

Von DeepL zu ChatGPT

Erfahrungen mit KI-Tools aus Sicht einer Fremdsprachenlehrerin

von Filomena Montemarano, Gymnasiallehrerin für Französisch und Italienisch, Dozentin für Fachdidaktik an der PH FHNW und der PH St. Gallen, Lehrmittelautorin («Tocca a te!») und LVB-Mitglied



«The development of AI [Artificial Intelligence] is as fundamental as the creation of the microprocessor, the personal computer, the Internet, and the mobile phone. It will change the way people work, learn, travel, get health care, and communicate with each other.»

Bill Gates, Blog GatesNotes, 21. März 2023

Von ersten Übersetzungstools zum Deep Learning

Als die ersten Übersetzungstools und Erfahrungen mit ChatGPT unsere Schulen erreichten, waren die Fremdsprachenlehrpersonen nicht sonderlich überrascht, da bereits vor der Pandemie, aber vor allem während des Distanzunterrichts, verschiedene maschinelle Übersetzungstools wie Google Translate oder DeepL im Schüleralltag Einzug gehalten hatten. Wie beim Chatbot GPT (Generative Pre-trained Transformer) basierten auch die Übersetzungsmaschinen auf immens grossen, neuronal angelegten Datenbanken, welche Deep Learning ermöglichen und somit eine ganz neue Textqualität erreichen können. Anstatt die Sätze Wort für Wort zu übersetzen, nimmt z.B. DeepL muttersprachliche Textbausteine als Modell und setzt sie zu einem flüssigen und grammatikalisch korrekten Text in der gewünschten Fremdsprache zusammen. Das Tool bedient sich also bei sprachlich perfekt formulierten Texten und generiert daraus eine Übersetzung in demselben perfekten Stil.

Die Auswirkungen dieser maschinellen Übersetzungstools auf den Unterricht liessen nicht lange auf sich warten: Plötzlich wurden Zusammenfassungen, Aufsätze und jegliche Art schriftlicher Hausaufgaben auf einem solch perfekten Sprachniveau eingereicht, das die Lernenden ohne Sprachaufenthalt nicht einmal ansatzweise hätten erreichen können. Da die fremdsprachigen Texte direkt online generiert wurden, gab es keine Quelle zum Nachverfolgen, geschweige denn einen möglichen Plagiatsnachweis. Der Frust unter den Fremdsprachenlehrpersonen war gross und es häuften sich Bemerkungen darüber, ob der Aufwand, in der Schule eine Fremdsprache über so viele Jahre mühsam zu erlernen, noch sinnvoll sei, da die neuen Apps qualitativ höhere Resultate ohne jeglichen Aufwand generieren konnten.

Einfluss auf den Fremdsprachenunterricht

Die Panikmacherei war jedoch unnötig, denn der Fremdsprachenunterricht bleibt wichtig wie eh und je.

Er hat sich aber gewandelt hin zur Handlungs- und Kompetenzorientierung. Es sind auf der gymnasialen Stufe neue (kreativere und komplexere) Text- und Prüfungsformen entstanden, welche mehr als deklaratives Wissen und Grammatikformen in den Schülerprodukten evaluieren. Denn auch mit den neuen digitalen Tools müssen viele Faktoren miteinbezogen werden: der Adressat, der Inhalt, das Genre, die Zweideutigkeiten, die verlässlichen Quellen und der Wahrheitsgehalt, kurz: Kontext und Media Literacy. So wie bei anderen Hilfsmitteln auch (Taschenrechner, Internet), muss man ziemlich genau wissen, was man sucht, um die richtige Information zu finden, um das erworbene Wissen zu integrieren und es danach auf seine Richtigkeit hin zu prüfen.

Unterdessen haben wir bereits gelernt, DeepL für unsere Zwecke zu nutzen und auch die Schüler:innen haben gemerkt, dass die eleganten Formulierungen und die schöne Sprache den Inhalt nicht ersetzen können, dass der unangebrachte Sprach-

stil an den Kontext angepasst werden muss und dass es nicht sinnvoll ist, die Lernphasen zu umgehen und unpräzise Texte in die Maschine einzutippen. Auch verliert man viel zu viel Zeit, wenn mittelmässige Resultate von ChatGPT an eine spezifische und kontextgebundene Aufgabe aus dem eigenen Unterricht angepasst werden müssen. Die Lernenden merken, dass sie für wirklich gute Resultate viel Energie ins Pre-Editing und Post-Editing investieren müssen, und vor allem, dass sie Lerngelegenheiten verpassen, wenn sie sich mit der Materie zu wenig auseinandersetzen und einfach die vorgeschlagene Antwort der KI unreflektiert übernehmen.

