

BABYLONIA

1|2025

*Rivista per l'insegnamento e
l'apprendimento delle lingue*

*Zeitschrift für Sprachunterricht
und Sprachenlernen*

*Revue pour l'enseignement et
l'apprentissage des langues*

*Rivista per instruir ed emprendre
linguatgs*

*A Journal of Language Teaching
and Learning*

WWW.BABYLONIA.ONLINE

**L'IA générative dans
l'enseignement des
langues étrangères**

**Generative AI in
foreign language
teaching**

**L'IA generativa
nell'insegnamento
delle lingue straniere**

**Generative KI im
Fremdsprachen-
unterricht**





L'IA générative dans l'enseignement des langues étrangères
Generative AI in foreign language teaching
L'IA generativa nell'insegnamento delle lingue straniere
Generative KI im Fremdsprachenunterricht

Responsabili della parte tematica:
Sabine Christopher, Laura Loder Buecher
& **Verónica Sánchez Abchi**

Babylonia

Rivista svizzera per l'insegnamento delle lingue

Trimestrale plurilingue
edito dalla
Associazione Babylonia Svizzera
cp 120, CH-6949 Comano
ISSN 1420-0007
no 3/anno XXX/2023

Con il sostegno di



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Dipartimento federale dell'interno DFI
Departament federal da l'intern DFI

Bundesamt für Kultur BAK
Office fédéral de la culture OFC
Ufficio federale della cultura UFC
Uffizi federal da cultura UFC



Repubblica e
Cantone Ticino



FONDATION
OERTLI
STIFTUNG

PHBern
Pädagogische Hochschule

Zürcher Hochschule
für Angewandte Wissenschaften



L'IA générative dans l'enseignement des langues étrangères

Generative AI in foreign language teaching

L'IA generativa nell'insegnamento delle lingue straniere

Generative KI im Fremdsprachenunterricht

6	Editoriale della redazione	22	Search 'A2, Liebesleben, 500 words'. Teachers' use of AI-based tools for the GFL reading class Anne Raveling
8	Introduzione Sabine Christopher Laura Loder Buechel Verónica Sánchez Abchi	28	Un estudio sobre el uso de ChatGPT para resolver tareas escritas en la clase de ELE Nicola Brocca Rebeca Iniesta Jiménez
10	Imparare le lingue con l'IA generativa: vincere il Giro d'Italia allenandosi su un'e-bike? Anna-Maria De Cesare	32	Integration KI-basierter Schreibtools im Fremdsprachenunterricht: Chancen und Herausforderungen für Lernende auf niedrigem Sprachniveau Brigitte Reber Meike Raaflaub
12	Apprendre des mots en FLE avec ChatGPT Raphael Berthele	38	Exploring the Potential and Challenges of ChatGPT in Language Education: An Investigation into Student Teachers' Critical Thinking, Interactions and Task-Based Activity Design Corinne Ca' Foscari Baradel
18	Un classificateur de cicéronianité pour travailler sur les idiomes latins Damien Cavaleri Micha Hersch Antje Kolde		

BABYLONIA

1|2025

44

Enlisting ChatGPT to Enhance Language Teacher Training

Holli Schaubert

50

Using Artificial Intelligence for Enhanced Lesson Planning and Material Development: Transforming Teaching Practices

Ferit Kılıçkaya

56

Comment développer la littératie digitale des enseignant-es et des apprenant-es? Conclusions du projet 'Digital Literacy in University Contexts'

Sara Cotelli Kureth
Alice Delorme Benites
Caroline Lehr
Hasti Noghrechi
Elizabeth Steele
Elana Summers

60

English Quest: Exploration du potentiel pédagogique des IA génératives dans l'apprentissage des langues

Guillaume Tschupp
Erica Borloz

66

Student Interactions with Educational AI Chatbots in Language for Specific Purposes: Insights from Usage and Perspectives

Alice Gruber

70

Project Tools@ Schools: immersive – Tasks for a critical use of generative AI in immersive foreign language sequences in subject teaching

Catherine Ferris
Sandrine Wild
Stefano Negrinelli

76

Per una didattica delle lingue straniere connettivista

Sara Alloatti
Filomena Montemarano

Babylonia, juste une belle idée...?

Lors de sa fondation en 1991, Babylonia était une idée. Une idée née dans les esprits d'un groupe d'humanistes engagés pour la construction d'une société *plurilingue, multiculturelle, ouverte et tolérante* passant selon eux par l'enseignement et l'apprentissage des langues étrangères.

Editoriale

Trente-quatre ans plus tard, Babylonia est toujours là, et son équipe de rédaction s'engage toujours et encore pour un dialogue constructif entre chercheur.es, didacticien.nes et enseignant.es ainsi que pour la promotion d'un enseignement des langues adapté aux besoins et particularités de chacun. Et ce en gardant à l'œil les évolutions positives et négatives de la société et en discutant de leur influence sur l'enseignement et l'apprentissage des langues – le présent numéro en est un bel exemple.

Pourtant, depuis plus de 30 ans, Babylonia doit aussi se battre pour assurer sa subsistance. Trente-quatre ans maintenant que les membres de la rédaction investissent leur temps pour produire des numéros de qualité. Et ceci bénévolement pour la plupart d'entre eux.

Un numéro de Babylonia représente des centaines d'heures de travail: réflexion sur les thématiques, choix des résumés les plus prometteurs, écriture des articles par leurs auteurs, relectures et révisions, traductions, mise en page, diffusion.

