sich u. a. daran, dass es auch formal defekte und kryptische Prompts verarbeitet (s. o.) und eine auf Grundlage statistischer Wahrscheinlichkeiten berechnete Fortsetzung des Prompts generiert. Die Kehrseite dieser den Sprachmodellen aufwändig antrainierten Hilfsbereitschaft ist ein Modell, „welches Fragen beantwortet, statt weitere Fragen hinzuzufügen“, weshalb es eben u. a. die Attribute „harmless, honest and helpful“ verdient (39). Weil die Person ChatGPT nicht anders kann, als sich so zuvorkommend und hilfsbereit zu verhalten, kann es die Rolle eines kooperativen Schreibpartners kaum oder nur sehr bedingt erfüllen. Die der Person ChatGPT antrainierten Attribute stehen in einem Spannungsverhältnis zu der Rollenerwartung als kooperativer Ko-Aktant, weil zu dieser Rolle auch das Formulieren von Rückfragen, das Vorbringen von kritischem Feedback sowie die Reflexion des eigenen Tuns und Könnens gehören (Stichwort „Rollendistanz“).

Aus dem Voranstehenden lassen sich die Herausforderungen genauer bestimmen, mit denen Schülerinnen konfrontiert sind, die gemeinsam mit ChatGPT einen Text zu überarbeiten haben. Für einen erfolgreichen und möglichst arbeitsökonomischen Arbeitsprozess benötigen die Nutzer:innen eines Sprachmodells einerseits Kenntnisse darüber, welche Art von Wissen das LLM überhaupt verfügbar machen kann, und andererseits, wie dieses Wissen im Modell repräsentiert wird und über die Chatfunktion abgerufen werden kann. Für das gezielte Abrufen von Wissen mithilfe von Prompts benötigen die Nutzer:innen ferner Kenntnisse über die dem Sprachmodell zuweisbaren Rollen und die Art und Weise, wie die mit einer sozialen Rolle assoziierten Erwartungen von dem Rollenträger ChatGPT realisiert werden können. In diesem Kontext müssen die jugendlichen Nutzer:innen auch und vor allem den in dem vorliegenden Artikel herausgearbeiteten Konflikt verstehen lernen, zwischen den Erwartungen an das LLM in der ihm zugewiesenen Rolle auf der einen Seite und auf der anderen Seite die an das LLM als Person gerichtete Erwartung, jederzeit unterstützend, fehlertolerant und – vermeintlich – hilfreich zu respondieren.

Für die kompetente, die eigene literale Entwicklung unterstützende Nutzung von ChatGPT bei der Überarbeitung von Texten – und sicherlich nicht nur dabei – benötigen die Nutzer:innen neben dem o. g. deklarativen Wissen über die Funktionsweise, die Potentiale und Grenzen des LLM produktive wie rezeptive Textkompetenzen. Insbesondere im Kontext des Textschreibunterrichts ist das kritische Lesen und Verstehen des mithilfe von Prompts evozierten textuellen Outputs als Bedingung für die zielgerichtete Weiterführung des Dialogs mit dem LLM zu verstehen. Und in der Formulierung von Folgeprompts manifestiert sich das Verstehen des von dem LLM entweder unmittelbar zuvor generierten textuellen Outputs und/oder das Verstehen gleich mehrerer von dem LLM im Dialog zuvor generierter Antworten.