DeepL & Co. haben im Fremdsprachenunterricht weitreichende Spuren hinterlassen. Die Lehrpersonen können nicht mehr wie sonst ihren Unterricht auf der Fehleranalyse aufbauen, da schriftliche Arbeiten nicht immer von den Lernenden selbst stammen. Wozu soll die Lehrperson ein Feedback zu einem KI-generierten, formal perfekten Text abgeben? Was ist der Lerneffekt für die Schüler:innen? Deshalb sind komplexere Aufgabenformen entstanden, welche, wie weiter oben erwähnt, die Handlungskompetenz messen und den Kontext sowie die (Media) Literacy im Fokus haben: schriftliche und mündliche Sprachmittlungen (nicht reine Übersetzungen); anspruchsvollere digitale Produkte wie Podcast, Erklärvideo, Wiki oder Concept Maps; mündliche Präsentationen, gefolgt von Kolloquien. Die Vielfalt der Aufgaben hat klar zugenommen und auch die Motivation der Lernenden, welche die Fremdsprache in realitätsnahen Kontexten nutzen lernen.

Möglichkeiten von ChatGPT

Dank der Gratisversion von ChatGPT 3 (unterdessen ChatGPT 3.5) sind Künstliche Intelligenz und Deep Learning nun auch in den anderen Fächern der Geistes- und Sozialwissenschaften angekommen. Mit einem simplen «Prompt» (Anfrage) lassen sich innerhalb von Sekunden blumige Texte in perfekter Sprache generieren, welche nach Belieben erweitert oder verein-

facht werden können. Dabei gibt uns der Chatbot das Gefühl, intelligent interagieren zu können, da sich die Maschine wie bei einem echten menschlichen Gespräch den gesamten Ge-

sprächsverlauf merken kann und sich bei jeder weiteren Anfrage auf Vorheriges bezieht. Zudem erscheinen die «Token» (Satzbausteine) auf dem Bildschirm wie von Geisterhänden ge-

Strategien im Umgang mit Chatbots und KI im Unterricht

1. Chatbots im Unterricht sinnvoll integrieren und kritisch thematisieren. Mit den Lernenden das Prompten üben (Pre-Editing und auch Post-Editing) und die Grenzen der Chatbots ausloten. Die generierten Texte nach Qualitätsmerkmalen analysieren (Quellen, Kontext, Präzision, Relevanz usw.).
2. Mit den Schüler:innen klare Regeln aufstellen. Die Lernenden wissen, wie und weshalb sie KI-generierte Textblöcke in ihren Produkten als Quelle markieren müssen (auch solche von Übersetzungstools) und üben eine kritische Distanz zu fixfertigen, eleganten Formulierungen.
3. Neue Prüfungsformate entwickeln und Bewertungsraster überdenken: weg von reinen Reproduktionsaufgaben, hin zu einer neuen Aufgabenkultur. Hier einige erste Vorschläge:
 - Vorträge in der Fremdsprache immer durch ein Kolloquium ergänzen, wo das Verständnis und die Vertiefung mit dem Thema geprüft werden kann, ebenso der Lernfortschritt und Umgang mit neuem (thematischem) Vokabular.
 - Nicht nur reine Übersetzungen verlangen, sondern Sprachmittlungen, bei denen der Kontext (Adressat, kultureller Hintergrund) miteinbezogen und deshalb in Inhalt und Sprachstil erkennbar werden muss.
 - Bei Matur- und anderen individuellen Vertiefungsarbeiten mindestens dreiteiliges Bewertungsraster vorsehen (Produkt, Prozess, Kolloquium). Durch eine regelmässige Begleitung bei der Fragestellung, bei der Ideenfindung und beim Schreibprozess soll der Missbrauch minimiert werden.
 - Komplexere Lernaufgaben verlangen, auch solche, die mit mehreren digitalen Tools verknüpft werden müssen (Lernjournal, Concept Map, Präsentation).
 - Vermehrt mündliche Prüfungen abhalten. Es ist jedoch illusorisch, nur noch mündliche Leistungen zu evaluieren, da das Schreiben eine wichtige basale Kompetenz in unserer Gesellschaft ist und für Lehre und Forschung unabdingbar bleibt. Auch wäre eine solche Lösung schlicht nicht umsetzbar, da zu zeitraubend (z.B. Semesterprüfungen an der Uni).
4. Selber ausprobieren (und neue Erkenntnisse in der Fachschaft teilen): die neuen Möglichkeiten ausprobieren und bei schriftlichen Aufgaben zuerst schauen, welche Produkte mit den KI-Tools möglich sind, um dann die Aufgaben anzupassen oder zu verändern.

