

Kapitel 13

Literarisches Übersetzen im Zeitalter der künstlichen Intelligenz: Möglichkeiten und Grenzen teilautomatisierter Arbeitsprozesse

Katharina Walter

Universität Innsbruck

Nach einem kurzen Vergleich neuronaler maschineller Übersetzungssysteme wie DeepL und großer Sprachmodelle wie ChatGPT aus technologischer und translatorischer Sicht widmet sich dieser Beitrag konkreten Beispielen von Human- und KI-Übersetzungen literarischer Texte. Die analysierten Textbeispiele zeigen, dass KI-Übersetzungen abgesehen von ihrer Tendenz zur Fehlerhaftigkeit auch durch größere Ausgangstextnähe und Homogenität gekennzeichnet sind, während Humanübersetzungen deutlich mehr Individualität aufweisen. Da eine Teilautomatisierung literarischer Übersetzungsprozesse bereits im Gange ist, müssen sich angehende und praktizierende literarische Übersetzer:innen mit den Möglichkeiten und Grenzen neuer KI-Technologien auseinandersetzen. Dies ermöglicht nicht nur effizient mit neuen KI-Werkzeugen zu arbeiten, sondern gegebenenfalls auch überzeugend gegen ihren Einsatz zu argumentieren – wo auch immer dies für nötig erachtet wird. Es deutet sich an, dass automatisierte Rohübersetzungen ganzer literarischer Texte mit anschließendem Post-Editing eher zu Qualitätseinbußen führen als der punktuelle Einsatz von KI-Anwendungen zur Lösung konkreter Übersetzungsprobleme.

1 Einleitung

Obgleich man noch vor wenigen Jahren annahm, dass sich das literarische Übersetzen der Automatisierung weitgehend entziehen würde, haben neuronale maschinelle Übersetzungssysteme (NMÜ-Systeme) wie DeepL und Large



Katharina Walter. 2026. Literarisches Übersetzen im Zeitalter der künstlichen Intelligenz: Möglichkeiten und Grenzen teilautomatisierter Arbeitsprozesse. In Oliver Czulo, Martin Kappus & Felix Hoberg (Hrsg.), *Digitale Translatologie*, 217–237. Berlin: Language Science Press. DOI: 10.5281/zenodo.17523056

Language Models (LLMs) wie ChatGPT längst auch in diese Domäne Einzug gehalten. Abgesehen von einer stetig wachsenden Anzahl an Publikationen zur (teil-)automatisierten literarischen Übersetzung (siehe beispielsweise Adhikari 2018, Hadley 2020, Kolb u. a. 2023, Macken u. a. 2022, Taivalkoski-Shilov 2019) dokumentieren diese Veränderung auch die Webauftritte mancher Verlagshäuser. Rossum Press beispielsweise bewirbt Übersetzungen literarischer Texte, die aufgrund einer effizienten Verknüpfung von KI-Anwendungen und Post-Editing in Teamarbeit nur einen Bruchteil von konventionell erstellten literarischen Übersetzungen kosten sollen. Der Verlag stellt außerdem in Aussicht, aufgrund moderner, technologiegestützter Arbeitsprozesse mehr Übersetzungen auf den Markt bringen zu können als andere Anbieter, und zwar ohne Qualitätseinbußen und ohne die Originalautor:innen finanziell im selben Maße zu belasten, wie das im Falle althergebrachter literarischer Humanübersetzungen erforderlich wäre (Rossum Press 2024).¹

Wenn man den Versprechungen von Rossum Press Glauben schenken möchte, kann man davon ausgehen, dass die reine Humanübersetzung literarischer Texte bald Geschichte sein wird. Auch wenn KI-Anwendungen die literarische Übersetzung schon verändert haben und auch weiter verändern werden, darf man jedoch bezweifeln, dass es möglich ist, maschinell vorübersetzte literarische Texte in Teamarbeit und mit einem Bruchteil der vormals erforderlichen Ressourcen so zu verbessern, dass diese mit qualitativ hochwertigen Humanübersetzungen auch nur annähernd mithalten können. Bisherige wissenschaftliche Publikationen zur (teil-)automatisierten Übersetzung literarischer Texte deuten ebenso wie die hier besprochenen Textbeispiele von Virginia Woolf (2017 [1921]) und Séamus Ó Conaill (2018) darauf hin, dass Humanübersetzer:innen bei der Anpassung literarischer Texte für die zielsprachlichen und -kulturellen Gegebenheiten weiterhin federführend sein müssen.

Nach einer knappen Einführung in Gemeinsamkeiten und Unterschiede von NMÜ-Systemen wie DeepL im Vergleich zu LLMs wie ChatGPT aus translatatorischer Sicht beleuchtet dieser Beitrag anhand zweier konkreter Textbeispiele und unter Einbeziehung einschlägiger wissenschaftlicher Publikationen mögliche Vor- und Nachteile beider Arten von Anwendungen für das literarische Über-

¹Im Originalwortlaut heißt es: „Using a system of AI-assisted team translation, our skilled editors are able to create high quality literary translations with a fraction of the resources which traditional methods require. Every word of the AI-generated draft translation is carefully weighed by a professional stylist of the target language, and we work closely with our authors at every step along the way. This novel translation method means that we can bring readers more titles than conventional publishers of translated fiction, while taking a much smaller slice of the pie away from authors.“ (Rossum Press 2024)

setzen. Abgesehen von den potenziellen Aufgaben von Menschen in Gegenüberstellung zu Maschinen im literarischen Übersetzen der Zukunft steht dabei auch die Frage im Raum, welche Rolle NMÜ-Systeme künftig im Vergleich zu LLMs einnehmen könnten.

Allgemein bleibt beim literarischen Übersetzen bislang ungeklärt, ob der Nutzen, der sich durch eine mögliche Produktivitätssteigerung aufgrund von KI-generierten Rohübersetzungen ergibt, die Nachteile aufwiegt, welche zum Beispiel durch das so genannte „Priming“ (siehe beispielsweise Kolb 2022: 20) bzw. die Einschränkung der Vielfalt an Übersetzungsmöglichkeiten durch den Einfluss der maschinellen Ausgabe entstehen können. Unter anderem wird in diesem Beitrag untersucht, ob (teil-)automatisierte literarische Übersetzungen die Kreativität und damit auch die Qualität literarischer Zieltexte verringern, indem sie zu homogeneren Übersetzungen führen, als sie Menschen ohne maschinelle Unterstützung anfertigen würden, oder ob KI-Anwendungen im Gegenteil die Kreativität auch fördern können. Dies wäre denkbar, wenn es Humanübersetzer:innen aufgrund von KI-generierten Rohfassungen beispielsweise möglich wäre, sich primär auffordernde Textabschnitte zu konzentrieren. Auch könnte die Tendenz von KI-Systemen, lexikalische und syntaktische Phänomene im Ausgangstext in einer Übersetzung zu reproduzieren, Post-Editor:innen vielleicht sogar dazu ermutigten, wörtlichere Übersetzungen unkonventioneller Ausdrücke zu verwenden, als es Humanübersetzer:innen normalerweise tun. So haben Kolb u. a. (2023) in ihrer Untersuchung einiger von Menschen und Maschinen übersetzten englischsprachigen Versionen von Nestroy's Stück *Der Talisman* (1840) festgestellt, dass NMÜ-Systeme wie DeepL und Google Translate manchmal wörtliche und damit neuartige zielsprachliche Äquivalente für Okkasionalismen anbieten. Dabei handelt es sich um Neologismen, die Nestroy oft nur einmal verwendete und für die Humanübersetzer:innen meist normalisierende Äquivalente wählten. Andere Studien über literarische Prosa (Kolb 2023, Guerberof-Arenas & Toral 2022, 2024) kamen jedoch zum Schluss, dass nachbearbeitete Texte insgesamt weniger kreativ seien und ihre Leser:innen auch weniger ansprechen würden als Humanübersetzungen. In diesem Beitrag wird übersetzerische Kreativität mit Bayer-Hohenwarter als eine Manifestation so genannter „creative shifts“ bzw. „kreativer Verschiebungen“ definiert, die auftreten, wenn Übersetzer:innen ihre Ausgangstexte etwa lexikalisch, grammatisch oder syntaktisch umgestalten (Bayer-Hohenwarter 2011: 669). Auch das von Guerberof-Arenas & Toral (2022: 3) diskutierte Zusammenspiel zwischen den einzelnen Dimensionen der Akzeptanz oder Zweckmäßigkeit, der Flexibilität oder Neigung zu kreativen Verschiebungen, der Neuartigkeit oder Einzigartigkeit und der Geläufigkeit oder Neigung

zu reproduktiven Strategien in einer Übersetzung spielt bei der Analyse ausgewählter Textpassagen eine Rolle.