L'intelligence artificielle pourra(it) nous aider pour une partie de ce travail, mais, pour l'instant du moins, nous faisons le pari de l'intelligence humaine pour alimenter les réflexions et pratiques de manière critique et informée.

Nous remercions les sponsors qui nous soutiennent depuis plus de 30 ans (en particulier l'office fédéral de la culture), nos partenaires institutionnels (Institut de plurilinguisme et CeDiLE), ainsi que les institutions qui voient l'importance de soutenir le travail humain: En 2024, la HEP Vaud, la PHZH, la PHZG, la BCUL et le canton du Tessin. En 2025, la PHZH, la PH Bern, la BCUL, et la Fondation Oertli (qui soutient ce numéro en particulier: merci!).

Merci aussi aux membres qui nous aident par leurs cotisations, ainsi qu'à certains généreux donateurs qui se reconnaîtront.

Comme le souligne l'une de nos lectrices en réaction à la mise en ligne graduelle de nos archives (*allez y jeter un œil!*), c'est «*génial que toute la richesse diffusée par votre revue soit désormais accessible à tout le monde!* 🍷»

Une richesse, oui! Mais une richesse pour laquelle nous nous battons quotidiennement – et, nous osons l'espérer, intelligemment.

Nous vous souhaitons une excellente lecture de ce numéro sur une thématique plus qu'actuelle, ainsi qu'une bonne plongée dans nos archives pour une compréhension diachronique de cette fameuse *idée* Babylonia!

Babylonia, solo una bella idea...?

Quando, nel 1991, è stata fondata Babylonia, si trattava di un'idea. Un'idea nata nella mente di un gruppo di umanisti impegnati a costruire *una società multilingue, multiculturale, aperta e tollerante* anche grazie all'insegnamento e l'apprendimento delle lingue straniere.

Trentaquattro anni dopo, Babylonia è ancora fra noi e la sua redazione è tuttora impegnata in un dialogo costruttivo tra ricercatori, pedagogisti e insegnanti, promuovendo un insegnamento delle lingue che si adatti alle esigenze e alle particolarità di ogni apprendente. Lo facciamo tenendo d'occhio gli sviluppi positivi e negativi della società e discutendo la loro influenza sull'insegnamento e sull'apprendimento – il presente numero ne è un bell'esempio.

Eppure, per oltre 30 anni, Babylonia ha dovuto lottare per la propria sopravvivenza. Da trentaquattro anni, i membri della redazione dedicano il loro tempo alla produzione di numeri di qualità. E la maggior parte di loro lo fa su base volontaria.

Realizzare un numero di Babylonia richiede centinaia di ore di lavoro: dalla riflessione sulla tematica, alla scelta delle proposte di contributo più promettenti, alla stesura degli articoli da parte degli autori, alla correzione delle bozze, all'editing, alle traduzioni, all'impaginazione e alla distribuzione.

L'intelligenza artificiale potrà/potrebbe aiutarci in parte in questo lavoro, ma, almeno per ora, ci affidiamo ancora all'intelligenza umana per fornire un contributo critico e informato alle nostre riflessioni e alla nostra pratica.

Ringraziamo gli sponsor che ci sostengono da oltre 30 anni (in particolare l'Ufficio federale della cultura), i nostri partner istituzionali (l'Istituto di plurilinguismo e il CeDiLE) e le istituzioni che riconoscono l'importanza del lavoro umano: nel 2024, l'HEP Vaud, la PHZH, la PHZG, la BCUL e il Canton Ticino. Nel 2025, la PHZH, la PH Bern, la BCUL e la Fondazione Oertli (che sostiene in particolare questo numero: grazie!).

Un ringraziamento speciale va anche alle socie e ai soci che ci sostengono con i loro abbonamenti e a alcuni generosi donatori che si riconosceranno.

Come ha detto un nostro lettore in risposta alla progressiva disponibilità online dei nostri archivi (*andate a dare un'occhiata!*), è «*geniale che tutta la ricchezza diffusa dalla vostra rivista sia ora accessibile a tutti!* 🍷»

Ricchezza, sì! Ma è una ricchezza per la quale lottiamo ogni giorno – e, osiamo sperare, con intelligenza.

Vi auguriamo una piacevole lettura di questo numero su un tema di grande attualità, e vi invitiamo a esplorare i nostri archivi per cogliere, anche in chiave diacronica, questa famosa *idea* di Babylonia!

BA

BY

Babylonia, nur eine schöne Idee...?

Als Babylonia 1991 gegründet wurde, ging es um eine Idee. Eine Idee, die in den Köpfen einer Gruppe von Humanisten geboren wurde, die sich für den Aufbau einer *mehrsprachigen, multikulturellen, offenen und toleranten Gesellschaft* engagierten, die ihrer Meinung nach durch das Lehren und Lernen von Fremdsprachen entstehen sollte.

34 Jahre später ist Babylonia immer noch da und das Redaktionsteam setzt sich weiterhin für einen konstruktiven Dialog zwischen Forschenden, Didaktiker:innen sowie Lehrpersonen ein und auch für einen auf die Bedürfnisse und Besonderheiten der einzelnen Lernenden abgestimmten Sprachunterricht. Dabei werden positive und negative Entwicklungen in der Gesellschaft im Blick behalten und deren Auswirkungen auf das Sprachenlernen und -lehren diskutiert – ein schönes Beispiel dafür ist die aktuelle Ausgabe.

Trotzdem muss Babylonia seit über 30 Jahren auch um das Fortbestehen kämpfen. Seit nun 34 Jahren investieren die Redaktionsmitglieder ihre Zeit, um qualitativ hochwertige Ausgaben zu produzieren. Und das für die Mehrheit unter ihnen unentgeltlich.