tippt, als ob man der Maschine beim Denken zuschauen würde. Tatsächlich generiert der Chatbot lediglich statistisch häufig beieinanderstehende Wortfolgen, so wie er sie in den riesigen Datenbanken vorfindet.

Die KI denkt nicht selbständig, sondern analysiert und synthetisiert die eingelesenen Trainingsdaten und gibt die am meisten verknüpften Resultate wieder (die Gratisversion von ChatGPT sogar nur mit Internetdaten bis 2021). So ist es verständlich, dass immer das mittelmässigste Resultat erzeugt wird, jenes nämlich, welches im Zusammenhang mit der eingegebenen Anfrage am wahrscheinlichsten existieren könnte.

Unterdessen sind weitere KI-Suchmaschinen auf dem Markt, welche ihre Informationen auch im Internet beziehen (Bing von Microsoft, Bard von Google, Ernie) und bereits wird KI in Apps integriert (Siri, Alexa, Spotify, Uber) und liefert ausser Texten auch Codes, Videos, Musik, Bilder (Midjourney, DALL-E, crayon) oder auch ganze Bibliografien (elicit.org) und weiterführende Fragestellungen mit genauen Quellenangaben (perplexity.ai). Die Anwendungs- und Verknüpfungsmöglichkeiten werden täglich mehr und nehmen uns die lästigen Routinearbeiten im Alltag ab (Mails, Protokolle, Tabellen, Forschungsberichte).

ChatGPT ist dort stark, wo es darum geht, Texte zu generieren: Das Tool kann zusammenfassen, vereinfachen, genrespezifisch schreiben und Schreibstile kopieren, Rollenspiele imitieren, Fragen aus einem Text extrapolieren, typische Lehrmittelaufgaben generieren (Lückentexte, Vokabellisten, Verständnis- und Prüfungsfragen, Multiple Choice) sowie Brainstorming simulieren und Pro-Contra Argumente auflisten. Unsere Schüler:innen und Studierenden nutzen ChatGPT vor allem als Ideensammlung, als Lese- oder Lernhilfe (Zusammenfassungen, Prüfungsfragen, Prüfungsgespräche im Rollenspiel) oder als Korrektur- und Stilassistent bei schriftlichen Arbeiten.

