

**Einsatz von generativer KI in der Hochschullehre:
Kritische Reflexion und nachhaltiger Lernerfolg durch eine
adaptierte *Comparative Judgement*-Methode**

Sarah Schwellenbach

2025

Ansprechpartner/in: Dr. Sarah Schwellenbach-Siwonia sarah.schwellenbach@uni-bielefeld.de	Fach/Fachgebiet: <i>Fakultät für Linguistik und Literaturwissenschaft</i> Lehreinheit Romanistik Universität Bielefeld	Zentrale Lehr-Themen: <i>(Klein)Gruppenarbeit</i> <i>Überfachliche Kompetenzen,</i> <i>Schlüsselkompetenzen</i> <i>Verbindung Theorie-Praxis</i>
Autorin: Sarah Schwellenbach		

Abstract (Zusammenfassung des Lehrprojekts)

Das Lehrprojekt untersuchte, wie der Einsatz von generativer künstlicher Intelligenz (KI) in der Hochschullehre didaktisch derart gestaltet werden kann, dass er sowohl den Lernerfolg als auch eine kritisch-reflektierte Nutzung durch Studierende fördert. Zielgruppe waren Studierende eines sprachwissenschaftlichen Seminars, die durch diverse methodische Zugänge an eine fachlich fundierte Analyse von Sprachmaterial herangeführt wurden. Erprobt wurde eine didaktisch adaptierte und methodisch modifizierte Form der *Comparative Judgement*-Methode (CJ-Methode), die als kritische Vergleichsanalyse zwischen sprachwissenschaftlich erarbeiteten und KI-generierten Analysen umgesetzt wurde. Während *Comparative Judgement* in der Assessmentforschung üblicherweise der Leistungsbewertung dient, wurde die Methode im Projekt zur Förderung von Reflexion, Fehlererkennung und Urteilskompetenz im Umgang mit KI weiterentwickelt. Die mithilfe vorstrukturierter Prompts generierten KI-Texte basierten auf derselben Aufgabenstellung und demselben bereitgestellten Vorwissen wie die studentischen Analysen und wurden über BIKI mit den Studierenden geteilt. Von mehreren erprobten Varianten wurden zwei Szenarien vertiefend betrachtet: In Variante A analysierten die Studierenden nach der Fachtextlektüre selbstständig ein sprachliches Korpus und bewerteten anschließend die KI-Ausgabe. In Variante B erfolgte die Präsentation der KI-Ausgabe vor der Fachtextlektüre, die Bewertung danach. Dabei zeigte sich ein deutlicher *order effect*: Wurden die Studierenden zuerst zur eigenen Analyse angeleitet (Variante A), bewerteten sie die KI-Ausgabe kritischer als in Variante B. Besonders wirkungsvoll erwies sich die Methode in Variante A, in der die Studierenden Halluzinationen der KI nicht nur erkannten, sondern auch fundiert analysierten und kritisch reflektierten. Dies führte zu einem vertieften Verständnis der Fachinhalte und einer nachhaltigen Stärkung ihrer Urteilskompetenz im Umgang mit KI-gestützten Texten. Die Ergebnisse zeigen, dass der KI-Einsatz den Lernerfolg fördern kann, wenn er didaktisch eingebettet, an Lernzielen orientiert und über strukturierte Prompts angeleitet wird. Im direkten Vergleich erkannten die Studierenden, dass die eigenständige wissenschaftliche Analyse gegenüber KI-generierten Ausgaben insbesondere hinsichtlich Validität, analytischer Genauigkeit und fachlich-argumentativer Tiefe deutlich überlegen ist. Die Auseinandersetzung mit beiden Zugängen förderte zudem ihr Bewusstsein für wissenschaftliche Qualität und Verantwortung im Umgang mit digitalen, insbesondere KI-gesteuerten, Werkzeugen.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

1. Übergeordnete Fragestellung des Lehrprojekts

Das Lehrprojekt untersuchte, mit welchen didaktischen Methoden generative Künstliche Intelligenz (KI) in der Hochschullehre derart eingesetzt werden kann, dass sie sowohl das Erreichen fachlicher Lernziele unterstützt als auch die Reflexions- und Urteilskompetenz der Studierenden im Umgang mit KI stärkt. Im Zentrum stand die Frage, wie KI-gestützte Lernprozesse so gestaltet werden können, dass Studierende nicht nur fachliche Konzepte und Analyseverfahren verstehen und anwenden, sondern auch befähigt werden, KI-generierte Informationen kritisch zu prüfen, einzuordnen und hinsichtlich ihrer fachlichen Qualität zu beurteilen.