Ganz allgemein werfen einige KI-kritische Publikationen in und außerhalb von Translationskontexten die Frage auf, ob KI-generierte oder KI-übersetzte Texte möglicherweise einen wesentlichen Grund für menschliches kreatives Schaffen, wozu auch das Übersetzen literarischer Texte gehört, *ad absurdum* führen. Der diesem menschlichen kreativen Schaffen zugrundeliegende Prozess stehe nämlich seit jeher auch mit der Suche nach Antworten auf fundamentale Fragen der Menschheit in Verbindung – mit einer Aufgabe also, die nicht von Maschinen übernommen werden könne (Mjölsnes 2022: 65, Kissinger u. a. 2021: 52). Schon die sumerische Keilschrift und ägyptischen Hieroglyphen behandelten vor mehreren tausend Jahren neben praktischen Dingen des alltäglichen Lebens auch religiöse und mythische Themen, um unter anderem den Sinn und Zweck des menschlichen Daseins zu hinterfragen. Demnach ist einer der zentralen Gründe für schriftliche Kommunikation seit jeher, in Worte zu fassen, was schwer zu begreifen ist – wozu gerade auch literarische Texte oft wesentlich beitragen. Analog dazu erfüllt auch das Übersetzen vieler literarischer Texte nicht nur historisch gesehen insbesondere den Zweck, Bestehendes in Frage zu stellen und Neues zu lernen. Von diesem Standpunkt aus betrachtet erscheint das automatisierte Erstellen bzw. Übersetzen literarischer Texte geradezu grotesk. Darum sollen allgemeine Fragen zur Intentionalität kreativen Schaffens und zu gesellschaftlichen Auswirkungen einer (Teil-)Automatisierung literarischer Übersetzungsprozesse in diesem Beitrag nicht außer Acht gelassen werden, obgleich es schwierig erscheint, darauf eindeutige Antworten zu geben.

2 NMÜ-Systeme und LLMs: Gemeinsamkeiten und Unterschiede aus technologischer und translatorischer Sicht

Bislang ist man in der Sprachtechnologieforschung uneins, ob so genannte General-Purpose LLMs wie ChatGPT mittelfristig NMÜ-Systeme wie DeepL ersetzen könnten. Dafür sprechen etwa laut Lee nicht nur die größeren Kontextfenster und die dadurch tendenziell bessere Textkohäsion in der maschinellen Ausgabe, sondern auch die ausgeprägte Flexibilität von LLMs, etwa im Hinblick auf die Anpassung stilistischer Textmerkmale im Zuge einer Übersetzung. Ein dezipierter Nachteil von LLMs gegenüber NMÜ-Systemen sei jedoch, dass aufgrund ebendieser höheren gestalterischen Flexibilität im Bereich der Textausgabe die Wiederholbarkeit von Übersetzungsprozessen in identischer oder zumindest

sehr ähnlicher Weise nicht wie bei der Verwendung von NMÜ-Systemen gewährleistet werden könne (Lee 2024: 2354–2361). Auch Lommel kommt auf der Grundlage eines Vergleichs der Übersetzungsleistung von Google Translate mit jener von ChatGPT 4.0 zum Schluss, dass LLMs wahrscheinlich NMÜ-Systeme nicht ersetzen werden, sondern vielmehr einen Anreiz für Unternehmen darstellen, diese zu verbessern. Ein mögliches Beispiel für eine solche Verbesserung wäre laut Lommel die Erweiterung der Kontextfenster in der NMÜ hin zu größeren Einheiten, also Absätzen statt bisher Sätzen. Obwohl die Übersetzungsergebnisse unter Verwendung von ChatGPT in Bezug auf die Textkohäsion möglicherweise besser seien und man mithilfe aufgabenspezifischer Prompts auch einen größeren Gestaltungsspielraum habe, betont Lommel, dass NMÜ-Systeme schneller übersetzen würden, einfacher zu trainieren und auch kostengünstiger seien als LLMs. Außerdem ließen sich NMÜ-Systeme besser mit CAT-Tools und anderen sprachtechnologischen Anwendungen kombinieren, und die Sicherheit der in NMÜ-Systeme eingespeisten Daten sei leichter zu gewährleisten (Lommel 2023). Es ist bekannt, dass beispielsweise DeepL bei Verwendung einer kostenpflichtigen Version seines Übersetzungssystems auf die Speicherung von Texteingaben zu Trainingszwecken verzichtet, während die Transparenz gängiger LLMs im Hinblick auf Datenschutz- und Urheberrechtsfragen bisher zu wünschen übrig lässt.

Ein kurz- oder mittelfristiges Verschwinden der NMÜ zugunsten von LLMs als Übersetzungsmaschinen ist daher angesichts der aktuellen Entwicklungen eher nicht zu erwarten. Das hängt nicht nur damit zusammen, dass z. B. DeepL als einer der international führenden Anbieter im Bereich NMÜ seinen Kund:innen seit kurzem für einige Sprachkombinationen und in kostenpflichtigen Versionen Übersetzungsdienstleistungen mit eigenem LLM anbietet (DeepL 2024) und sich damit wohl für die nächste Zeit einen Platz unter den führenden KI-Unternehmen im Sprachdienstleistungssektor gesichert hat. Auch wenn wegen möglicher Umwälzungen, die LLMs in vielen Bereichen der Alltags- und Berufswelt bewirken könnten, jede Prognose wohl mit Vorsicht zu genießen ist, sieht es derzeit so aus, als würden NMÜ-Systeme weiterhin ein wichtiges Hilfsmittel für die literarische Übersetzung im Allgemeinen bleiben, während LLMs sich vor allem als Inspirationsquelle für die Übersetzung besonders experimenteller Textpassagen oder auch in gebundener Rede verfasster Texte eignen. Letzteres soll am Beispiel eines Reims aus einem zeitgenössischen irischen Kinderbuch verdeutlicht werden.

An dieser Stelle erscheint es angebracht, NMÜ-Systeme und LLMs hinsichtlich ihrer Übersetzungsleistung in groben Zügen zu vergleichen. NMÜ-Systeme wie DeepL oder Google Translate erstellen die bestmögliche Übersetzung jedes

Eingabesegments auf der Grundlage der habituellen Nähe von Wörtern in der Ausgangssprache im Vergleich zur habituellen Nähe von Wörtern in der Zielsprache. Das 2017 eingeführte Transformer-Modell ist derzeit die beliebteste Architektur in der NMÜ. Nach der Umwandlung der quellsprachlichen Eingabe in Vektoren im Encoder wird im Decoder durch eine erneute Umwandlung von Vektoren in Text eine zielsprachliche Ausgabe generiert (Rothwell u. a. 2023: 104). Im Gegensatz zur seriellen Textverarbeitung der Vorgängermodelle bewirkt der Aufmerksamkeitsmechanismus, dass der Transformer zuerst das Wichtigste im Text erfasst, weshalb auch längere Sätze inzwischen immer besser übersetzt werden können (Siegel 2023: 313–314). Die Leistung gängiger NMÜ-Systeme ist zumindest für ressourcenreiche Sprachen wie Englisch, Russisch oder Deutsch inzwischen beachtlich. Dennoch bereiten Übersetzungsfehler, terminologische Inkonsistenzen, die Voreingenommenheit der Trainingsdaten (Rothwell u. a. 2023: 105), ein weniger nuanciertes Sprachregister in der NMÜ-Ausgabe und eine Reduktion der lexikalischen Vielfalt im Vergleich zu den eingespeisten Ausgangstexten weiterhin Probleme (Toral u. a. 2019). Letzteres Phänomen kann vor allem in der literarischen Übersetzung, bei der sowohl Textinhalt als auch Textform eine wesentliche Rolle spielen, zu Qualitätseinbußen führen, da eine stilistische Vereinheitlichung von Übersetzungen unumgänglich ist, sofern keine umfangreiche Nachbearbeitung erfolgt.