Eine Babylonia-Ausgabe bedeutet hunderte Arbeitsstunden: Überlegungen zu den Themen, Auswahl der vielversprechendsten Abstracts, Verfassen der Artikel durch die Autor:innen, Korrekturlesen und Überarbeitungen, Übersetzungen, Layout, Verbreitung.

Künstliche Intelligenz kann (könnte) uns bei einem Teil dieser Arbeit unterstützen, aber zumindest für den Moment setzen wir auf menschliche Intelligenz, um kritische und fundierte Diskussionen und Praktiken anzuregen.

Wir danken den Sponsor:innen, die uns seit über 30 Jahren unterstützen (insbesondere das Bundesamt für Kultur), unseren institutionellen Partner:innen (Institut für Mehrsprachigkeit und CeDiLE) sowie den Institutionen, die die Bedeutung der Unterstützung von menschlicher Arbeit erkennen: Im Jahr 2024 die HEP Vaud, die PHZH, die PHZG, die BCUL und der Kanton Tessin. Im Jahr 2025 die PHZH, die PH Bern, die BCUL und die Oertli-Stiftung (welche diese Ausgabe unterstützt: danke!).

Ein grosses Dankeschön geht auch an die Mitglieder, die uns mit ihren Beiträgen unterstützen, sowie an einige grosszügige Gönner:innen, die sich selbst wiedererkennen werden.

Wie eine unserer Leserinnen als Reaktion auf die schrittweise Onlinestellung unserer Archive ([schauen Sie doch mal rein!](#)) feststellte, ist es „*einfach toll, dass all die Fülle, die eure Zeitschrift bietet, jetzt für alle zugänglich ist!*“ 🍌

Eine Fülle, ja! Aber eine Fülle, für die wir uns täglich und hoffentlich auf intelligente Weise einsetzen.

Wir wünschen Ihnen eine spannende Lektüre dieser Ausgabe zu einem höchst aktuellen Thema und einen guten Streifzug durch unsere Archive, um die diachrone Entwicklung der oben genannten Idee Babylonia besser zu verstehen!

LO

Babylonia, mo ina bial'idea...?

Cu Babylonia ei vegnida fundada 1991 s'extractava ei d'in'idea. In'idea ch'ei naschida els tgaus d'ina gruppa da humanists che s'engaschavan per la creaziun d'ina *societad plurilingua, multiculturala, aviarta e toleranta* che dueva tenor els crescher entras l'instrucziun ed igl emprender da lungatgs jasters.

34 onns pli tard ei Babylonia aunc adina cheu ed il team da redacziun s'engascha vinavon per in dialog constructiv denter perscrutader-ra-s, didacticher-a-s e scolast-a-s ed era per in'instrucziun da lugatgs adattada als basegns ed allas particularitads da mintgin. En quest connex vegnan svilups positivs e negativs ella societad teni en egl e las consequenzas da quels sigl emprender e sill'instrucziun da lungatgs discutas – in bi exempel ei l'ediziun actuala.

E tuttina ha Babylonia dapi varga 30 onns era da sbatter per sia existenza. Dapi 34 onns investeschan ils commembers dalla redacziun lur temps per producir ediziuns d'alta qualitat. E quei senza paga per la gronda part dad el-la-s.

In'ediziun da Babylonia munta tschiens uras lavur: ponderaziuns davart ils temas, elecziun dallas resumaziuns las pli empermettontas, il scriber dils artechels davart dils-dallas autur-a-s, lecturas e correcturas, translaziuns, layout, diffusiu.

L'intelligenza artificziala sa(vess) sustener nus tier ina part da quella lavur, mo silmeins pil mument seconcentrein nus sill'intelligenza humana per promover discussiuns e praticas criticas e fundadas.

Nus engraziein als sponsurs che sustegnan nus dapi passa 30 onns (cunzun igl uffez federal da cultura), a nos partenaris instituziunals (Institut da plurilinguitad e CeDiLE) sco era allas instituziuns che renconuschan l'impurtonza dil sustegn dalla lavur humana: Egl onn 2024 la HEP Vaud, la PHZH, la PHZG, la BCUL ed il cantun Tessin. Egl onn 2025 la PHZH, la PH Bern, la BCUL e la fundaziun Oertli (che sustegn quest'ediziun: engrazie!).

In grond engraziel era als commembers che sustegnan nus cun lur contribuziuns sco era ad in pèr fautors generus che vegnan a renconuscher sesezs.

Sco ina da nossas lecturas ha constatau sco reacziun sin il metter online pass per pass da nos archivs ([mei a mirar inaga!](#)), eisi “*semplamein super che tut l'abundanza che vossa revista porscha ei ussa accessibla a tut-ta-s!*” 🍌

In'abundanza, gie! Mo in'abundanza per la quala nus s'engaschein mintga di – e speronza era perdertamein.

Nus giavischain a Vos ina lectura captivonta da quest'ediziun sur d'in tema fetg actual ed ina buna scuvretga da nos archivs per capir meglier il svilup diacron dall'idea Babylonia!

NIA

L'IA GÉNÉRATIVE DANS L'ENSEIGNEMENT DES LANGUES ÉTRANGÈRES

● Sabine Christopher
| Osservatorio
linguistico della
Svizzera italiana (OLSI)
Laura Loder Buechel
| HEP Zurich
Verónica Sánchez
Abchi
| l'IRD (Neuchâtel) et
Université de Genève.