2. Besondere Herausforderungen

Bei der Konzeption des Lehrprojekts wurden sowohl die spezifischen Anforderungen im Umgang mit KI als auch die strukturelle Offenheit eines hybriden Lehrformats in Kleingruppen berücksichtigt. Die für das Lehrprojekt didaktisch bedeutendste Herausforderung lag im Umgang mit der bekannten Problematik des Halluzinierens, das heißt der Erzeugung scheinbar kohärenter, jedoch inhaltlich unzutreffender Aussagen von KI-Chatbots. Um dieser Herausforderung didaktisch derart zu begegnen, dass die Überprüfbarkeit der KI-Inhalte durch Studierende grundsätzlich gewährleistet war, wurden Anwendungssituationen entwickelt, in denen die Studierenden das relevante Wissen gezielt durch Fachlektüre erwarben und die KI-Ausgaben auf dieser Grundlage – in unterschiedlicher Reihenfolge von Lektüre und KI-Begegnung – kritisch beurteilten. Zugleich wurde untersucht, wie sich diese unterschiedliche Abfolge auf die Bewertung von KI-Ausgaben auswirkte. Eine besondere Rolle spielte dabei die gezielte Gestaltung und gemeinsame Nutzung von Prompts: Über das universitäts-eigene Tool BIKI wurden durch die Lehrkraft erstellte Prompts mit identischer Aufgabenstellung und Wissensgrundlage mit den Studierenden geteilt, um eine methodisch kontrollierte und fachlich fundierte Vergleichbarkeit zwischen Studierenden- und KI-Leistungen zu gewährleisten. In einer Prompt-Werkstatt wurden zudem grundlegende Prinzipien des kritisch-reflektierten KI-Einsatzes erarbeitet und Studierende angeleitet, Prompts derart zu formulieren, dass sie die Transparenz, Überprüfbarkeit und fachliche Anschlussfähigkeit der KI-Ausgaben begünstigten.

Das hybride Format in Kleingruppen, dessen didaktische Besonderheiten in einem separaten Lehrprojekt ausführlicher adressiert wurden, erforderte zudem eine besonders flexible Planung. Da die Zahl der digital zugeschalteten Teilnehmenden variierte, wurden die Gruppenformate situativ angepasst und je nach Konstellation der Teilnehmenden entsprechend als Präsenz-, Onlinegruppen (in separaten Breakout-Räumen) und/oder als hybride Gruppenformate mit Präsenz- und DTEN-Teilnehmenden umgesetzt. Unterschiedliche digitale Tools (wie Padlet, Zoom-Whiteboard und Mentimeter) ermöglichten eine gleichberechtigte und aktive Einbindung aller Studierenden unabhängig vom Lernort. Diese Flexibilität wurde gezielt genutzt, um Lernräume methodisch differenziert zu gestalten und kollaborative Prozesse zu fördern. Durch einen Raumwechsel ergaben sich zudem bessere akustische Bedingungen und eine der Gruppengröße angemessene Lernumgebung, die sich im Verlauf des Seminars als besonders förderlich für den Einsatz des DTEN-Systems erwies.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

3. Ziele

3a. Lehrziele

Ziel der Lehrveranstaltung war es, die Studierenden im Sinne einer kompetenzorientierten Hochschullehre dazu zu befähigen, sprachwissenschaftliche Analyseverfahren nicht nur anzuwenden, sondern die Qualität und Aussagekraft eigener sowie KI-generierter Analysen kritisch zu reflektieren und zu bewerten. Im Zentrum stand die Förderung folgender Kompetenzen:

- *Fachkompetenz*: Anwendung grundlegender sprachwissenschaftlicher Analyseverfahren zur Untersuchung lautlicher, morphosyntaktischer und lexikalisch-semantischer Merkmale sowie charakteristischer Strukturmerkmale von Varietäten, Misch- und Kreolsprachen anhand authentischer Sprachdaten.
- *Methodenkompetenz*: Strukturierte Auswertung, Reflexion und Vergleich eigener und KI-generierter Analyseprozesse (CJ-Methode).
- *Urteilskompetenz*: Kritische Bewertung von KI-Ausgaben im Hinblick auf fachliche Richtigkeit, wissenschaftliche Anschlussfähigkeit und fachlich-argumentativer Tiefe.
- *Critical Digital Literacy / Critical AI Literacy*: Entwicklung eines reflektierten, handlungsorientierten Umgangs mit KI im akademischen Kontext – einschließlich der Fähigkeit, Prompts gezielt zu gestalten, Ergebnisse kritisch zu prüfen und digitale Werkzeuge verantwortungsvoll zu nutzen.
- *Transferkompetenz*: Übertragung der erworbenen Analyse-, Reflexions- und Bewertungskompetenzen auf weitere fachliche Kontexte (wie Analyse neuer Varietäten, kritische Bewertung von Lehrwerken aus sprachwissenschaftlicher Perspektive), auf zukünftige Anforderungen im Studium und in der Forschung sowie auf außerwissenschaftliche berufliche Handlungsfelder.

3b. Antizipation der studentischen Perspektive

Die Studierenden profitierten in mehrfacher Hinsicht von dem Lehrprojekt: Sie vertieften nicht nur ihr sprachwissenschaftliches Wissen, sondern erwarben zugleich Kompetenzen für einen reflektierten und verantwortungsvollen Umgang mit generativer KI in Studium, Forschung und berufspraktischen Kontexten. Sie lernten, KI-Ausgaben systematisch zu hinterfragen, fachlich zu prüfen, einzurichten und mit eigenen Analysen zu vergleichen. Im direkten Vergleich erkannten sie, dass KI-generierte Ausgaben die fachliche Qualität wissenschaftlicher Arbeit nicht erreichen können und vielmehr einer sorgfältigen Prüfung und Einordnung bedürfen. Die Gegenüberstellung ihrer eigenen Analysen mit den KI-Ausgaben ermöglichte ihnen unter Anleitung eine vertiefte kritisch-reflexive Auseinandersetzung, in der sie ihre analytischen Fähigkeiten und ihre Urteilskompetenz gezielt schärfen konnten. Dies stärkte sowohl ihr Verständnis für wissenschaftliche Methoden als auch ihre Selbstwirksamkeit im Umgang mit digitalen, insbesondere KI-gestützten, Werkzeugen und ihre Fähigkeit, zwischen algorithmisch erzeugtem Wissen und fachlich begründeten Urteilen zu unterscheiden. Auf diese Weise wurden Schlüsselkompetenzen wie *Critical Digital Literacy* und *Critical AI Literacy* sowie überfachliche Kompetenzen, wie Reflexionsfähigkeit, Urteilskompetenz und ein Bewusstsein für wissenschaftliche Integrität und Verantwortung im Umgang mit KI, gefördert. Diese