LLMs wie ChatGPT wurden nicht zum Übersetzen entwickelt, aber es hat sich schnell gezeigt, dass mit ihrer Hilfe Übersetzungen von ähnlicher Qualität wie mit NMÜ-Systemen erstellt werden können. Eine Gegenüberstellung von NMÜ-Systemen und LLMs hinsichtlich ihrer Architektur und Leistungsfähigkeit durch Hendy u. a. (2023) gibt Einblicke in die Möglichkeiten und Grenzen beider Systemtypen, die auch für die Übersetzungspraxis aufschlussreich sein können. Beide Anwendungstypen basieren auf der von Vaswani u. a. (2017) vorgestellten Transformer-Architektur, wobei LLMs nicht die bei NMÜ-Systemen übliche Zweiteilung in Encoder versus Decoder aufweisen, sondern nur mit einem Decoder ausgestattet sind. In diesem Zusammenhang ist auch relevant, dass LLMs im Gegensatz zu NMÜ-Systemen nicht mit mehrsprachigen, parallelen Korpora, sondern mit einsprachigen Daten trainiert werden, was eine stärkere Gewichtung ressourcenreicher Sprachen, allen voran Englisch, unumgänglich macht. Dementsprechend sind bei der Verwendung von LLMs als Übersetzungssysteme deutlich mehr Trainingsdaten erforderlich, bis eine starke kontextabhängige mehrsprachige Leistung nachgewiesen werden kann (Hendy u. a. 2023).

Bisherige Publikationen bewerten die Leistung von LLMs als Übersetzungsanwendungen im Vergleich zu NMÜ-Systemen unterschiedlich. Gao u. a. (2024)

betonen, dass LLMs zwar insgesamt schlechtere Ergebnisse liefern als NMÜ-Systeme, dass aber aufgabenspezifische Prompts die Leistung von LLMs erheblich steigern können. Solche Prompts sollten die Übersetzungsaufgabe (z. B. Ausgangs- und Zielsprache) und die Domäne (z. B. Recht versus Technik) spezifizieren und auch Part-of-Speech-Tags oder grammatischen Informationen enthalten. In einer weiteren Studie zum Prompt-Engineering in Übersetzungskontexten kommt He zum Schluss, dass die Übersetzungsergebnisse bei der Verwendung von ChatGPT auch dann verbessert werden können, wenn dem LLM eine Expert:innenrolle zugewiesen werde, wie im folgenden Beispiel: „You are a professional popular science translator“ (He 2024: 4). Jiao u. a. weisen einerseits auf unterschiedliche Übersetzungsergebnisse je nach verwendeter ChatGPT-Version hin, andererseits soll ChatGPT bessere Übersetzungen als NMÜ-Systeme produzieren, wenn die übersetzten Texte die gesprochene Sprache widerspiegeln, aber bei biomedizinischen Abstracts immer noch von NMÜ übertraffen werden (Jiao u. a. 2023). Lee hebt zwar die höhere Randomisierung der LLM-Übersetzungsergebnisse als Nachteil hervor, kommt aber insgesamt zum Schluss, dass ChatGPT mindestens so gut übersetze wie die verbreiteten NMÜ-Systeme (Lee 2024: 2352). Folglich gibt er auch eine sehr positive Zukunftspрогнose für LLMs im Sprachdienstleistungssektor ab: „Considering that LLMs are only at their nascent stage, it is almost certain that MT will be able to operate with increasing autonomy, at least with respect to specialized technical communications“ (Lee 2024: 2352). Sánchez-Gijón und Palenzuela-Badiola zeigen darüber hinaus, dass LLMs wie ChatGPT nicht nur für das Übersetzen selbst, sondern auch für die Vor- und Nachbereitung von Übersetzungen nützlich sein können. Sie nennen unter anderem die erfolgreiche Verwendung von LLMs zur Text- und Datenanalyse, Glossarerstellung und zur Korrektur von Übersetzungen (Sánchez-Gijón 2023: 238–239).

3 Human- und KI-Übersetzungen literarischer Kurztexte im Vergleich

Im Folgenden wird ein Auszug aus Virginia Woolfs literarischer Skizze „Blue & Green“ im englischsprachigen Original (2017 [1921]) und in der deutschen Übersetzung von Christel Kröning (2021 [1921]) zwei KI-Übersetzungen gegenübergestellt, die mit DeepL (2023) und ChatGPT 3.5 (2023) erstellt wurden. Anschließend werden die KI-Übersetzungen vom November 2023 mit zwei weiteren vom September 2024 verglichen, wobei bei der DeepL-Zweitübersetzung erstmals das Sprachmodell der nächsten Generation verwendet wurde, das wenige Wochen

zuvor unter anderem für die Sprachkombination Deutsch-Englisch eingeführt worden war. Bei der ChatGPT-Zweitversion wurde versucht, das Ergebnis durch zusätzliches aufgabenspezifisches Prompt-Engineering zu optimieren. Abschließend wird ein kurzer, gereimter Textauszug aus einem irischen Kinderbuch von Séamus Ó Conaill mit dem Titel *Spuds and the Spider* (Ó Conaill 2018) analysiert, um exemplarisch die gebundene Rede als einen Anwendungsbereich zu thematisieren, für den bis auf weiteres (wenn überhaupt) vor allem LLMs nützliche Übersetzungsergebnisse generieren. Aufgrund des begrenzten Textkorpus, dessen Analyse der Darstellung unterschiedlicher Tendenzen in der Human- und KI-Übersetzung und insbesondere einer möglichen Differenzierung zwischen NMÜ-Systemen (DeepL) und LLMs (ChatGPT) in diesem Kontext dienen soll, wird auf den Einsatz von Metriken zur Übersetzungsbewertung bzw. zur Quantifizierung linguistischer Diversität wie beispielsweise der Type-Token-Relation verzichtet. Es sei jedoch darauf hingewiesen, dass Korpusanalysen von nachbearbeiteten literarischen Übersetzungen im Vergleich zu reinen Humanübersetzungen seit einigen Jahren wertvolle Beiträge zum Verständnis der veränderten kognitiven Übersetzungsprozesse im Zeitalter der KI liefern (siehe z. B. Macken u. a. 2022).

In einem ersten Schritt werden in Tabelle 1 das englischsprachige Original von Woolf sowie die Human- und die beiden KI-Übersetzungen vom November 2023 dargestellt. Der Prompt für die erste Übersetzung mit ChatGPT am 05.11.2023 lautete wie folgt: „Please translate the following piece of flash fiction by Virginia Woolf into German.“ Um die Übersetzungsleistung dieses großen, nicht aufgabenspezifischen Sprachmodells zu optimieren, wurden Informationen zur Textsorte, Autorin und Zielsprache angegeben.

Die Humanübersetzung unterscheidet sich erwartungsgemäß deutlich von den beiden KI-Versionen. Während Kröning sich insgesamt weiter vom Ausgangstext entfernt, um einen idiomatischen literarischen Text in deutscher Gegenwartssprache zu produzieren, bleiben beide KI-Applikationen im Hinblick auf Wortwahl, Syntax und Stil sehr nah am Ausgangstext, weisen also erheblich weniger „creative shifts“ und somit deutlich mehr reproduzierende Übersetzungsstrategien auf als die Humanübersetzung (siehe Bayer-Hohenwarter 2011, Guerberof-Arenas & Toral 2022). Besonders fällt dies bei den jeweiligen Übersetzungen für „snub-nosed“ und „black tarpaulin of his hide“ ins Auge. Um eine Assoziation des Monstergesichts mit Kindlichkeit durch die Verwendung der Standardübersetzung zu vermeiden, wählt Kröning das neutralere „flachnasig“ statt des Adjektivs „stupsnasig“, das in beiden KI-Übersetzungen enthalten ist. Die KI-Übersetzung „schwarze Plane seiner Haut“ entspricht einer Übertragung des Ausgangssegments auf der Wortebene, während Kröning mit der Wahl von