L'intelligence artificielle (IA) constitue un champ de recherche depuis plusieurs décennies, mais les récents progrès ont accéléré son évolution de manière spectaculaire. Depuis l'annonce de la création de ChatGPT en 2022, toutes les alarmes ont été déclenchées, et la rapidité avec laquelle de nouveaux outils d'IA, toujours plus performants, sont développés interroge et transforme tous les secteurs, de la science à l'art, en passant par de nombreux aspects de notre vie quotidienne. Le domaine de l'enseignement des langues n'est bien sûr pas en reste.

L'IA soulève des questions éthiques, écologiques, techniques, sociales et même existentielles indéniables. Les institutions proposent des outils et des guides pour mieux aborder les enjeux de l'IA dans l'éducation¹. Mais quelle que soit la perspective adoptée, elle révolutionne profondément notre expérience en tant qu'enseignant.e.s de langues. Dans ce numéro, nous nous penchons sur les expériences de recherche et d'enseignement qui aident les enseignant.e.s à utiliser

efficacement l'IA générative. Notre objectif est d'explorer l'impact – potentiellement bénéfique – que l'IA peut apporter à notre façon d'enseigner, d'apprendre, de travailler et de faire de la recherche dans le domaine des langues.

Nous ouvrons la discussion avec un entretien avec Anna Maria de Cesare, chercheuse qui a récemment participé à la création d'une revue de linguistique spécialisée dans les études sur les textes produits par l'IA. Dans cet entretien, elle nous parle de l'intérêt de l'IA pour la linguistique en général et pour la linguistique appliquée en particulier.

Les différents articles du numéro explorent des thématiques très diverses. Il est néanmoins possible d'identifier trois axes principaux autour desquels les contributions se regroupent :

1. **L'IA en salle de classe**
2. **L'IA au service de la formation**
3. **Les nouveaux outils d'IA**

¹ Au niveau internationale, le « [Guide pour l'IA générative dans l'éducation et la recherche](#) » de l'UNESCO; au niveau Suisse, les « [Recommandations pour une utilisation réfléchie et mesurée dans l'enseignement](#) » de la CIIP, entre autres, disponible [ici](#).

« L'IA soulève des questions éthiques, écologiques, techniques, sociales et même existentielles indéniables. [...] Quelle que soit la perspective adoptée, l'IA révolutionne profondément notre expérience en tant qu'enseignant.es de langues. »

L'IA en salle de classe

Ce premier axe présente des expériences concrètes d'enseignement qui mobilisent l'IA et analysent son impact sur l'apprentissage.

L'article de Berthele explore l'apport de ChatGPT dans l'apprentissage lexical d'une langue étrangère, comparé à des techniques classiques, comme les cartes de vocabulaire.

Le travail de Cavaleri, Hersch et Kolde, qui propose un outil pour évaluer les caractéristiques stylistiques de productions écrites en latin, montre comment exploiter l'IA pour l'enseignement des langues anciennes.

La contribution de Raveling propose des pistes pour utiliser LATILL dans l'enseignement de l'allemand langue étrangère.

De leur côté, Brocca & Iniesta Jimenez examinent l'impact de ChatGPT sur les tâches de production écrite en espagnol, tandis que Raaflaub et Reber proposent un guide pratique pour intégrer des instruments de traduction, basés sur l'IA, dans les tâches d'écriture dans l'enseignement de niveau débutant.

L'IA au service de la formation

Ce deuxième axe explore comment l'IA peut être mise au service de la formation des enseignant.es, un domaine central pour la profession.

L'article de Baradel, celui de Schaubert, ainsi que la contribution de Kilickaya s'inscrivent dans cette perspective. Ces travaux analysent comment l'IA peut alléger et faciliter la planification et la production d'activités pédagogiques, tout en soulignant la nécessité d'une approche critique et réflexive. Dans cette même

lignée, le travail de Cotelli et de son équipe met en avant l'importance d'une « littératie numérique » et la nécessité de fournir des outils « méta » pour une gestion appropriée de l'IA dans la formation.

Les nouveaux outils d'IA

Le dernier axe regroupe les articles qui introduisent des outils novateurs pour l'enseignement des langues et nous donnent l'occasion d'inaugurer une nouvelle section dans la revue, *Nuovi strumenti*.

Tschupp et Borloz présentent *English Quest*, un jeu vidéo dédié à l'apprentissage de l'anglais, tandis que Gruber décrit *Mizou*, un outil pour l'apprentissage de l'anglais des affaires. Ferris et al., pour leur part, proposent une réflexion sur la façon dont l'IA peut créer un environnement d'apprentissage linguistique « immersif ».

Finalement, et de manière transversale, sans entrer dans un axe spécifique, la contribution d'Alloatti et Montemarano analyse plusieurs enjeux : comment l'IA reconfigure l'enseignement des langues, quel est le rôle de l'enseignant.e, et ce que l'IA suppose en termes de conception pédagogique.

En somme, ces contributions montrent que, même si l'IA pourrait remplacer certains aspects du travail des enseignant.es, elle demeure avant tout un outil au service de la pratique. Ce sont les pratiques qui se transforment, le rôle de l'enseignant.e qui évolue, et, en fin de compte, la nature même du métier qui se reconfigure progressivement. Voici notre modeste contribution pour encourager la réflexion dans ce processus de transformation.

Sabine Christopher (Phd en Sciences de la communication) est collaboratrice scientifique à l'Osservatorio linguistico della Svizzera italiana (OLSI)

Laura Loder Buechel (Phd en Linguistique) est enseignante de didactique de l'anglais à la HEP Zurich.

Verónica Sánchez Abchi (Phd en Linguistique) est collaboratrice scientifique à l'IRDP (Institut de Recherche et de documentation pédagogique) et enseignante de didactique de l'espagnol à l'Université de Genève.