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

Kompetenzen sind sowohl für das wissenschaftliche Arbeiten als auch für außerwissenschaftliche Berufsfelder von grundlegender Bedeutung.

3c. Lernziele, Aktivitäten und Überprüfung

Lernziele

Die Studierenden sollten durch das Lehrprojekt in die Lage versetzt werden,

- grundlegende Konzepte und Analyseverfahren der Sprachwissenschaft sicher anzuwenden,
- eigene Analysen zu authentischem Sprachmaterial (u.a. Korpora) systematisch zu erstellen,
- KI-generierte Analyseergebnisse hinsichtlich fachlicher Richtigkeit, Tiefe und Nachvollziehbarkeit kritisch zu beurteilen,
- grundlegende Konzepte generativer KI (wie Prinzipien probabilistischer Textgenerierung, Halluzinationen, Prompt-Sensitivität und Grenzen der Validität) zu verstehen,
- eine *Critical AI Literacy* zu entwickeln

Lernaktivitäten

- Vorbereitende Lektüre ausgewählter wissenschaftlicher Texte
- Interaktive studentische Referate mit kooperativen Arbeitsphasen
- Erstellung eigener Korpusanalysen (insbesondere zu lautlichen, morphosyntaktischen und lexikalisch-semantischen Merkmalen der jeweiligen Varietät, Kreol- oder Mischsprache)
- Gezielte Erprobung und Analyse der Wirkung unterschiedlicher Prompts (Prompt-Werkstatt) zur Förderung von Reflexionsfähigkeit und Verständnis der Steuerungsmechanismen generativer KI
- kritisch-vergleichende Gegenüberstellung eigener Analysen mit KI-Ausgaben
- Gezielte weitere Befragung von BIKI zur Nachvollziehbarkeit und Begründung der Analyseentscheidungen
- Kritisch-reflexive Bewertung der Qualität und wissenschaftlichen Anschlussfähigkeit der KI-Ausgaben
- Nutzung kollaborativer digitaler Tools (u. a. Padlet, Zoom-Whiteboard, Mentimeter, BIKI)

Überprüfung und Rückmeldung

- Teaching Analysis Poll (TAP) zur qualitativen Rückmeldung durch Studierende
- Fakultätsübergreifende und fachspezifische kollegiale Hospitation mit Diskussion
- Ergebnissicherung über Padlet- und Zoom-Whiteboard-Beiträge
- Vergleichende Diskussion der KI- und Studierendenanalysen
- Beobachtung und Analyse der Argumentationsprozesse sowie der Fähigkeit der Studierenden, Fehler und Unstimmigkeiten in KI-Ausgaben kritisch zu identifizieren und fachlich zu begründen



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

Sechs-Lernebenen-Checkliste

Lernebene	Lernziele	Lern- und Studienaktivitäten	Rückmeldungen und Prüfungen
1. Fachwissen	Studierende verstehen grundlegende Begriffe und Konzepte der Varietäten- und Kontaktlinguistik sowie die linguistischen, sozialen und historischen Rahmenbedingungen, unter denen Varietäten, Misch- und Kreolsprachen entstehen. Sie können ihre jeweiligen charakteristischen Merkmale benennen und erläutern und kennen die wesentlichen Grundlagen wissenschaftlichen Arbeitens. Darüber hinaus verstehen sie grundlegende Konzepte generativer KI (wie Prinzipien probabilistischer Textgenerierung, Halluzinationen, Prompt-Sensitivität und Grenzen der Validität).	<ul style="list-style-type: none"> Vorbereitende Lektüre und gemeinsame Erarbeitung wissenschaftlicher Texte zur Varietäten- und Kontaktlinguistik sowie zu den phonologischen, morphosyntaktischen und lexikalischen Merkmalen ausgewählter Varietäten, Misch- und Kreolsprachen, interaktive Referate. 	<ul style="list-style-type: none"> Mündliches Feedback durch die Lehrkraft, Diskussionen im Plenum zur Verständnissicherung.
2. Anwendung	Studierende analysieren sprachliche Daten aus einem Korpus und vergleichen eigene Analysen mit KI-generierten Ausgaben (CJ-Methode). Sie wenden ihr linguistisches Fachwissen an, um KI-gene-	<ul style="list-style-type: none"> Gruppenarbeit zur Analyse sprachlicher Daten aus einem Korpus mit arbeitsteiliger Fokussierung (wie phonologische, lexikalistische und morphosyntaktische Merkmale der Varietät), 	<ul style="list-style-type: none"> Formatives Feedback durch Lehrkraft und Peers, Diskussion der im Padlet festgehaltenen Ergebnisse im Plenum, Reflexive Überprüfung des Fachverständnisses und der eigenen