13 Literarisches Übersetzen im Zeitalter der künstlichen Intelligenz

Tabelle 1: Human- und KI-Übersetzungen von Woolfs „Blue“

Virginia Woolf	Christel Kröning	DeepL (2023)	ChatGPT 3.5 (2023)
<p>BLUE</p> <p>The snub-nosed monster rises to the surface and spouts through his blunt nostrils two columns of water, which, fiery-white in the centre, spray off into a fringe of blue beads. Strokes of blue line the black tarpaulin of his hide. Slushing the water through mouth and nostrils he sings, heavy with water, and the blue closes over him dowsing the polished pebbles of his eyes. Thrown upon the beach he lies, blunt, obtuse, shedding dry blue scales. Their metallic blue stains the rusty iron on the beach. Blue are the ribs of the wrecked rowing boat. A wave rolls beneath the blue bells. But the cathedral's different, cold, incense laden, faint blue with the veils of madonnas.</p> <p>(118 Wörter)</p>	<p>BLAU</p> <p>Als das flachnasige Ungeheuer den Wasserspiegel durchbricht, stößt es durch jede Nüster eine Fontäne aus, die sich, im Innern grellweiß, auffächert zu einem blauen Perlenvorhang. Blau Pinselstriche überziehen das Teerschwarz seiner Haut. Wasser durch Maul und Nüstern schlürfend, sinkt es, vollgesogen, und das Blau schließt sich über ihm, schließt sich über den glänzenden Kieselaugen. An den Strand gespült liegt es da, glanzlos, stumpf, verliert eine spröde Schuppe nach der anderen. Ihr Metallblau färbt den Eisenrost des Strandes. Blau sind die Rippen des Ruderbootwracks. Eine Welle rollt unter den Glockenblumen hindurch. Die Kathedrale jedoch ist anders, kalt, weihrauchschwer, blassblau von Madonnenschleiern.</p> <p>(101 Wörter)</p>	<p>BLAU</p> <p>Das stupsnasige Ungeheuer steigt an die Oberfläche und spuckt durch seine stumpfen Nasenlöcher zwei Wassersäulen aus, die in der Mitte feurig-weiß sind und in einem Saum aus blauen Perlen verspritzen. Blaue Striche säumen die schwarze Plane seiner Haut. Er spült das Wasser durch Mund und Nasenlöcher und singt, schwer von Wasser, und das Blau schließt sich über ihm und tränkt die polierten Kieselsteine seiner Augen. Auf den Strand geworfen liegt er, stumpf, stumpf, trockene blaue Schuppen abwerfend. Ihr metallisches Blau befleckt das rostige Eisen am Strand. Blau sind die Rippen des zerstörten Ruderbootes. Eine Welle rollt unter den blauen Glocken. Aber die Kathedrale ist anders, kalt, weih- rauchgeschwängert, blassblau mit den Schleieren der Madonnen.</p> <p>(114 Wörter)</p>	<p>BLAU</p> <p>Das stupsnasige Ungeheuer taucht an die Oberfläche und spuckt durch seine stumpfen Nasenlöcher zwei Wasserfontänen aus, die in der Mitte feurig weiß sind und sich in einen Saum aus blauen Perlen versprühen. Striche aus Blau ziehen sich über die schwarze Plane seiner Haut. Er schwappat das Wasser durch Mund und Nüstern, singt, schwer vom Wasser, und das Blau schließt sich über ihm und löscht die polierten Kiesel seiner Augen aus. An den Strand geworfen liegt er, stumpf, abgestumpft, trocknet blau schuppige Haut ab. Ihr metallisches Blau färbt das rostige Eisen am Strand. Blau sind die Rippen des zerstörten Ruderbootes. Eine Welle rollt unter den blauen Glocken. Aber die Kathedrale ist anders, kalt, vom Weihrauch erfüllt, schwach blau mit den Schleieren der Madonnen. (123 Wörter)</p>
<p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220.</p>	<p>Woolf, Virginia. 1921/2021. <i>Virginia Woolf – Meistererzählungen / Collected Stories: Zweisprachige Ausgabe (Dt./Engl.) Neu Überetzt von Christel Kröning</i>. Anaconda, 175–176.</p>	<p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Überetzt mit DeepL. 05.11.2023.</p>	<p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Überetzt mit ChatGPT 3.5. 05.11.2023.</p>

„Teerschwarz“ einerseits eine Normalisierung der im Englischen durchaus ungewöhnlichen Kollokation vornimmt und andererseits historisches Wissen über die Herstellung wasserabweisender Planen zu Beginn des 20. Jahrhunderts mithilfe von Teer in ihre Übersetzung einfließen lässt.

Die KI-Übersetzungen unterscheiden sich nur geringfügig voneinander, so dass die Beobachtung von Lommel, dass LLMs aufgrund größerer Kontextfenster bei kohäsiven Elementen wie Genusmarkierungen bessere Übersetzungen produzieren, in diesem Übersetzungsbeispiel nicht bestätigt wird (Lommel 2023). Weder DeepL noch ChatGPT geben in den deutschen Textversionen das für das Nomen „Monster“ adäquate Neutrum-Pronomen „es“ aus, was auf die Tatsache zurückzuführen ist, dass Woolf wie im Englischen bei personifizierten Tieren häufig eine Geschlechtszuweisung vornimmt. Das maskuline Pronomen wird in beiden KI-Übersetzungen aus dem Ausgangstext übernommen, was die mehrfach geäußerte These nicht bestätigt, dass LLMs aufgrund größerer Kontextfenster u. a. Geschlechtsmarkierungen kohärenter übersetzen sollten (siehe Lommel 2023). Die große Nähe zum Ausgangstext führt teilweise zu erheblichen stilistischen Mängeln und Fehlern im Deutschen, wie die Verdoppelung deutscher Adjektive in „stumpf, stumpf“ (DeepL 2023) bzw. „stumpf, abgestumpft“ (ChatGPT 3.5 2023) sowie die Verwendung einer falschen Präposition in den Kollokationen „blass blau mit den Schleiern der Madonnen“ (DeepL 2023) und „schwach blau mit den Schleiern der Madonnen“ (ChatGPT 3.5 2023) belegen. Im Bereich der Idiomatik treten sowohl in der DeepL- als auch in der ChatGPT-Übersetzung Fehler deutlich hervor. In der DeepL-Ausgabe stößt der monsterartige Fisch „zwei Wassersäulen aus, die in der Mitte feurig-weiß sind und in einem Saum aus blauen Perlen verspritzen“ (2023). Der Fehler liegt hier in der Verwendung von „verspritzen“ als intransitives Verb. In der deutschsprachigen ChatGPT-Version sind weder die im Wasser verschwindenden Augen des eintauchenden Fisches („das Blau schließt sich über ihm und löscht die polierten Kiesel seiner Augen aus“) noch das Abwerfen der Schuppen nach dem Strand („trocknet blau schuppige Haut ab“) erkennbar (ChatGPT 3.5 2023).

Interessant sind auch Krönings Eingriffe an zwei Stellen des Textes, die im Englischen höchstwahrscheinlich fehlerhaft sind und deren Nichtbeachtung im Deutschen daher zwangsläufig zu Übersetzungsfehlern führt. Im Gesamtzusammenhang erscheint es schlüssiger, dass das Monster im Text „singt“ (Woolf 2021 [1921]) und nicht „sinkt“ (DeepL 2023 und ChatGTP 3.5 2023), und es ist auch wahrscheinlich, dass Woolf mit „blue bells“ eigentlich „bluebells“ bzw. „Glockenblumen“ (Woolf 2021 [1921]) meinte und nicht das in beiden KI-Übersetzungen ausgegebene Äquivalent auf Wortebene, „blaue Glocken“ (DeepL 2023 und ChatGTP 3.5 2023).

Um den Textvergleich möglichst übersichtlich zu gestalten, werden in Tabelle 2 die wichtigsten Beobachtungen zu den einzelnen Übersetzungen dargestellt. Im Anschluss wird in Tabelle 3 die ursprüngliche DeepL-Version mit einer neueren Version verglichen, die mit dem Sprachmodell der nächsten Generation erstellt wurde, das DeepL im September 2024 für einige Sprachkombinationen, darunter Englisch-Deutsch, vorgestellt hat.