IMPARARE LE LINGUE CON L'IA GENERATIVA: VINCERE IL GIRO D'ITALIA ALLENANDOSI SU UN'E-BIKE?

● Anna-Maria De Cesare
| TU Dresden

Testi generati dall'IA sotto la lente della ricerca linguistica

L'interesse di ricerca di Anna-Maria De Cesare è rivolto alla grammatica dell'italiano, anche in prospettiva contrastiva con altre lingue, alla sociolinguistica e alla linguistica del testo. Per *Babylonia* ha recentemente co-curato il numero tematico 3/2021 *La rappresentazione delle donne nelle pratiche didattiche, nei discorsi, nelle lingue* (assieme a Matteo Casoni). Ha al suo attivo numerose pubblicazioni su un'ampia varietà di generi testuali e attualmente si focalizza sui testi generati automaticamente. Le rivolgiamo tre domande sulle sue attuali attività di ricerca.

1. Qual è l'interesse dell'IA sia per la linguistica generale sia per la linguistica applicata, in particolare per l'insegnamento e l'apprendimento delle lingue?

L'interesse è molto grande in entrambi i casi. Per quanto riguarda innanzitutto la linguistica generale,

l'IA cosiddetta *generativa* ci chiede di tornare a riflettere su alcune domande cruciali. Tra queste: qual è la differenza tra l'acquisizione di una lingua (nativa) da parte di un essere umano (bambino) e da parte di una macchina, in particolare dei modelli linguistici di grandi dimensioni, come GPT-3.5 e GPT-4 (alla base di ChatGPT)? O ancora: quali lingue, o varietà di lingue, sono in grado di riprodurre in modo soddisfacente questi modelli linguistici, programmati per riprodurre la scrittura umana?

Nel campo della linguistica applicata, l'IA generativa può essere sfruttata a fini molto diversi. I modelli linguistici possono aiutarci a pianificare il programma di un intero corso, ma anche a creare esercizi per una specifica lezione, di grammatica, di lettura e così via. ChatGPT, inoltre, può anche fungere da partner 'conversazionale', in scambi dialogici che si svolgono per iscritto. Un uso dei modelli linguistici che trovo particolarmente



Anna-Maria De Cesare è Professoressa Ordinaria di linguistica romanza presso la Technische Universität Dresden.

rilevante consiste nel proporre unità didattiche che mirano a valutare le abilità scritte di ChatGPT, che si pongono cioè come obiettivo quello di allenare lo spirito critico di chi impara a scrivere. Le domande che l'insegnante si può porre assieme a chi è in classe sono, tra le altre, le seguenti: qual è la qualità dei testi generati da ChatGPT? ChatGPT è in grado di correggere un testo scritto da chi è in classe? Dove e come corregge? Queste domande, a loro volta, permettono di riflettere su questioni di portata più generale: è utile usare strumenti come ChatGPT per scrivere o correggere un proprio testo?

2. Hai recentemente lanciato una nuova rivista scientifica, *AI-Linguistica. Linguistic Studies on AI-Generated Texts and Discourses (AI-Ling)*, dedicata ai testi generati dall'intelligenza artificiale considerati nell'ottica delle varie branche della linguistica. Puoi descrivere il progetto?

Il principale obiettivo della rivista *AI-Linguistica*, nata all'inizio del 2024, consiste nel promuovere lo studio, nei vari indirizzi della linguistica, di una nuova tipologia testuale: quella dei testi generati in modo automatico. Per via della sua novità, questa tipologia ha confini esterni e interni ancora poco netti. Ne fanno parte non solo i testi scritti generati *ex novo* con sistemi basati sull'intelligenza artificiale, come quelli prodotti con l'ausilio di ChatGPT, ma anche le traduzioni automatiche basate su reti neurali e i messaggi vocali prodotti da diversi tipi di chatbot e assistenti virtuali, come Alexa e Siri.

AI-Linguistica accoglie contributi incentrati sia su fenomeni linguistici, preferibilmente legati alle lingue romanze e germaniche, sia sui discorsi che ruotano attorno ai testi generati. Oltre ai classici articoli scientifici, accetta anche contributi brevi (tra le 3'000 e le 6'000 parole), che fanno il punto su un fenomeno o un concetto specifico. Il primo articolo breve pubblicato nella rivista è una riflessione critica sul concetto e sul termine di *allucinazione* (<https://doi.org/10.62408/ai-ling.v1i1.9>).

La rivista è *Diamond Open Access*: non ha fini di lucro, è gratuita sia per chi pubblica un contributo sia per chi lo legge. Invito chi volesse saperne di più a consultare il sito della rivista: https://ai_ling.journals.qucosa.de/ai_ling/about

3. Alla TU Dresden hai organizzato la seconda edizione del convegno "Automated texts In the ROMance languages and beyond" (AI-ROM-II) che si è tenuta il 2 e 3 settembre 2024. Ci puoi parlare di questo convegno: quali sono le linee di ricerca e quali acquisizioni più recenti potrebbero avere delle ripercussioni sull'apprendimento e l'insegnamento delle lingue?