Literatur

- Baecker, Dirk. 2005. *Form und Formen der Kommunikation*. Frankfurt am Main: Suhrkamp.
- Becker-Mrotzek, Michael, und Ingrid Böttcher. 2012. *Schreibkompetenz entwickeln und beurteilen*. 4. Auflage. Berlin: Cornelsen.
- Boscolo, Pietro, and Suzanne Hidi. 2007. "The Multiple Meaning of Motivation to Write". In *Writing and Motivation*, edited by Idd., 1-14. Oxford: Elsevier.
- Dahrendorf, Ralf. 2006. *Homo Sociologicus. Ein Versuch zur Geschichte, Bedeutung und Kritik der Kategorie der sozialen Rolle*. 16. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dix, Annika. 2017. *Berichte und Berichten als didaktische Gattungen. Eine Textform zwischen Erwerb und schulischer Norm*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Gätje, Olaf, und Tobias Weindel. 2025. „Schreibend und lesend Texte schreiben mit dem Writing-Ko-Aktant ChatGPT“. In *DeutschGPT – Deutschunterricht im Dialog mit Künstlicher Intelligenz*, herausgegeben von Hans-Georg Müller und Maurice Fürstenberg. Berlin: Frank & Timme.
- Goffman, Erving. 1973. *Interaktion: Spaß am Spiel. Rollendistanz*. München: Piper.
- Hayes, John R. 2012. "Modeling and Remodeling Writing". *Written Communication* vol. 29, no. 3: 369-88.
- Hirschauer, Stefan. 2016. „Verhalten, Handeln, Interagieren. Zu den mikrosoziologischen Grundlagen der Praxistheorie“. In *Praxistheorie. Ein soziologisches Forschungsprogramm*, herausgegeben von Hilmar Schäfer, 45-68. Bielefeld: transcript. doi: 10.1515/9783839424049-003.
- Kessel, Thomas, Alexander Brandt, Jonas Offermatt, et al. 2025. *ChatGPT und Large Language Models? Klare Antworten aus erster Hand*. Tübingen: UVK Verlag.
- Lehnen, Katrin. 2023. „Kooperatives digitales Schreiben in und außerhalb der Schule. Ko-Autorschaft und Textfeedback zwischen Vergemeinschaftung und automatisierter Textproduktion“. *Der Deutschunterricht* Nr. 5: 18-28.
- Luhmann, Niklas. 1987. *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Frankfurt am Main: Suhrkamp.

- . 1991. „Die Form ‚Person‘ “. *Soziale Welt* Bd. 42, Nr. 2: 166-75.
- . 1998. *Die Gesellschaft der Gesellschaft*. Frankfurt am Main: Suhrkamp.
- Navigli, Roberto, Simone Conia, and Björn Ross. 2023. “Biases in Large Language Models: Origins, Inventory, and Discussion”. *ACM Journal of Data and Information Quality* vol. 15, no. 2: 1-21. doi: 10.1145/3597307.
- Nicolas Gandalf, und Aylin Caliskan 2024. “A Taxonomy of Stereotype Content in Large Language Models”. *arXiv preprint* doi: arXiv:2408.00162.
- Reinert, Johannes. 2012. *Die Inhaltsangabe im Deutschunterricht. Eine Video- und Interviewstudie zum Umgang mit einer viel diskutierten Textform*. PhD Dissertation, Hildesheim: Universitätsbibliothek Hildesheim.
- Steinhoff, Torsten 2025. „Künstliche Intelligenz als Ghostwriter, Writing Tutor und Writing Partner. Zur Modellierung und Förderung von Schreibkompetenzen im Zeichen der Automatisierung und Hybridisierung der Kommunikation am Beispiel des Schreibens mit ChatGPT in der 8. Klasse“. In *Personale und funktionale Bildung im Deutschunterricht. Theoretische, empirische und praxisbezogene Perspektiven*, herausgegeben von Christian Albrecht, Jörn Brüggemann, Tabea Kretschmann, et al., 85-99. Stuttgart: Metzler.
- Tekin, Özlem 2023. „ChatGPT im Unterricht Deutsch als Fremdsprache“. *Alman Dilive Kültürü Araştırmaları Dergisi* Bd. 5, Nr. 2: 135-63. doi: 10.55143/alkad.1390420.
- Thompson, Adrian. 2024. *ChatGPT for Conversational AI and Chatbots: Learn How to Automate Conversations with the Latest Large Language Model Technologies*. Birmingham: Packt Publishing.
- Titscher, Stefan, Michael Meyer, Ruth Wodak, et al. (eds). 2002 [2000]. *Methods of Text and Discourse Analysis*. London-Thousand Oaks-New Delhi: Sage Publications.
- Vara, Vauhini. 2025. “ChatGPT May Be Polite, but It’s not Cooperating with You”. *The Guardian*, 13 May. <<https://www.theguardian.com/technology/ng-interactive/2025/may/13/chatgpt-ai-big-tech-cooperation>> (03/2026).
- Yin, Ziqi, Hao Wang, Kaito Horio, et al. 2024. “Should we Respect LLMs? A Cross-Lingual Study on the Influence of Prompt Politeness on LLM Performance”. In *Proceedings of the Second Workshop on Social Influence in Conversations (SICoN 2024)*, edited by James Hale, Kushal Chawla, and Muskan Garg, 9-35. <<https://aclanthology.org/2024.sicon-1.2/>> (03/2026).