Tabelle 2: Human- und KI-Übersetzungen von Woolfs „Blue“ im Vergleich

Christel Kröning	DeepL	ChatGPT 3.5
Kürzer	Länger	Länger als Original (Spiegelung des humanübersetzerischen Phänomens der Explizierung?)
Übersetzerin entfernt sich oft vom Originalwortlaut (z. B. „flachnasig“, „Teerschwarz seiner Haut“ für „snub-nosed“ und „black tarpaulin of his hide“)	Bleibt näher am Originalwortlaut (z. B. „stupsnasig“, „schwarze Plane seiner Haut“ für „snub-nosed“ und „black tarpaulin of his hide“)	Bleibt näher am Originalwortlaut (z. B. „stupsnasig“, „schwarze Plane seiner Haut“ für „snub-nosed“ und „black tarpaulin of his hide“)
Keine Fehler	Fehler (z. B. Wassersäulen, die „in einem Saum aus blauen Perlen verspritzten“, maskulines Pronomen „[e]r“ zur Beschreibung des Monsters, Interferenz des Englischen in „blassblau mit den Schleiern der Madonnen“ für „faint blue with the veils of madonnas“)	Fehler (z. B. maskulines Pronomen „[e]r“ zur Bescheibung des Monsters, Übersetzen von „dowsing“ als „löscht ... aus“ und „shedding“ als „trocknet ... ab“, Interferenz des Englischen in „schwach blau mit den Schleiern der Madonnen“ für „faint blue with the veils of madonnas“)
Keine stilistischen Schwächen	Stilistische Schwächen (z. B. „stumpf, stumpf“ für „blunt, obtuse“)	Stilistische Schwächen (z. B. „stumpf, abgestumpft“ für „blunt, obtuse“)
Verbessert potenzielle Schwächen im Originaltext (z. B. „sinkt“ für „sings“, „Glockenblumen“ für „blue bells“)	Keine Verbesserung von potenziellen Schwächen im Originaltext (z. B. „singt“ für „sings“, „blaue Glocken“ für „blue bells“)	Keine Verbesserung von potenziellen Schwächen im Originaltext (z. B. „singt“ für „sings“, „blaue Glocken“ für „blue bells“)

Das Ergebnis dieses Übersetzungsversuchs mit dem neuen DeepL-Sprachmodell deutet nicht auf eine merkliche Verbesserung der Ausgabequalität hin. Auf lexikalischer Ebene weist die neuere DeepL-Version im Vergleich zur älteren einige Unterschiede auf, die jedoch nicht zu einer Qualitätssteigerung führen (z. B. „steigt“ (DeepL 2023) versus „erhebt sich“ (DeepL 2024) am Anfang des Textes). Die ungeschickte Verdoppelung von „stumpf, stumpf“ (DeepL 2023) für „blunt, obtuse“ (Woolf 2017 [1921]) wurde in der neueren DeepL-Version durch ein ansprechenderes Adjektivpaar mit Alliteration ersetzt („stumpf, starr“, DeepL 2024).

Tabelle 3: Das klassische Sprachmodell von DeepL und die nächste Generation

Virginia Woolf	DeepL Classic Language Model	DeepL Next-Generation Language Model
<p>BLUE</p> <p>The snub-nosed monster rises to the surface and spouts through his blunt nostrils two columns of water, which, fiery-white in the centre, spray off into a fringe of blue beads. Strokes of blue line the black tarpaulin of his hide. Slushing the water through mouth and nostrils he sings, heavy with water, and the blue closes over him dowsing the polished pebbles of his eyes. Thrown upon the beach he lies, blunt, obtuse, shedding dry blue scales. Their metallic blue stains the rusty iron on the beach. Blue are the ribs of the wrecked rowing boat. A wave rolls beneath the blue bells. But the cathedral's different, cold, incense laden, faint blue with the veils of madonnas.</p> <p>(118 Wörter)</p> <p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220.</p>	<p>BLAU</p> <p>Das stupsnasige Ungeheuer steigt an die Oberfläche und spuckt durch seine stumpfen Nasenlöcher zwei Wassersäulen aus, die in der Mitte feurig-weiß sind und in einem Saum aus blauen Perlen verspritzen. Blaue Striche säumen die schwarze Plane seiner Haut. Er spült das Wasser durch Mund und Nasenlöcher und singt, schwer von Wasser, und das Blau schließt sich über ihm und tränkt die polierten Kieselsteine seiner Augen. Auf den Strand geworfen liegt er, stumpf, stumpf, trockene blaue Schuppen abwerfend. Ihr metallisches Blau befleckt das rostige Eisen am Strand. Blau sind die Rippen des zerstörten Ruderbootes. Eine Welle rollt unter den blauen Glocken. Aber die Kathedrale ist anders, kalt, weihrauchgeschwängert, blassblau mit den Schleieren der Madonnen. (114 Wörter)</p> <p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Übersetzt mit DeepL. 05.11.2023.</p>	<p>BLAU</p> <p>Das stupsnasige Monster erhebt sich an die Oberfläche und speit durch seine stumpfen Nasenlöcher zwei Wassersäulen, die in der Mitte feuerweiß sind und in einem Kranz aus blauen Perlen versprühen. Blaue Striche säumen die schwarze Plane seines Verstecks. Er schlürft das Wasser durch Mund und Nasenlöcher und singt, schwer vom Wasser, und das Blau schließt sich über ihm und bedeckt die polierten Kiesel seiner Augen. Er liegt am Strand, stumpf, starr, und lässt trockene blaue Schuppen fallen. Ihr metallisches Blau färbt das rostige Eisen am Strand. Blau sind die Rippen des zerstörten Ruderbootes. Eine Welle rollt unter den blauen Glocken. Aber die Kathedrale ist anders, kalt, mit Weihrauch geschwängert, mattblau mit den Schleieren der Madonnen. (116 Wörter)</p> <p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Übersetzt mit DeepL Pro (Sprachmodell der nächsten Generation).</p> <p>28.09.2024.</p>

Andererseits bleiben einige Übersetzungsfehler aus der DeepL-Erstversion auch in der neueren Übersetzung in ähnlicher Form bestehen: Die Phrase „in einem Kranz aus blauen Perlen versprühen“ (DeepL 2024) enthält wieder ein transitives Verb in intransitiver Verwendung, und „mattblau mit den Schleieren der Madonnen“ (DeepL 2024) beruht wieder auf der direkten Übertragung der Präposition aus dem Englischen, die in der vorliegenden Kollokation im Deutschen jedoch falsch ist. Da das englische „hide“ (Woolf 2017 [1921]) in der Zweitversion fälsch-

licherweise als „Versteck“ übersetzt wurde (DeepL 2024), kommt hier sogar ein zusätzlicher Fehler vor. Erste Beobachtungen deuten also darauf hin, dass bei Verwendung des DeepL-Sprachmodells der nächsten Generation mit einer leicht erhöhten Fehlerquote in der maschinellen Ausgabe zu rechnen ist, während gleichzeitig auch eine höhere Anzahl an „creative shifts“ (vgl. Bayer-Hohenwarter 2011, Guerberof-Arenas & Toral 2022, 2024) auftreten dürfte. Bei Übersetzungen mit dem herkömmlichen Sprachmodell von DeepL ist hingegen tendenziell häufiger mit reproduzierenden Übersetzungslösungen zu rechnen, wobei dieses Sprachmodell auch zuverlässiger ist. Ähnliches scheint auch für den Einsatz von LLMs für Übersetzungen ohne aufgabenspezifisches Prompting zu gelten.

In Tabelle 4 wird auch die bereits besprochene Version von ChatGPT 3.5 mit einer neueren Übersetzung verglichen, die mit ChatGPT 4o erstellt wurde. Um eine möglichst hohe Übersetzungsleistung zu gewährleisten, enthielt der Prompt für die Erstübersetzung mit ChatGPT 3.5 am 05.11.23 als aufgabenspezifische Informationen den Namen der Originalautorin, die Bezeichnung des Genres und die Angabe der Zielsprache. Für die ChatGPT-Zweitübersetzung wurde dem Modell 4o, wie von He (2024: 4) und OpenAI (2024) empfohlen, zusätzlich eine Expert:innenrolle zugeschrieben. Durch die explizite Aufforderung an ChatGPT, die Übersetzungsaufgabe langsam auszuführen, wurde ein weiterer Hinweis von OpenAI für effektives Prompt-Engineering berücksichtigt (OpenAI 2024). So lautete der Prompt vom 27.09.24: „You are an author and literary translator specializing in early 20th century fiction from the United Kingdom. Please translate the following piece of flash fiction by Virginia Woolf from English into German and focus on capturing the stylistic nuances of the source text while using idiomatic German. Take your time.“ Das Ergebnis wird in Tabelle 4 neben dem Original von Woolf und der ersten ChatGPT-Übersetzung dargestellt.