Anche il convegno, come la rivista *AI-Linguistica*, mira a promuovere e stimolare la ricerca sui testi generati nel campo della linguistica romanza. Nei due ambiti citati (apprendimento e insegnamento delle lingue) è importante conoscere bene sia le potenzialità (di cui si è già parlato sopra) sia i limiti dell'IA generativa. Si sostiene a volte, in modo semplicistico, che i testi prodotti dai modelli linguistici siano paragonabili o addirittura indistinguibili da quelli scritti da esseri umani. La ricerca attuale sta però mostrando che esistono in realtà molte spie linguistiche (di tipo lessicale, interpuntivo e grammaticale) e testuali che permettono di distinguere la scrittura prodotta da una macchina da quella di una persona in carne e ossa. **Conoscere l'inventario dei tratti distintivi dell'italiano generato si rivela oggi essenziale sia per chi insegna una lingua (anche come L2) sia per chi la impara.**

A chi impara una lingua (nel caso di cui stiamo parlando, nella sua veste scritta), questo inventario permette

di capire che i testi generati non possono servire da modello di buona scrittura: essendo il prodotto di un algoritmo, questi testi tendono ad assomigliarsi e a essere molto ripetitivi a tutti i livelli: nella scelta del lessico, nella costruzione delle frasi e perfino nella forma delle unità testuali (per esempio, i blocchi di testo tendono ad avere una dimensione simile). A chi insegna permette invece di valutare se il testo che sta leggendo è della mano di una persona che impara a scrivere o se è stato generato da un algoritmo. **In ambito educativo bisogna far capire che non vale la pena usare i testi generati in modo abusivo: l'acquisizione della scrittura, a qualsiasi livello, richiede una continua esercitazione e dunque molto impegno.** Non si può pretendere di vincere il Giro d'Italia allenandosi su un'e-bike...

Un altro grosso limite dei modelli linguistici, almeno di quelli attualmente disponibili, è che anche quelli più avanzati (come GPT-4o) sono addestrati su una base dati in cui l'inglese è sovrarappresentato. Gli studi mostrano che la natura dei dati di addestramento lascia delle 'impronte' più o meno visibili: i testi generati in italiano (ma lo stesso vale per quelli prodotti in francese, spagnolo ecc.) sono costellati di strutture che ricalcano quelle della lingua inglese. Si punta spesso il dito contro le cosiddette allucinazioni, relative agli errori di tipo contenutistico, ma sono altrettanto gravi gli errori e le distorsioni di tipo linguistico e testuale. Conoscere i limiti dell'italiano generato è molto importante, soprattutto per chi impara a scrivere in italiano come lingua seconda.

«Conoscere l'inventario dei tratti distintivi dell'italiano generato si rivela oggi essenziale sia per chi insegna una lingua (anche come L2) sia per chi la impara.»

APPRENDRE DES MOTS EN FLE AVEC CHATGPT

Ce studio s'intéresse à l'usage de ChatGPT pour l'apprentissage du vocabulaire en français comme langue étrangère. Sachant que l'apprentissage du vocabulaire à l'aide de cartes (flashcard) constitue l'un des méthodes les plus efficaces, nous le comparons avec une approche d'apprentissage basée sur l'utilisation de ChatGPT. Les participants sont 51 jeunes adultes en formation (niveau tertiaire, universitaire ou pédagogique HEP), pour la plupart germanophones.

Dans un premier moment, les participants devaient apprendre 20 mots avec deux supports différents (Quizlet et ChatGPT) et en quatre listes distinctes, afin de varier les mots étudiés avec les deux supports. La moitié des participants a travaillé d'abord avec Quizlet puis avec ChatGPT, tandis que l'autre moitié a suivi la procédure dans l'ordre inverse. Aux participants a été demandé d'utiliser les mots appris pour rédiger une lettre de réclamation concernant la Mensa de leur université. La rétention des mots a été finalement évaluée par un post-test une semaine plus tard.

Les résultats montrent que les mots appris avec Quizlet sont utilisés davantage dans le devoir écrit que les mots étudiés avec ChatGPT. Le post-test indique également que les participants se souviennent mieux des mots appris avec Quizlet. Nous discutons de ces résultats en nous posant la question de savoir dans quelle mesure les deux conditions d'apprentissage ont favorisé la création et l'approfondissement des traces mnémotechniques, en nous référant aux concepts de profondeur de traitement (*depth of processing*), de charge d'implication (*involvement load*), et aux théories qui considèrent les ressources limitées dans l'apprentissage du vocabulaire.

[traduit par ChatGPT]

Tema

Raphael Berthele | Université de Fribourg



Raphael Berthele est professeur de plurilinguisme à l'université de Fribourg. Il a étudié et travaillé aux universités de Fribourg, Tübingen, Berkeley et Berne. Il a cofondé l'Institut de plurilinguisme fribourgeois en 2008. Il est responsable de deux programmes de master en sciences du plurilinguisme et en didactique des langues étrangères. Ses recherches portent sur divers aspects du plurilinguisme, du plan cognitif au plan social. Ces dernières années, il s'est consacré à l'étude empirique du plurilinguisme réceptif, à l'acquisition de la littératie plurilingue, à l'expression des relations informationnelles et spatiales chez les bi- et plurilingues, et aux différences individuelles dans l'apprentissage des langues.

Contexte et questions de recherche

Les enquêtes menées auprès des enseignants-e-s de langues vivantes révèlent souvent que le développement des ressources lexicales en langue cible demeure un défi majeur. En comparaison de l'apprentissage incident du vocabulaire qui repose, par exemple, sur la lecture ou le visionnage de matériel audiovisuel, les études empiriques démontrent la plus grande efficacité de l'apprentissage délibéré et ciblé de mots, notamment par l'emploi de fiches (parfois appelées *fiches bristol*, angl. *flashcards*; cf. Webb et al. 2020). Dans le projet de recherche présenté ici, notre but était de comparer l'efficacité de l'apprentissage par fiches à une méthode alternative qui implique l'interaction des apprenant-e-s avec un système comme ChatGPT, désormais largement accessible depuis que les agents conversationnels se sont généralisés. Concrètement, nos deux questions étaient les suivantes :

Est-ce que l'interaction avec ChatGPT en tant que 'coach lexical' pour des mots difficiles...