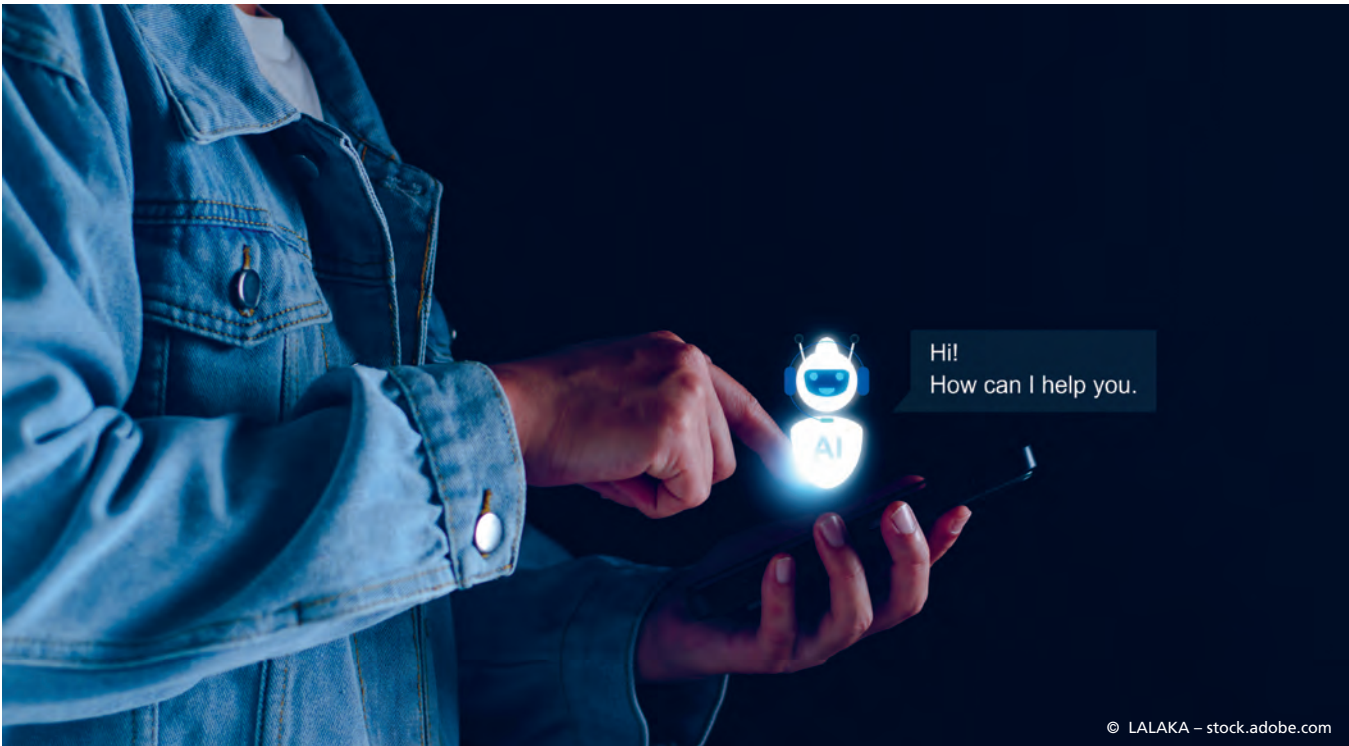
ChatGPT als Lehrperson nutzen

Aber auch die Lehrpersonen können für ihren Unterricht vom Schreibassistenten profitieren: eine Theorie oder eine Textpassage vereinfachen oder kurz erklären lassen, Beispielsätze erfinden mit einer bestimmten Verbform, differenzierte Aufgabenstellungen ausformulieren usw. Der Chatbot kann ebenso als Korrekturasistent (mit Bewertungskriterien) genutzt werden und Teile des Feedbacks an die Lernenden übernehmen. Die KI

hilft auch dabei, Sprachspiele zu erfinden (Taboo, Rollenspiele) oder erstellt die gesamte Projektplanung mit Lernzielen, Zeitplan und Bewertungsvorschlägen. Die Einsatzmöglichkeiten sind je nach Fach verschieden und es gilt, diese vorher selbst auszuloten, um mögliche KI-generierte Schülerprodukte in der Bewertung schon voranzuplanen und die Aufgabenstellung dementsprechend zu verändern oder komplexer zu gestalten.

Einsatzmöglichkeiten von KI und Chatbots im Sprachunterricht

- Die Schüler:innen gegen die KI antreten lassen, z.B. für literarische Analysen, Sprachmittlung oder Interpretation: Welcher Text geht besser auf den Inhalt und den geschichtlichen Hintergrund ein? Welche Analyse ist kreativer und eröffnet neue Gedankengänge?
- Mit der Klasse über die Qualität von Texten diskutieren: Wie unterscheidet sich ein mittelmässiger von einem qualitativ guten Text? Kann eine glänzende Oberfläche den Inhalt ersetzen?
- Chatbot als Lernpartner bei Prüfungsvorbereitungen nutzen: Prüfungsfragen erstellen lassen und Antworten verfassen, danach mit jenen des Chatbots vergleichen.
- Grammatik/Theorie erklären und passende Beispielsätze oder Lernkontrollen generieren lassen.
- Chatbot für Rollenspiele nutzen, um Dialoge einzuüben. Allerdings kennt die KI momentan noch nicht die GER-Sprachniveaus und formuliert meist zu perfekt resp. zu schwierig für die Lernenden mit Niveau A2 oder B1.
- Aufsätze korrigieren und bewerten lassen (auch das Kriterienraster eingeben) und die eigene Korrektur mit jener des Chatbots vergleichen: Welches Feedback ist in welcher Phase nützlicher für den Schreibprozess? Wie lassen sich die unterschiedlichen Korrekturen erklären?
- Schreibstil, Aussprache, Vokabular durch ChatGPT korrigieren und kommentieren lassen: Die Schüler:innen verbessern den Text nach den Anweisungen der Maschine.
- Präzises Prompting einüben: z.B. mehrere Anfragen zeitversetzt mit demselben Prompt starten und vergleichen. Variante: verschieden formulierte Anfragen zu demselben Thema gegeneinander abwägen.
- Gute Fragestellungen üben: die Antworten verschiedener Chatbots vergleichen (ChatGPT, perplexity.ai, Bing), die bessere Fragestellung wählen und die Auswahl rechtfertigen.