Auch wenn die zweite deutschsprachige Version von ChatGPT kaum Übersetzungsfehler enthält und als Zieltext idiomatischer ist als die erste Version, besteht durchaus noch Überarbeitungsbedarf, wenn eine dem Original qualitativ ebenbürtige Übersetzung erstellt werden soll. Auf lexikalischer Ebene weist die zweite deutschsprachige ChatGPT-Version (ChatGPT 4o 2024) keine Übersetzungsfehler auf, während potentielle Fehler im Ausgangstext („sings“ statt „sinks“ bzw. „blue bells“ statt „bluebells“) auch in dieser Übersetzung bestehen bleiben. Auch das maskuline Pronomen „er“ als ungeeignetes Äquivalent für das im Neutrum verwendete Nomen „Monster“ bleibt in der neuen Übersetzung erhalten. Andererseits deutet die Übersetzung von „snub-nosed“ als „stumpfnasig“ in der deutschsprachigen Version vom 27.09.2024 möglicherweise darauf hin, dass ChatGPT 4o in der Lage sein könnte, bestimmte Bedeutungsnuancen kohärenter wiederzugeben, als dies in der ChatGPT-Erstversion der Fall ist.

Tabelle 4: ChatGPT-Übersetzungen von Woolfs „Blue“

Virginia Woolf	ChatGPT 3.5	ChatGPT 4o
<p>BLUE</p> <p>The snub-nosed monster rises to the surface and spouts through his blunt nostrils two columns of water, which, fiery-white in the centre, spray off into a fringe of blue beads. Strokes of blue line the black tarpaulin of his hide. Slushing the water through mouth and nostrils he sings, heavy with water, and the blue closes over him dowsing the polished pebbles of his eyes. Thrown upon the beach he lies, blunt, obtuse, shedding dry blue scales. Their metallic blue stains the rusty iron on the beach. Blue are the ribs of the wrecked rowing boat. A wave rolls beneath the blue bells. But the cathedral's different, cold, incense laden, faint blue with the veils of madonnas.</p> <p>(118 Wörter)</p>	<p>BLAU</p> <p>Das stupsnasige Ungeheuer taucht an die Oberfläche und spuckt durch seine stumpfen Nasenlöcher zwei Wasserfontänen aus, die in der Mitte feurig weiß sind und sich in einen Saum aus blauen Perlen versprühen. Striche aus Blau ziehen sich über die schwarze Plane seiner Haut. Er schwapppt das Wasser durch Mund und Nüstern, singt, schwer vom Wasser, und das Blau schließt sich über ihm und löst die polierten Kiesel seiner Augen aus. An den Strand geworfen liegt er, stumpf, abgestumpft, trocknet blau schuppige Haut ab. Ihr metallisches Blau färbt das rostige Eisen am Strand. Blau sind die Rippen des zerstörten Ruderbootes. Eine Welle rollt unter den blauen Glocken. Aber die Kathedrale ist anders, kalt, vom Weihrauch erfüllt, schwach blau mit den Schleieren der Madonnen. (123 Wörter)</p>	<p>BLAU</p> <p>Das stumpfnasige Ungeheuer erhebt sich an die Oberfläche und sprüht durch seine stumpfen Nüstern zwei Wasserfontänen, die im Zentrum feurig-weiß sind und sich zu einem Saum aus blauen Tropfen vernebeln. Blaue Striche durchziehen die schwarze Plane seines Fells. Schwere Wasserströme durch Mund und Nüstern spülend, singt er, schwer beladen mit Wasser, und das Blau schließt sich über ihm und taucht die polierten Kiesel seiner Augen in Dunkelheit. An den Strand geworfen, liegt er da, stumpf, dumpf, wirft trockene blaue Schuppen ab. Ihr metallisches Blau befleckt das rostige Eisen am Strand. Blau sind die Rippen des gestrandeten Ruderbooten. Eine Welle rollt unter den blauen Glocken. Doch die Kathedrale ist anders, kalt, von Weihrauch erfüllt, ein blasses Blau, verschleiert von Madonnen. (123 Wörter)</p> <p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Übersetzt mit ChatGPT 3.5. 05.11.2023.</p> <p>Prompt: „Please translate the following piece of flash fiction by Virginia Woolf into German.“</p> <p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220. Übersetzt mit ChatGPT 4o. 27.09.2024.</p> <p>Prompt: „You are an author and literary translator specializing in early 20th century fiction from the United Kingdom. Please translate the following piece of flash fiction by Virginia Woolf from English into German and focus on capturing the stylistic nuances of the source text while using idiomatic German. Take your time.“</p>
<p>Woolf, Virginia. 1921/2017. „Blue & Gree“. In <i>Monday or Tuesday</i>. Project Gutenberg. https://www.gutenberg.org/ebooks/29220.</p>		

Als letzte Herausforderung im Vergleich von Human- und KI-Übersetzungen soll hier ein Reim besprochen werden, der in einem 2018 veröffentlichten Kinderbuch vorkommt. *Spuds and the Spider* von Séamus Ó Conaill mit Illustrationen von Daniele Archimede handelt von einer koboldartigen Figur aus der irischen Folklore, einem *leprechaun* namens Spuds Potsofgold, und seiner Frau Rose Goodytwoshoes. Spuds gerät in einen Konflikt mit einem Spinnerich namens Leggers McWeb, der überall in seinem Haus Spinnweben hinterlässt. Am Ende rettet Leggers mit einem seiner Fäden Spuds, der in einen tiefen Brunnenschacht gefallen ist, und die beiden versöhnen sich.

Um die Streitigkeiten zwischen Spuds und Leggers zu schlichten und ihren Mann zur Ruhe zu bringen, spricht Rose immer wieder dieselben gereimten Worte, deren Human- und KI-Übersetzungen in der Tabelle 5 zusammengefasst sind. Die Humanübersetzungen wurden in Lehrveranstaltungen zum literarischen Übersetzen (Wintersemester 2021–22) und zum kreativen Schreiben (Sommersemester 2023) von Studierenden an der Universität Innsbruck erstellt, während die KI-Übersetzungen jeweils am 18.01.2024 mit DeepL, ChatGPT 3.5 und ChatGPT 4.0 generiert wurden. Die Tabelle zeigt, dass bei den vorliegenden Humanübersetzungen grundsätzlich die Tendenz besteht, die Textform, also die Reime und Rhythmisierung des Textes, in den Vordergrund zu stellen, während bei den KI-Übersetzungen zunächst ausschließlich die Bedeutungsebene berücksichtigt wurde. Während bei Verwendung von DeepL vorerst keine Möglichkeit besteht, eine Übersetzung hinsichtlich der Textform anzupassen, kann man mit ChatGPT seit jeher gereimte Zieltexte erstellen. Allerdings war es beim hier präsentierten Übersetzungsversuch nicht möglich, in einem Schritt den Text in gebundener Rede vom Englischen ins Deutsche zu übertragen, obwohl der Ausgangstext im Prompt explizit als „rhyme“ bezeichnet wurde.

Wenngleich beispielsweise das Wortpaar „Verdruss“ und „Diskurs“ als Reimvorschlag für einen Text der Kinderliteratur nur bedingt tauglich ist und vermutlich keiner der hier angeführten KI-Übersetzungsvorschläge unverändert übernommen werden könnte, eröffnet die Möglichkeit, auch Texte in gebundener Rede teilautomatisiert zu übersetzen, vielversprechende neue Perspektiven. Auch Kolb u. a. weisen auf mögliche positive Effekte von KI-Anwendungen auf die Übertragung von Wortneuschöpfungen sowie experimentellen Texten oder Textpassagen hin. Da die maschinelle Ausgabe in der Regel lexikalisch und syntaktisch näher beim Ausgangstext bleibt als Humanübersetzungen, können KI-Anwendungen ihre Nutzer:innen dazu einladen, originelle Übersetzungsideen aufzugreifen, und so der Normalisierung stilistischer Besonderheiten im Zieltext entgegenwirken (Kolb u. a. 2023).