(1) ... aide les apprenant-e-s à utiliser ces mots dans une tâche écrite ?

(2) ... permet d'apprendre ces mots de manière efficace ?

Notre hypothèse était que le travail avec un agent conversationnel permettrait aux apprenant-e-s de développer et d'approfondir leurs connaissances lexicales avec davantage d'engagement. En raison de la nouveauté de l'outil et de son caractère ludique, nous partions du principe qu'une activité interactive avec ChatGPT entraînerait un degré d'engagement plus élevé dans la tâche. Cette intuition se voyait renforcée par le fait que le traitement cognitif de l'output généré par cet outil devait naturellement inciter l'apprenant-e à évaluer la justesse et la pertinence du texte produit, ce qui conduirait à une meilleure mémorisation des items lexicaux, selon la théorie

de la charge d'implication ('involvement load'; Yanagisawa & Webb, 2022). L'implication, telle qu'elle est conçue dans cette approche théorique, comporte trois dimensions qui chacune contribue à l'apprentissage de nouveaux mots : Le besoin d'exprimer une idée dans la langue cible ('need'), la recherche de moyens linguistiques permettant de l'exprimer ('search') et la vérification du résultat de cette recherche lors de la mise en mots ('evaluation'). Dans le cadre de notre projet, les deux dimensions recherche et évaluation semblent particulièrement pertinentes. De plus, l'interaction avec l'agent conversationnel devrait, selon la théorie de la profondeur du traitement (Craik & Tulving, 1975), également favoriser la mémorisation, puisque la tâche décrite ci-dessous met l'accent principalement sur la dimension sémantique des mots.

Tâche et déroulement

La langue cible de notre expérience didactique était le français langue étrangère. Les participant-e-s étaient des apprenants adultes majoritairement germanophones, tous étudiant-e-s au sein d'une haute école (pédagogique ou universitaire) en Suisse. Nous avons conçu une tâche écrite dans une orientation communicative. Cette tâche consistait à rédiger un courriel au rectorat de la haute école, afin de contester les prix trop élevés pratiqués par la Mensa. De plus, la consigne prévoyait que d'autres doléances doivent figurer dans cette lettre (environ 250 mots), notamment la mise à disposition de davantage d'espace de pique-nique et de fours à micro-ondes.

Avant de récolter les données, nous avons fait générer à l'IA ChatGPT (gpt-3.5-turbo) plusieurs lettres sur la base de cette consigne. Dans ces lettres, nous avons identifié des entités lexicales difficiles à apprendre, d'abord de manière introspective et intuitive. Nous nous sommes abstenus de pré-tester la connaissance de ces mots, car un tel pré-test aurait potentiellement déjà contribué à l'apprentissage des items. D'autre part, pour l'expérience finale, nous n'avons sélectionné que des mots qui, selon l'outil Flelex¹, sont enseignés à des niveaux avancés (C1, C2). Ceci afin de minimiser la probabilité que ces mots fassent déjà partie du vocabulaire productif des apprenants.

« Dans la condition ChatGPT, un éventuel effet positif dû à l'output idiomatique en langue cible illustrant l'emploi des mots en contexte serait donc contrecarré par une surcharge informationnelle. »

*abordable / accroître / afin de / bondé /
chronophage / défavoriser / en dépit de
/ en outre / impérativement / judicieux
/ l'agacement (m.) / l'aménagement /
l'entrave (f.) / l'épanouissement / l'in-
fléchissement (m.) / le gaspillage / no-
tamment / pallier / soucieux / subvenir
aux besoins de qqn*

Au moyen de ces 20 items, nous avons composé quatre listes différentes qui ont été attribuées aux participant-e-s de manière aléatoire. Chaque participant-e a tenté de mémoriser la moitié des mots durant 15 minutes à l'aide des questions spécifiques posées au ChatGPT et l'autre moitié des items, pour la même durée, avec des fiches numériques (sur la plateforme Quizlet). Une moitié des participant-e-s a travaillé d'abord avec Quizlet et ensuite avec Chat GPT, alors que l'autre moitié des participants a suivi la procédure inverse.

Les apprenant-e-s étaient libres de travailler avec des fiches Quizlet en version français-allemand, français-italien ou français-anglais. La réalisation de la tâche avec ChatGPT était standardisée dans la mesure où les participant-e-s étaient priés d'utiliser pour chaque mot des consignes (*prompts*) préformulées. La première consigne consistait à demander à ChatGPT d'*expliquer le mot* «...» *de manière simple*. Puis, il fallait demander à l'outil d'*écrire 3 phrases avec le mot* «...» afin d'obtenir des exemples de son usage en contexte. Les réponses étaient sauvegardées au moyen d'un formulaire dans lequel les participant-e-s collaient l'output obtenu (voir l'encadré pour un exemple).

¹ <https://cental.uclouvain.be/cefrlex/flelex/search/>

(1) Explique le mot suivant en langage simple : **pallier**

Le mot « pallier » signifie trouver une solution temporaire à un problème ou une situation difficile, sans pour autant résoudre complètement le problème à long terme. C'est un peu comme mettre un pansement sur une blessure, cela aide à limiter les dégâts momentanément, mais cela ne guérit pas complètement la blessure.