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

	<p>rierte Ausgaben systematisch zu überprüfen. Sie erkennen, begründen und korrigieren Halluzinationen oder Fehlklassifikationen der KI auf Grundlage fachwissenschaftlicher Kriterien. Zudem entwickeln sie die über BIKI geteilten Prompts weiter und reflektieren deren Wirkung auf die Qualität und Nachvollziehbarkeit der KI-Ausgaben.</p>	<ul style="list-style-type: none"> • Ergebnissicherung der Gruppenarbeiten im Padlet, • Gruppenarbeit zur kritischen Beurteilung KI-gestützter Ausgaben (BIKI), • eigenständige Erstellung und Weiterentwicklung von Prompts. 	<p>Analyseentscheidungen im Vergleich zwischen studentischen Analysen und KI-Ausgaben (CJ-Methode),</p> <ul style="list-style-type: none"> • vergleichende Diskussion der Analyseergebnisse mit KI-Ausgaben im Plenum.
3. Verknüpfung	<p>Studierende erkennen Zusammenhänge zwischen Sprachkontakt, Varietätenbildung, Sprachwandel und Identitätsbildung und übertragen ihr erworbene Wissen auf neue Varietäten und fachliche Kontexte (wie Analyse unbekannter Sprachdaten, Reflexion schulischer Lehrwerke und Unterrichtsgestaltung). Zudem reflektieren sie den Stellenwert digitaler Bildung und entwickeln Transferperspektiven für den Einsatz von KI im Studium und Fremdsprachenunterricht in der Schule.</p>	<ul style="list-style-type: none"> • Kommentierte Diskussionen im Plenum, • Analyse schulischer Lehrwerke und Transferfragen zu Einsatzmöglichkeiten im Schulunterricht, • Anwendung des Analyseansatzes auf neue Sprachbeispiele, • Analyse und Bewertung bestehender Prompt-Anleitungen aus Schulen im Rahmen der Prompt-Werkstatt. 	<ul style="list-style-type: none"> • Kommentierte Diskussion im Plenum, • Transferaufgaben (Reflexionsfragen zu Unterrichtsbeispielen).



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

4. Menschliche Dimension	Studierende reflektieren ihre eigene Rolle im Lernprozess, arbeiten inklusiv und respektvoll im hybriden Setting zusammen.	<ul style="list-style-type: none"> • Kollaboratives Arbeiten in Präsenz-, Online- und hybriden Gruppen, • Peer-Feedback zur Zusammenarbeit, • Austausch zu Rollen und Teamdynamik. 	<ul style="list-style-type: none"> • Beobachtung der Gruppeninteraktion, • Peer-Feedback, • kurze mündliche Reflexionsimpulse.
5. Werte	Studierende entwickeln ein Bewusstsein für verantwortungsvollen KI-Einsatz sowie für Fragen der Sprachgerechtigkeit und Inklusion.	<ul style="list-style-type: none"> • Diskussionen zu Bias und KI-Ethik • Reflexion zu Sprachpolitik, sprachlicher Vielfalt, kulturellem Kontext und Identität von Sprechenden 	<ul style="list-style-type: none"> • Reflexionsrunden und Abschlussreflexion zu Wertfragen (Padlet, Zoom-Whiteboard).
6. Lernen, wie man lernt	Studierende lernen, KI gezielt als Werkzeug im Lernprozess zu nutzen, Prompts präziser und zielgerichtet zu formulieren, deren Qualität zu bewerten und selbstgesteuert zu arbeiten.	<ul style="list-style-type: none"> • Anleitung zur Prompt-Entwicklung, • Anwendung in BIKI, Auswertung der Resultate, • Selbstlernphasen nach der Sitzung. 	<ul style="list-style-type: none"> • Diskussion über Promptqualität, • Lernziel-Check („Was habe ich im kritischen Vergleich zur KI-Analyse erkannt?“), • Feedback zu Selbstlernstrategien.

4. Methodik

Das Lehrprojekt folgte einem konstruktivistischen Lernverständnis (vgl. Piaget 1975; von Glaserfeld 1997; Reich ⁵2012), bei dem Wissen nicht passiv aufgenommen, sondern aktiv durch die Lernenden selbst konstruiert wurde. In kooperativen Analysephasen, bei der Arbeit mit authentischem Sprachmaterial und in der kritisch-reflektierten Nutzung von KI-Anwendungen erarbeiteten sich die Studierenden Inhalte selbstständig und in angeleiteten Gruppen- und Diskussionsphasen.

Die Methode, die sich im Zuge der übergeordneten Fragestellung des Lehrprojektes als besonders geeignet für den kritisch-reflexiven Umgang mit KI herausstellte, war der Einsatz einer modifizierten *Comparative Judgement*-Methode (CJ-Methode), bei der die Studierenden ihre selbst erarbeiteten sprachwissenschaftlichen Analysen mit KI-generierten Analysen verglichen. Ursprünglich in der Assessmentforschung entwickelt, um durch wiederholte Paarvergleiche die Leistungsbewertung besonders reliabel und effizient zu gestalten (vgl. Thurstone 1927; Pollitt 2012), wurde die Methode im Projekt didaktisch umgedeutet und weiterentwickelt. Statt summativer Bewertung stand der kritische Vergleich zwischen wissenschaftlich erarbeiteten Analysen der Studierenden und KI-generierten Analysen im Mittelpunkt.