© LALAKA – stock.adobe.com

Als Instrument für den inklusiven Unterricht ist ChatGPT unschlagbar: Geduldig erklärt das Tool wieder und wieder die problematischen Textstellen, zeigt Schritt für Schritt den Lösungsweg an, vereinfacht die schwierige Theorie oder die sprachlich zu komplexen Fragestellungen. Dank dieser Hilfestellungen können auch Lernende mit Lernschwächen oder ungenügenden Sprachkenntnissen im Lernprozess unterstützt werden. Maschine und Lehrperson können in einer solchen individualisierten Lernbegleitung quasi «Hand in Hand» arbeiten. Interessanterweise geben die Lernenden in Umfragen an, dass sie einerseits das zeitnahe Feedback der KI schätzen, jedoch möchten sie keinesfalls einzig und allein durch eine Maschine beurteilt werden. Sie empfinden die Aussagen als zu generalisierend und ohne Bezug zu ihrer Person oder zum Kontext – also genau dieselbe Kritik, welche wir gegenüber KI-generierten schriftlichen Erzeugnissen haben. Hier treffen sich die Einschätzungen der Lernenden mit jenen der Lehrenden.

Grenzen von ChatGPT

Die KI hat keinen wirklichen Bezug zur physischen respektive physikalischen

Realität, ihr Weltwissen besteht aus Wortblöcken und nicht aus echten Erfahrungen und Beziehungen. Deshalb kann weder DeepL noch ChatGPT auf den Kontext eingehen, kennt weder Pragmatik oder Empathie, noch kann sie Mehrdeutigkeiten interpretieren oder übersetzen. Ja, man könnte sagen, dass dem Tool ein gewisses Mass an gesundem Menschenverstand fehlt, um in spezifischen Situationen adäquat zu reagieren.

Zudem ist ChatGPT so trainiert, dass es, ungeachtet der Richtigkeit des Resultats, wortreich und enthusiastisch zu allem eine Antwort liefert. Deswegen ist es unumgänglich, beim Prompten nebst den klaren Instruktionen auch immer genug Informationen zur Situation anzugeben sowie zur Rolle, die die Maschine in ihrer Antwort einnehmen soll. Ansonsten besteht die Gefahr, dass der Chatbot «halluziniert», d.h. dass er aus Wortblöcken eine virtuelle (oder falsche) Wahrheit generiert.

Berechtigtes Unbehagen

Die allgemeine Angst vor der Maschine ist teilweise berechtigt. Wenn wir nicht wissen, mit welchen Daten eine KI gefüttert wurde und keine Trans-

parenz im Umgang mit diesen Daten gewährleistet wird, dann kann die Maschine dubiose oder klischeehafte Inhalte reproduzieren und so wichtige Entscheidungen beeinflussen, z.B. bei der medizinischen Versorgung, Gerichtsurteilen, Bewerbungen usw. Und obwohl ChatGPT durch OpenAI auf einen extrem höflichen Output programmiert wurde, mit Leitplanken der Political Correctness, so gibt es keine Garantie, dass die KI-Tools nicht diskriminierende Allgemeinplätze reproduzieren.