Tabelle 5: Human- und KI-Übersetzungen eines Reims

Original	Humanübers. WS 2021-22 (Lit. Üs) und SS 2023 (Kreat. Schr.)	DeepL (18.01.24)	ChatGPT 3.5 (18.01.24)	ChatGPT 4 (18.01.24)
<p><i>He means no harm, Let him go, Just be kind, You never know. (Rose Goodytwoshoes versucht zwischen Spuds und Leggers zu vermitteln)</i></p>	<p><i>Er schadet nicht, Lass ihn gehn, Sei einfach nett, Man wird sehn. Alternative 1: <i>Er</i> meint's doch nicht böse, lass ihn einfach gehen. <i>Nett</i> zu sein lohnt sich, <i>du</i> wirst es schon sehen. Alternative 2: <i>Er</i> meint's nur gut, lass ihn geh'n <i>d'rum</i> hab nur Mut, <i>du</i> wirst schon sehn.</i></p>	<p><i>Er</i> meint es nicht böse, <i>Lass</i> ihn gehen, <i>Sei</i> einfach nett, <i>Man</i> kann nie wissen.</p>	<p>Prompt: „Please translate this rhyme into German.“</p>	<p>Prompt: „Please translate this rhyme into German.“</p>

4 Risikomanagement: Eine relevante Größe in der literarischen Übersetzung?

Seitdem zunächst NMÜ-Systeme und in der Folge auch LLMs Einzug in den Sprachdienstleistungssektor gehalten haben, sind die damit verbundenen persönlichen und gesellschaftlichen Risiken sowie die sich verändernden Kompetenzprofile von Translator:innen Gegenstand wissenschaftlicher Diskurse (siehe beispielsweise Nitzke u. a. 2019, Sandrini 2022, Hackenbuchner & Krüger 2023). Nitzke u. a. weisen unter anderem auf Risiken im Zusammenhang mit Datenschutzverletzungen und Qualitätseinbußen hin. Auch ein möglicher Verlust der Kontrolle über die verwendeten Hilfsmittel aus Sicht von Kund:innen sowie deren mögliche Haftung im Falle von Übersetzungsfehlern werden angeführt (Nitzke u. a. 2019: 241). Dementsprechend werden zahlreiche neue Kompetenzen definiert, die in der Sprachdienstleistungsbranche benötigt werden und sozial-

unternehmerische, sprachlich-übersetzerische und technologische Komponenten umfassen (Nitzke u. a. 2019: 248-250). Auch Hackenbuchner und Krüger betonen die steigende Relevanz einer MÜ-Kompetenz sowohl für Laien als auch für Sprachdienstleister:innen, wobei vor allem letztere in der Lage sein müssten, die KI-Übersetzung technologisch, sprachlich und wirtschaftlich-gesellschaftlich kompetent zu verorten (Hackenbuchner & Krüger 2023: 285-287). Im Hinblick auf technologische Kompetenzen verweisen Hackenbuchner und Krüger neben allgemeinen Aspekten der Sprachtechnologie insbesondere auch auf Datenkompetenz und eine kognitive MÜ-Kompetenz, die sich auf den Einfluss der MÜ auf Übersetzer:innen bezieht. Darüber hinaus betonen Hackenbuchner und Krüger die Bedeutung eines Verständnisses für die wirtschaftlichen und gesellschaftlichen Implikationen der MÜ, wobei zu letzterem Bereich auch ethische Fragen gehören (Hackenbuchner & Krüger 2023: 285-292).

Dass eine Verwendung von KI-Übersetzungen ohne menschliche Kontrolle etwa im staatlichen Gesundheitssystem des Vereinigten Königreichs (NHS) mit erheblichen Risiken verbunden ist, dürfte außer Zweifel stehen. Dort wurde der Verzicht auf Humandolmetscher:innen unter anderem mit einer möglichen gesundheitlichen Schädigung von Patient:innen in Verbindung gebracht (Hill 2023). Aber auch im Bereich der literarischen Übersetzung birgt eine fehlende oder unzureichende Nachbearbeitung automatisierter Übersetzungen Risiken. Schlechte literarische Übersetzungen können laut Taivalkoski-Shilov dem Ansehen der übersetzten Originalautor:innen, deren Werken und dem Lesepublikum schaden, das in allen Sprachen Anspruch auf höchste literarische Qualität haben sollte (Taivalkoski-Shilov 2019: 693).

Nichtsdestotrotz ist die KI-Übersetzung aus dem literarischen Übersetzen nicht mehr wegzudenken, und es gibt Bestrebungen in der Übersetzungspraxis und -wissenschaft, Wege für eine möglichst positive und nachhaltige Vernetzung von Mensch und Maschine zu finden. Macken u. a. (2022) äußern sich grundsätzlich positiv über die Möglichkeit, durch effiziente KI-gestützte Workflows qualitativ hochwertige und bezahlbare literarische Übersetzungen zu erstellen. Als Beispiel führen Macken u. a. den Verlag Nuanxed an, bei dem literarische Zieltexte unter Einbeziehung von KI-Übersetzung, anschließendem Post-Editing (zweisprachig) und abschließendem Lektorat (einsprachig) erstellt werden. Die Untersuchung eines ersten Textkorpus ergab, dass am Ende der dritten Phase kein nennenswerter Qualitätsunterschied zwischen einer literarischen Humanübersetzung und einer post-editierten KI-Übersetzung nachweisbar war (Macken u. a. 2022).

5 Resümee und Ausblick

Neben translationswissenschaftlichen und -praktischen Perspektiven, die den Einzug neuer Technologien in das literarische Übersetzen zu legitimieren versuchen, werden auch Stimmen laut, die sich dezidiert gegen eine (Teil-)Automatisierung des literarischen Übersetzens aussprechen. Ein offener Brief der Berufsverbände der Übersetzer:innen aus Österreich, Deutschland und der Schweiz schließt mit einem Manifest für menschliche Sprache. In Punkten fünf und sechs wird beispielsweise darauf hingewiesen, dass KI-Übersetzungen ihre Leserschaft im Hinblick auf ihre Authentizität täuschen und gleichzeitig das Ökosystem des Literaturbetriebs schädigen würden (VdÜ / A*ds / IGÜ 2024). Auch Mjölsnes argumentiert, dass das Produkt der maschinellen Übersetzung letztlich „stumm“ sei, da es nur eine Textoberfläche ohne Verankerung im Hinterland gelebter Erfahrung schaffe. Bedeutungsvoll könnte eine KI-Übersetzung nur in einem nachfolgenden Schritt, im Zuge eines Post-Editing, werden, wodurch ein ursächlicher Grund für das Schreiben selbst in Frage gestellt werde. Mjölsnes schreibt: „In dieser Trennung liegt ihr Grundproblem, denn sie widerspricht den Vorgängen, dank denen die Sprachen Sinn schöpfen und Erkenntnis hervorbringen. Deswegen ist die maschinelle Übersetzung letzten Endes stumm, egal, wie eloquent ihr Wortgeräusch daherkommt.“ (Mjölsnes 2022: 65)

Obwohl die oben genannten kritischen Einwände überaus wichtige ethische und gesellschaftliche Fragen berühren, die ebenso wie Fragen des Datenschutzes und des Urheberrechts mit großer Aufmerksamkeit und Vorsicht zu behandeln sind, ist der technologische Fortschritt weder aufzuhalten, noch erscheint es aus translatorischer Sicht wünschenswert, auf vielseitige und leistungsstarke Werkzeuge wie NMÜ-Systeme und LLMs zu verzichten. Die hier diskutierten Übersetzungsbispielze zeigen, dass man mit NMÜ-Systemen und LLMs möglicherweise brauchbare Rohübersetzungen literarischer Texte erstellen und vor allem spezifische Übersetzungsprobleme lösen kann, zu denen beispielsweise die Übertragung von Textpassagen in gebundener Rede oder Wortneuschöpfungen gehört. Unabhängig von der Qualität der maschinellen Ausgabe obliegt es jedoch qualifizierten Humanübersetzer:innen beziehungsweise Post-Editor:innen, maschinelle Textfassungen zu publizierbaren literarischen Übersetzungen zu machen.