Um valide Vergleichsgrundlagen zu schaffen, wurden die KI-Ausgaben mithilfe umfangreich vorstrukturierter Prompts erzeugt, die auf derselben Aufgabenstellung und demselben von der



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](https://creativecommons.org/licenses/by-sa/4.0/) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

Lehrkraft bereitgestellten Vorwissen zu der jeweils untersuchten Varietät des Französischen, Misch- oder Kreolsprache basierte. Dieses bereitgestellte Vorwissen, welches auch konkrete Hinweise auf die von den Studierenden zu lesenden wissenschaftlichen Texten enthielt, diente der KI als erweiterte Grundlage für die Analyse des sprachlichen Korpus und stellte so eine zumindest annähernd vergleichbare Ausgangssituation zu den Studierenden her. Über das universitätseigene Tool BIKI wurden anschließend die von der Lehrperson erstellten Prompts und daraus resultierenden KI-Ausgaben den Studierenden zugänglich gemacht. Durch das Teilen des BIKI-Chats konnten die Studierenden die KI eigenständig weiter befragen, wodurch ein vertiefter Reflexionsprozess über die Analyseergebnisse angeregt wurde.

Im Rahmen des Projekts wurden zwei Varianten erprobt, die sich in der Reihenfolge der Präsentation unterschieden, um einen Einfluss eines möglichen *order effect* (vgl. Bruine de Bruin & Keren 2003) zu untersuchen. In Variante A erfolgte zunächst die Fachtextlektüre und die eigenständige Analyse der sprachlichen Daten aus dem Korpus, bevor die Studierenden die KI-Ausgabe bei gleicher Aufgabenstellung bewerteten. In Variante B wurde dagegen die KI-Ausgabe bereits vor der Fachtextlektüre präsentiert, die Bewertung der KI-Ausgabe erfolgte danach. Die Urteile im CJ-Verfahren orientierten sich an Kriterien wie fachliche Richtigkeit, Argumentationstiefe, Terminologietreue und Nachvollziehbarkeit. Ergänzend dazu bot eine Prompt-Werkstatt den Studierenden die Möglichkeit, die Wirkung von Prompts systematisch zu erproben, diese präziser zu formulieren, deren Wirkung zu evaluieren und die Qualität von KI-Ausgaben kritisch zu beurteilen.

5. Rollenreflexion

Das Lehrprojekt reflektiert ein Lehrverständnis, das auf partizipativer, forschungsnaher und reflexiver Hochschuldidaktik beruht. Anstelle einer reinen Wissensvermittlung stand die Begleitung von Lernprozessen im Vordergrund, die Räume für kritische Auseinandersetzung, Eigenverantwortung und nachhaltiges fachliches Verstehen eröffnete. Der Fokus lag dabei nicht allein auf der Vermittlung sprachwissenschaftlicher Fachinhalte, sondern ebenso auf der praktischen Anwendung sprachwissenschaftlicher Analyseverfahren auf authentisches Sprachmaterial sowie auf der Reflexion dieser Verfahren im Rahmen der Vergleichsanalyse zwischen den von den Studierenden eigenständig erstellten Analysen und den KI-generierten Texten, die auf vergleichbarer Wissensgrundlage basierten. Besonderes Augenmerk galt dabei einem didaktisch gerahmten, ethisch reflektierten und kritisch-dialogischen Umgang mit KI, der Studierende zu aktiver Befragung, Reflexion und Bewertung der KI-Ausgaben anregte. Die eingesetzten Methoden, wie insbesondere das vergleichende Arbeiten mit eigenen Analysen und KI-generierten Ausgaben, zielten darauf, Reflexionsfähigkeit, Urteilskompetenz und ein selbstbewusstes, aktives Lernverständnis zu fördern.

Das hybride Setting erforderte zudem eine bewusste Inklusion aller Studierenden, sowohl auf technischer als auch didaktischer Ebene. Das zugrunde liegende Lehrverständnis begreift Inklusion und Diversität dabei als Ressource, die durch Perspektivenvielfalt und diverse Lern- und Erfahrungs-zugänge Kreativität fördert und zu vertieftem, nachhaltigem Lernerfolg beiträgt.

Die Rolle der Lehrperson im Projekt bestand im Wesentlichen darin, Lernprozesse zu begleiten, didaktisch zu strukturieren sowie inklusiv und verantwortungsvoll zu gestalten. Dazu gehörte insbesondere, die Interaktion mit KI didaktisch und ethisch zu rahmen und Studierende zu einer reflek-



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

tierten und verantwortungsbewussten Nutzung digitaler Werkzeuge im wissenschaftlichen Kontext zu befähigen.

6. Evaluation und Feedback

Zur Sicherung der Qualität und zur Überprüfung der Lernwirksamkeit wurde das Lehrprojekt sowohl kollegial als auch studentisch evaluiert. Dabei kamen unterschiedliche Feedbackinstrumente zum Einsatz.