Auch die Gefahr von Deep Fakes ist real: Es wird eine immer grössere Herausforderung, dies alles zu überschauen und zu durchschauen. Wer kann dann noch sagen, ob ein Foto echt ist oder ob die Stimme der Freundin computererzeugt ist? Wenn ein Mail in einem uns bekannten Soziolekt geschrieben wurde, erweckt dies unser Vertrauen und öffnet die Pforten für Betrug und Verschwörungstheorien. Bereits jetzt verunsichern und verstören uns erste KI-generierte Videos und Bilder in hyperrealistischer Optik und nur der gesunde Menschenverstand und unsere Erfahrung mit der realen Welt entlarven diese Produkte noch als Deep Fakes.

Konsequenzen für Schulen, Unterricht und Prüfungen

Wie sollen Bildungsinstitute damit umgehen, dass schriftliche Hausarbeiten und Projektarbeiten nun auch ohne Eigenleistung und mit wenig Aufwand abgegeben werden können? Der Ruf nach Verboten ist eine verständliche, jedoch nicht zielführende Reaktion, da sich technische Innovationen weder stoppen noch umgehen lassen. Auch die Plagiats-erkennung (mit oder ohne digitales Wasserzeichen) gestaltet sich äusserst schwierig und laufend werden neue Tools entwickelt, die diese Kontrollmechanismen umgehen können. Einige Bildungsinstitute setzen bei digitalen Prüfungen auf erhöhte Überwachung, andere sperren gewisse Seiten. Jedoch wird die KI bald in allen Suchmaschinen integriert sein und täglich entstehen neue, auf Deep Learning basierte Apps und Anwendungen.

ChatGPT und DeepL werden unsere Schule nicht umstürzen, aber wir werden diese Tools mit der Zeit in den Unterricht integrieren und unsere Ar-

beitsmethoden an die neuen Begebenheiten anpassen müssen. Zu unserem Bildungsauftrag gehört, dass wir unsere Schüler:innen auf das Studium und auf die Arbeitswelt vorbereiten, wo solche KI-Tools zum Standard gehören werden. Gleichzeitig sollen sie aber lernen, diese Technologien kritisch zu betrachten und sinnvoll einzusetzen.

Darum werden wir weiterhin die basalen Kompetenzen trainieren, trotz Textgeneratoren, Übersetzungstools, trotz Wikipedia, Taschenrechner oder anderer Hilfsmittel. Der Chatbot kann die Formulierungen zum Teil übernehmen, Gedanken und Ideen werden jedoch vom Menschen geliefert. Bei individuellen Vertiefungsarbeiten (Maturarbeiten) richtet sich der Fokus nun auf die Ideenfindung, den Lösungsweg und den Schreibprozess, also auf die kritische Auseinandersetzung mit einer komplexen Fragestellung, anstatt vor allem auf das fertige schriftliche Produkt. Die Zusammenarbeit von Mensch und Maschine lässt komplexere Aufgaben zu,

welche man früher so nie hätte lösen können. Und auch die ganze Prüfungsproblematik wird überdacht. Wer weiss: Vielleicht verzichten wir in Zukunft auf die eine oder andere «rein formelle» Schreibübung und konzentrieren uns auf die Sinnhaftigkeit, die Prüfung des Inhalts und die Eigenleistung?

Auch der Lehrplan 21 mit der Kompetenzorientierung hat den Fremdsprachenunterricht massgeblich verändert, weg von reproduziertem Wissen, hin zu produktivem, kreativem, kollaborativem Arbeiten und konkreten Anwendungen. Dank DeepL wurden bestimmte Prüfungen durch Lerntagebücher und Portfolios ersetzt. Nun ermöglichen auch ChatGPT und Co. neue Prüfungsformate, in denen nicht die Form allein massgeblich ist für die Evaluation, sondern die Kompetenz und die Lernprogression. Die Chatbots sollen nicht das Denken übernehmen, sondern beim Denken helfen.

Links zum Ausprobieren (nur zum Teil als Gratisversion verfügbar)

Textgeneratoren:

- ChatGPT3
- ChatGPT4 (OpenAi)
- Bing (Microsoft)
- Bard (Google)

Bildergeneratoren:

- DALL-E
- Crayon
- Midjourney

Quellenrecherche:

- Elicit.org

Schreibassistent:

- DeepL Write (Beta-Version)
- YouWrite

Fragestellungen:

- perplexity.ai