Trotz vieler Vorteile ist auch im Falle eines umfassenden literarischen Post-Editing mit einer Homogenisierung literarischer Übersetzungen und dadurch auch mit Einbußen im Bereich der übersetzerischen Kreativität zu rechnen (siehe Kolb 2022, 2023, Guerberof-Arenas & Toral 2024). Wie vor allem Kolb in Publikationen zur Post-Editing-Prozessforschung für literarische Übersetzungen zeigt (Kolb 2022, 2023), ist eine umfassende Nachbearbeitung und Personalisierung

maschinellem Übersetzungsvorschläge außerdem mit erheblichem Aufwand verbunden. Die Zeitsparnis durch Teilautomatisierung, die häufig als entscheidender Vorteil des Einsatzes von KI-Systemen auch in der literarischen Übersetzung hervorgehoben wird, dürfte somit deutlich geringer ausfallen als vielfach angenommen. Darum sollte es – sofern keine urheberrechtlichen Vorbehalte bestehen – Humanübersetzer:innen überlassen bleiben, ob und für welche Zwecke sie KI-Anwendungen einsetzen. Auch sollte das literarische Post-Editing als anspruchsvolle und ressourcenintensive Tätigkeit angemessen vergütet werden.

Literatur

- Adhikari, Megharaj. 2018. Literary translation in the age of machine. *Nepalese Translation* 2. 27–36.
- Bayer-Hohenwarter, Gerit. 2011. 'creative shifts' as a means of measuring and promoting translational creativity. *Meta: Jounal des traducteurs* 3(56). 662–692.
- DeepL. 2024. *Next-generation Language Model for DeepL Translator*. <https://support.deepl.com/hc/en-us/articles/14241705319580> - Next - generation - language-model-for-DeepL-Translator (25 September, 2024).
- Gao, Yuan, Ruili Wang & Feng Hou. 2024. How to design translation prompts for ChatGPT: An empirical study. In *Proceedings of the 6th ACM international conference on multimedia in asia workshops*, 1–7. DOI: 10.1145/3700410.3702123.
- Guerberof-Arenas, Ana & Antonio Toral. 2022. Creativity in translation: Machine translation as a constraint for literary texts. *Translation Spaces* 11(2). 184–212.
- Guerberof-Arenas, Ana & Antonio Toral. 2024. To be or not to be: A translation reception study of a literary text translated into Dutch and Catalan using machine translation. *Target* 36(2). 215–244.
- Hackenbuchner, Janica & Ralph Krüger. 2023. DataLitMT – Teaching data literacy in the context of machine translation literacy. In Mary Nurminen, Judith Brenner, Maarit Koponen, Sirkku Latomaa, Mikhail Mikhailov, Frederike Schierl, Tharindu Ranasinghe, Eva Vanmassenhove, Sergi Alvarez Vidal, Nora Aranberri, Mara Nunziatini, Carla Parra Escartín, Mikel Forcadá, Maja Popović, Carolina Scarton & Helena Moniz (Hrsg.), *Proceedings of the 24th Annual Conference of the European Association for Machine Translation*, 285–293. Tampere, Finland: European Association for Machine Translation. <https://aclanthology.org/2023.eamt-1.28> (27 Mai, 2024).
- Hadley, James. 2020. Literary machine translation: Are the computers coming for our jobs? *Counterpoint* (4). 14–18.

- He, Sui. 2024. Prompting ChatGPT for translation: A comparative analysis of translation brief and persona prompts. In *Proceedings of the 25th Annual Conference of the European Association for Machine Translation (volume 1)*, 316–326. Sheffield, UK: European Association for Machine Translation. DOI: 10.48550/arxiv.2403.00127.
- Hendy, Amr, Mohamed Abdelrehim, Amr Sharaf, Vikas Raunak, Mohamed Gabr, Hitokazu Matsushita, Young Jin Kim, Mohamed Afify & Hany Hassan Awadalla. 2023. *How good are GPT models at machine translation? A comprehensive evaluation*. DOI: 10.48550/arXiv.2302.09210.
- Hill, Matthew. 2023. NHS interpreting service problems contributed to patient deaths. *BBC News*. <https://www.bbc.com/news/uk-england-bristol-66605536> (26 September, 2025).
- Jiao, Wenxiang, Wenxuan Wang, Jen-tse Huang, Xing Wang, Shuming Shi & Zhaopeng Tu. 2023. *Is ChatGPT a good translator? Yes with GPT-4 as the engine*. DOI: 10.48550/arXiv.2301.08745.
- Kissinger, Henry, Eric Schmidt & Daniel Huttenlocher. 2021. *The Age of AI*. London: John Murray.
- Kolb, Waltraud. 2022. Welche Rolle können Maschinen in der Literaturübersetzung spielen? *UNIVERSITAS* 1/22. 19–23.
- Kolb, Waltraud. 2023. ‘I am a bit surprised’: Literary translation and post-editing processes compared. In Andrew Rothwell, Andy Way & Roy Youdale (Hrsg.), *Computer-assisted literary translation: The state of the art*, 53–68. London: Routledge.
- Kolb, Waltraud, Wolfgang U. Dressler & Elisa Mattiello. 2023. Human and machine translation of occasionalisms in literary texts: Johann Nestroy’s ‘Der Talmann’ and its English translations. *Target: International Journal of Translation Studies*. DOI: 10.1075/target.21147.kol.
- Lee, Tong King. 2024. Artificial intelligence and posthumanist translation: ChatGPT versus the translator. *Applied Linguistics Review* 15(6). 2351–2372. DOI: <https://doi.org/10.1515/applrev-2023-0122>.
- Lommel, Arle. 2023. *Why generative AI is the best thing to happen to machine translation*.
- Macken, Lieve, Bram Vanroy, Luca Desmet & Arda Tezcan. 2022. Literary translation as a three-stage process: machine translation, post-editing and revision. In *Proceedings of the 23rd Annual Conference of the European Association for Machine Translation*, 101–110. Ghent: European Association for Machine Translation.
- Mjølsnes, Ettore. 2022. *Der stumme Text: Eine Kritik der maschinellen Übersetzung*. Küsnacht: Digiboo.

- Nitzke, Jean, Silvia Hansen-Schirra & Carmen Canfora. 2019. Risk management and post-editing competence. *The Journal of Specialised Translation* 31(1). 239–259.
- Ó Conaill, Séamus. 2018. *Spuds and the Spider*. Dublin: Gill Books.
- OpenAI. 2024. *Prompt engineering*.
- Rossum Press. 2024. *Rossum Press*. <https://rossumpress.org/> (26 September, 2025).
- Rothwell, Andrew, Joss Moorkens, María Fernández-Parra, Joanna Drugan & Frank Austermuehl. 2023. Current machine translation technologies. In *Translation Tools and Technologies*, 97–113. London: Routledge.
- Sánchez-Gijón, Pilar. 2023. Analysis and evaluation of ChatGPT-induced HCI shifts in the digitalised translation process. In *Proceedings of the International Conference on Human-Informed Translation and Interpreting Technology*, 227–267. Shoumen, Bulgaria: INCOMA Ltd. DOI: 10.26615/issn.2683-0078.2023_021.
- Sandrini, Peter. 2022. Die neuronale didaktische Herausforderung: Berufsbild und Kompetenzen im Wandel. In Gernot Hebenstreit & Philipp Hofeneder (Hrsg.), *Translation im Wandel: Gesellschaftliche, konzeptuelle und didaktische Perspektiven*, 115–134. Frank & Timme. <https://doi.org/10.57088/978-3-7329-9129-7>.
- Siegel, Melanie. 2023. B 14 Maschinelle Übersetzung. In Rainer Kuhlen, Dirk Lwandowski, Wolfgang Semar & Christa Womser-Hacker (Hrsg.), *Grundlagen der Informationswissenschaft*, 307–316. Berlin; Boston: De Gruyter Saur.
- Taivalkoski-Shilov, Kristiina. 2019. Ethical issues regarding machine(-assisted) translation of literary texts. *Perspectives* 27(5). 689–703. DOI: 10.1080/0907676X.2018.1520907.
- Toral, Antonio, Mikel Forcada & Andy Way. 2019. Post-editese: An exacerbated translationese. In *Proceedings of MT Summit XVII*, 273–281. Dublin. <http://arxiv.org/abs/1907.00900>.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser & Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*, Bd. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fb053c1c4a845aa-Paper.pdf.
- VdÜ / A*ds / IGÜ. 2024. Offener Brief zur KI-Verordnung. Online.
- Woolf, Virginia. 2021 [1921]. *Meistererzählungen. Collected Stories. Zweisprachige Ausgabe*. Übers. von Christel Kröning. original-date: 1921. Köln: Anaconda.
- Woolf, Virginia. 2017 [1921]. *Monday or Tuesday by Virginia Woolf*. original-date: 1921. Project Gutenberg.