Im Rahmen einer sowohl fachspezifischen als auch fakultätsübergreifenden kollegialen Hospitation wurde insbesondere der didaktische Einsatz der KI im Seminar beobachtet und reflektiert. Die kollegialen Rückmeldungen bestätigten den didaktischen Mehrwert des gewählten Ansatzes im Rahmen der von der Lehrkraft didaktisch weiterentwickelten CJ-Methode. Besonders hervorgehoben wurden die zielgerichtete und über BIKI mit den Studierenden geteilte Prompt-Gestaltung durch die Lehrperson, die Einbettung des KI-Einsatzes in bereits erarbeitetes Fachwissen sowie die kritische Vergleichsanalyse von KI- und Studierendenleistungen als wirksame Reflexionsstrategie. Diese Form des KI-Einsatzes wurde von den hospitierenden Lehrkräften als besonders lernförderlich, methodisch durchdacht und didaktisch verantwortungsvoll bewertet.

Zur Erhebung der studentischen Perspektive wurde zur Semesterhalbzeit ein Teaching Analysis Poll (TAP) durch das Zentrum für Lehren und Lernen (ZLL) durchgeführt. Die Ergebnisse des TAP spiegelten ein deutlich erhöhtes Problembewusstsein der Studierenden im Umgang mit KI wider, insbesondere hinsichtlich der Fähigkeit, halluzinierte Inhalte als solche zu erkennen. Als besonders erkenntnisfördernd („Aha-Effekt“) beschrieben die Studierenden die Einsicht, den Ausgaben der KI nicht uneingeschränkt vertrauen zu können. Dieses Bewusstsein entstand insbesondere im Vergleich der KI-Antworten mit den eigenen wissenschaftlich erarbeiteten Analysen, bei dem sie auf ihr im Seminar erworbenes Wissen zurückgriffen und dieses kritisch zur Beurteilung der KI-Ausgaben anwandten.

Fazit zur Evaluation

Sowohl das kollegiale als auch das studentische Feedback bestätigte, dass der gezielte und strukturierte KI-Einsatz über BIKI, eingebettet in eine Lehrkonzeption, die die kritische Auseinandersetzung mit den KI-Ausgaben systematisch anleitet, das Lernen unterstützt und wesentlich zur Erreichung der Lernziele des Seminars beigetragen hat. Die Evaluation hebt zugleich die Bedeutung von geleiteten Prompt-Settings (wie dem Erstellen und Teilen von Prompts mit vorbereiteter Wissensgrundlage über BIKI), von vergleichenden Analyseformaten (wie dem Vergleich studentischer und KI-generierter Analysen mittels der modifizierten CJ-Methode) sowie von der kritisch-reflektierten Auseinandersetzung mit KI-Ausgaben und deren Halluzinationen hervor, die sich als tragfähige und wirksame didaktische Bausteine für die Hochschullehre erwiesen haben.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

7. Ergebnisse

Die Ergebnisse des Lehrprojekts zeigen, dass Künstliche Intelligenz (KI) das fachliche Lernen vor allem dann wirksam unterstützt, wenn sie in ein didaktisch fundiertes Lehrkonzept eingebettet ist, bei dem Studierende durch zuvor erarbeitetes Fachwissen in die Lage versetzt werden, fachbezogene KI-Ausgaben kritisch zu überprüfen und ihr Wissen im Rahmen einer reflektierenden Auseinandersetzung anzuwenden und zu vertiefen. Drei zentrale Erkenntnisse lassen sich hervorheben:

1. Wirksamkeit der modifizierten *Comparative Judgement*-Methode (CJ-Methode):

Die modifizierte CJ-Methode erwies sich als besonders geeignet, um die fachlichen Lernziele zu erreichen und die Reflexions- sowie Urteilskompetenz der Studierenden im Umgang mit KI zu stärken. Durch den systematischen Vergleich eigener mit KI-generierten Analysen entwickelten die Studierenden die Fähigkeit, Analysefehler der KI in Bezug auf die jeweils untersuchte sprachliche Varietät, Misch- oder Kreolsprache zu erkennen, ihre eigenen Analysen fachlich zu begründen und Qualitätskriterien wissenschaftlicher Arbeit gezielt anzuwenden. Dadurch wurden nicht nur Fachwissen und Analysekopetenz vertieft, sondern auch metakognitive Fähigkeiten sowie Urteils- und Reflexionskompetenz weiterentwickelt.

Zudem zeigte die Arbeit mit KI-generierten Texten eine im Vergleich zum üblichen Peer-Feedback veränderte Feedback-Dynamik: Studierende formulierten Rückmeldungen zu den KI-Texten wesentlich direkter, ehrlicher und analytisch präziser, als sie es im direkten Austausch mit Kommiliton*innen üblicherweise tun. Da die KI als nicht-menschliche Instanz wahrgenommen wurde, entfiel die sozial konforme Zurückhaltung und die damit verbundene Indirektheit in den Rückmeldungen (zum Konversationsprinzip der sozialen Konformität siehe auch Schwellenbach 2025), was eine kritischere Auseinandersetzung mit inhaltlichen Schwächen und fachlichen Fehlern begünstigte. Die CJ-Methode erwies sich somit zugleich als wirkungsvolles Instrument zur Förderung sachlicher Argumentationskompetenz und metakognitiver Reflexion.

Die Untersuchung der beiden Varianten A und B lieferte zusätzliche Einblicke in die Lernprozesse und bestätigte das Auftreten eines *order effect*: Wurden die Studierenden zuerst zur eigenständigen Analyse angeleitet (Variante A), bewerteten sie die KI-Ausgabe kritischer und reflektierter als in Variante B. Es lässt sich vermuten, dass die Studierenden in Variante B, in der die KI-Ausgabe zuerst präsentiert wurde, stärker von den KI-Ergebnissen beeinflusst und dadurch in ihrer eigenen Urteilsbildung voreingenommener waren. In Variante A hingegen verfügten sie durch den eigenständigen Analyseprozess über ein klares Bewusstsein der eigenen Schritte und Entscheidungsprozesse, was eine fundiertere und kritischere Bewertung der KI-Ausgabe begünstigte. Dies unterstreicht die Bedeutung einer sorgfältigen didaktischen Rahmung und bewussten Reihenfolge von Fachtextlektüre, eigener Analyse und anschließendem Vergleich mit der KI-Ausgabe, um eine reflektierte und eigenständige Urteilsbildung zu fördern. Insgesamt zeigte sich, dass die Kombination aus *Comparative Judgement* und gezielter Promptgestaltung sowohl den fachlichen Kompetenzerwerb stärkte als auch die Entwicklung von Schlüsselkompetenzen im Sinne einer *Critical AI Literacy* wirksam unterstützte.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

2. Bedeutung von Vorwissen und Promptgestaltung:

Der didaktische Mehrwert der Interaktion mit KI-Chatbots zeigte sich vor allem dann, wenn die Studierenden über eine vorbereitete Wissensgrundlage verfügten, die ihnen eine kritische Überprüfung der KI-Inhalte ermöglichte und sie befähigte, Halluzinationen der KI zuverlässig zu erkennen. Zugleich erwies sich die Gestaltung der Prompts als zentrale Steuerungsgröße: Die über das universitätseigene Tool BIKI geteilten, von der Lehrkraft vorbereiteten Prompts ermöglichten methodisch kontrollierte und transparente Lernbedingungen. Präzise und informationshaltige Prompts verbesserten die fachliche Qualität und Nachvollziehbarkeit der KI-Ausgaben, konnten jedoch erwartungsgemäß das Auftreten von Halluzinationen nicht vollständig verhindern. Gerade diese fehlerhaften, aber scheinbar kohärenten KI-Ausgaben erwiesen sich als besonders lernwirksam, da sie die Studierenden dazu anregten, vermeintlich sprachwissenschaftliche Analysen kritisch zu überprüfen, eigene Analysewege zu begründen und die Funktionsweise generativer KI differenziert zu reflektieren. In der Prompt-Werkstatt wurden diese Zusammenhänge systematisch vertieft und trugen maßgeblich zur Entwicklung von Schlüsselkompetenzen im Sinne einer *Critical AI Literacy* bei.

3. Technologische Voraussetzungen als didaktische Bedingung:

Die Erfahrungen mit BIKI verdeutlichten, dass die Wahl der KI-Version einen wesentlichen Einfluss auf die fachliche Angemessenheit, Kohärenz und Plausibilität der generierten Analysen hat. Niedrigere Modelle (wie ChatGPT mini) halluzinierten erwartungsgemäß stark und lieferten selbst bei Grundbegriffen unzuverlässige oder inkonsistente Antworten. Erst der Einsatz leistungsfähigerer Modelle (ab ChatGPT 4.1) reduzierte die Zahl fehlerhafter Ausgaben deutlich und schuf eine günstigere Grundlage für die kritische Auseinandersetzung im Seminar. Ebenso erwies sich die Möglichkeit, über BIKI vorbereitete Prompts und KI-Antworten mit den Studierenden zu teilen, als wichtige technologische Voraussetzung für die Vergleichbarkeit von Studierenden- und KI-Leistungen sowie für die Nachvollziehbarkeit der Analyseprozesse. Für die hochschuldidaktische Praxis bedeutet dies, dass technologische Infrastruktur keine bloße Begleiterscheinung darstellt, sondern wesentlich zum didaktischen Nutzen des KI-Einsatzes beiträgt.

Transfer:

Im Bereich des Transfers des Gelernten zeigte sich, dass einige Studierenden die im Seminar entwickelte Methode des kritischen Vergleichs bereits eigeninitiativ anwandten. Besonders deutlich wurde dies in einem Referat, in dem eine Studierende zunächst eine KI-generierte Definition präsentierte, diese kritisch mit einer fachwissenschaftlich fundierten Definition verglich und die Unterschiede gemeinsam mit der Gruppe diskutierte. Das Beispiel zeigt, dass die im Projekt angelegte kritische Urteilsbildung über die Seminarsituation hinauswirkte und einen nachhaltigen Einfluss auf das wissenschaftliche Arbeitsverständnis der Studierenden hatte.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

8. Perspektiven und Empfehlungen

Das Lehrprojekt soll in kommenden Lehrveranstaltungen weitergeführt und inhaltlich ausgeweitet werden, um das Potenzial der modifizierten CJ-Methode als didaktisch gerahmtes Instrument zur Wissensvertiefung und kritisch-reflektierten Auseinandersetzung weiter zu erproben. Beibehalten wird der gezielte Einsatz der modifizierten CJ-Methode in Kombination mit vorbereiteten und über BIKI geteilten Prompts. Da präzise gestaltete Prompts die Qualität der KI-Ausgaben zwar verbessern, die bekannte Problematik des Halluzinierens jedoch nicht vollständig beheben können, werden diese Fehlleistungen der KI auch künftig weiterhin gezielt didaktisch als Reflexionsimpuls genutzt, um Studierende zur kritischen Analyse, fachgerechten Korrektur und reflektierten Bewertung der inhaltlichen Aussagekraft algorithmisch erzeugter Analysen anzuleiten. Ebenso wird das hybride Setting mit flexiblen Gruppenarbeitsformaten beibehalten, um auch zukünftig eine nahezu gleichwertige Beteiligung der Studierenden zu ermöglichen. Für kommende Durchführungen ist außerdem geplant, die Prompt-Werkstatt noch stärker mit praxisnahen Transferaufgaben zu verbinden, um die Anwendungskompetenz der Studierenden im Umgang mit KI-Tools weiter zu vertiefen.

Anderen Lehrenden kann die Kombination aus gezielter Promptgestaltung, dem Teilen der Prompts über BIKI mit den Studierenden sowie einer kritisch-reflexiven Vergleichsanalyse im Rahmen der modifizierten CJ-Methode besonders empfohlen werden. Diese Elemente fördern nicht nur ein vertieftes Verständnis und die reflektierte Anwendung fachwissenschaftlichen Wissens im Analyse- und Bewertungsprozess, sondern auch Schlüsselkompetenzen wie *Critical Digital Literacy* und *Critical AI Literacy*, reflektiertes Urteilsvermögen und die Fähigkeit, digitale Werkzeuge verantwortungsvoll einzusetzen.

Der didaktische Einsatz von KI verlangt, das Halluzinieren nicht nur als technisches Problem, sondern als pädagogische Chance zu begreifen. Das Erkennen inhaltlicher Inkonsistenzen und das fachgerechte Richtigstellen fehlerhafter KI-Ausgaben mithilfe der im Lernprozess erworbenen Fachkompetenzen schärfen das Bewusstsein der Studierenden für die Grenzen KI-generierter Analysen, fördern metakognitive Reflexion und stärken ihre Fähigkeit, wissenschaftliche Qualität eigenständig zu beurteilen.

9. Rahmenbedingungen der vorgestellten Veranstaltung, Materialien und Literatur

9.1 Studiengang und -art, für den das Projekt geplant wurde

Das Seminar richtete sich in erster Linie an Studierende der BA- und MA-Studiengänge Französisch innerhalb der Lehreinheit Romanistik an der Fakultät für Linguistik und Literaturwissenschaft (LiLi). Darüber hinaus nahm zudem eine Masterstudierende der Anglistik teil.

9.2 Lehrveranstaltungstyp und -dauer

Das zwei Semesterwochenstunden umfassende Seminar für Bachelor- und Masterstudierende wurde im SoSe 2025 angeboten und als einmalige Lehrveranstaltung durchgeführt.

9.3 Gruppengröße

An der Lehrveranstaltung haben fünf Studierende teilgenommen.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte

9.4 Beteiligte bei der Durchführung des Projekts

An der Durchführung des Lehrprojektes war eine Lehrende beteiligt.

9.5 Materialien und Literatur

Materialien

BIKI (Bielefelder KI-Portal; integrierte Auswahlumgebung mit verschiedenen ChatGPT-Versionen sowie Modellen wie DeepSeek R1 Distill Llama 70B, Llama 3.1 und 3.3., Qwen 2.5 und 3, Codestral 22B)

DTEN D 7 (interaktives Videokonferenzsystem)

Padlet (digitale Pinnwand für Ergebnissicherung)

Zoom-Whiteboard (digitale Tafel für kollaboratives Arbeiten im hybriden Setting)

Mentimeter (digitales Tool für Live-Umfragen und interaktive Abstimmungen)

Literatur

Bruine de Bruin, Wändi & Gideon Keren. 2003. Order effects in sequentially judged options due to the direction of comparison. *Organizational Behavior and Human Decision Processes* 92(1–2), 91–101. [https://doi.org/10.1016/S0749-5978\(03\)00080-3](https://doi.org/10.1016/S0749-5978(03)00080-3)

Piaget, Jean. 1975. *L'équilibration des structures cognitives. Problème central du développement.* Paris: PUF.

Pollitt, Alastair. 2012. The method of Adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice* 19(3), 281–300.

Reich, Kersten. 2012. *Konstruktivistische Didaktik*, Weinheim und Basel: Beltz.

Schwellenbach, Sarah. 2025. Revising and extending Gricean maxims: The TRICS-Principles. *Journal of Linguistics*, 1–28. <https://doi.org/10.1017/S002226725100765>

Thurstone, Louis L. 1927. A law of comparative judgment. *Psychological Review* 34(4), 273–286.

von Glaserfeld, Ernst. 1997. *Radical Constructivism. A Way of Knowing and Learning*. London und Washington, D.C.: Falmer Press.



Die Inhalte dieses Dokuments sind – sofern nicht anders vermerkt – lizenziert unter einer [Creative Commons Namensnennung - Weitergabe unter gleichen Bedingungen 4.0 International Lizenz](#) (CC BY-SA 4.0).

Universität Bielefeld | Zentrum für Lehren und Lernen | ZLL | Zertifikat „Professionelle Lehrkompetenz für die Hochschule“ | www.uni-bielefeld.de/lehrkonzepte