(2) Fais trois phrases en français qui contiennent ce mot : **pallier**

1. Il faut trouver une solution pour pallier ce manque de personnel.

2. Les vitamines peuvent aider à pallier les carences alimentaires.

3. Les mesures d'urgence ont été mises en place pour pallier les conséquences de la catastrophe naturelle.

Les fiches Quizlet, rédigées de façon minimale, comportaient uniquement la traduction du mot français vers l'allemand, l'anglais ou l'italien. Dans l'élaboration de notre expérience, nous avons évoqué des alternatives, par exemple d'ajouter des phrases contenant le mot cible ou des images. Finalement, nous avons opté pour la solution la plus minimaliste afin de maximiser l'effet (hypothétique) de l'information beaucoup plus substantielle fournie par ChatGPT sur l'utilisation et l'apprentissage des mots.

Dans le cas de mots polysémiques, seule la traduction pertinente pour la tâche décrite ci-dessus était indiquée (p. ex. *pallier* – *ausgleichen/ovviare/compensate for*).

Afin de s'assurer que l'output généré par l'IA soit lu, les participant-e-s devaient répondre à de simples questions concernant les réponses fournies par ChatGPT (qualité de la description du sens ou pertinence des trois exemples). Les réponses des participant-e-s n'ont pas été analysées.

Dans la phase finale de la récolte, nous avons demandé aux participant-e-s d'écrire la lettre de réclamation concernant la mensa en les priant explicitement d'utiliser autant de mots que possible. Ces lettres manuscrites ont été transcrites par nos soins. L'analyse a ensuite principalement consisté à répertorier, dans chaque lettre, les emplois de mots appris. De manière plus approfondie, nous avons également observé si les emplois de ces mots avaient donné lieu à des erreurs et, le cas échéant, celles-ci ont été catégorisées en différents types (p. ex. orthographe, grammaire). Exactement une semaine après la rédaction de la lettre, les participant-e-s ont passé un post-test, afin de vérifier si les nouveaux mots français avaient bien été mémorisés. Ce test a opérationnalisé la profondeur des connaissances lexicales en demandant aux participants de traduire le mot cible en L1, et en leur donnant la possibilité de former une phrase dans la langue cible contenant ce mot.

Échantillon

Plusieurs groupes d'apprenant-e-s ont participé au projet de recherche. Les participant-e-s étaient des étudiant-e-s de divers programmes et facultés de l'Université de Fribourg ainsi que des HEP (Hautes Écoles pédagogiques) de Zurich et de Berne. Au total, 51 personnes ont pris part à l'étude. La moyenne d'âge était de 22,7 ans. Le niveau de français de ces apprenant-e-s se situait entre B1 et B2. Les participant-e-s avaient majoritairement l'allemand – seule ou combinée à une seconde langue – pour L1 (N = 38), mais d'autres L1 étaient également représentées, à savoir : l'italien (N=12), l'albanais (N=2), l'anglais et le tamoul (N=1 respectivement). Cinq participant-e-s se considéraient comme bilingues avec l'allemand et une autre langue (auto-évaluation).

35 étudiant-e-s (groupe expérimental) ont suivi le protocole décrit ci-dessus, alors que 16 autres n'ont travaillé que

« Nous avons constaté que les apprenant-e-s qui travaillaient avec ChatGPT étaient confronté-e-s à une activité complexe nécessitant le traitement d'une grande quantité d'informations. »

10 mots de la liste, mais au moyen des deux plateformes (Quizlet et ChatGPT). La constitution de ce 'groupe contrôle' était en fait une erreur en matière de design expérimental. Celle-ci s'est cependant avérée utile puisque ce groupe nous permet de comparer l'utilisation et la mémorisation de mots non-travaillés ainsi que de mots ayant été appris avec les deux méthodes.

Mots employés dans les textes

Les analyses des lettres rédigées par les participant-e-s indique que les mots appris dans Quizlet (moyenne de 3.3 mots utilisés) sont davantage employés que ceux étudiés grâce à ChatGPT (moyenne de 1.3 mots). La figure 1 permet de comparer l'emploi de chacun des 20 mots dans les deux conditions. Seul un mot (*afinde*) apparaît plus fréquemment dans les textes lorsqu'il est appris selon la condition ChatGPT. Le groupe contrôle (données non présentées dans le graphique), qui a travaillé les mêmes 10 mots avec les deux outils, utilise en moyenne 5.9 mots dans les textes rédigés. Les mots absents de la liste du groupe contrôle ne sont employés que sporadiquement (moyenne de 0.20).

Bien que les résultats que nous donnent à voir la statistique descriptive soient d'ores et déjà éloquentes, nous avons voulu vérifier l'impact de la condition sur l'utilisation des mots appris lors de la rédaction des lettres au moyen d'un modèle logistique mixte. Ce modèle a pour fonction de mesurer la probabilité de l'emploi d'un mot spécifique des listes de vocabulaire mémorisé dans le texte écrit. Le facteur fixe correspond à la condition (Quizlet vs. ChatGPT), alors que les facteurs aléatoires sont l'item lexical (ordonnées à l'origine aléatoires), ainsi que le/la participant-e (ordonnées à l'origine aléatoires et pentes par condition aléatoires). Le résultat montre un effet significatif ($p < 0.001$) de la condition (les chances logarithmiques augmentent de 1.5 pour la condition Quizlet). L'inclusion des L1 comme covariable ne contribue pas au modèle de manière notable.

Dans la plupart des cas, lorsqu'un mot est employé, les erreurs sont plus nombreuses dans la condition ChatGPT que dans la condition Quizlet. Nous pouvons donc conclure qu'au moment de

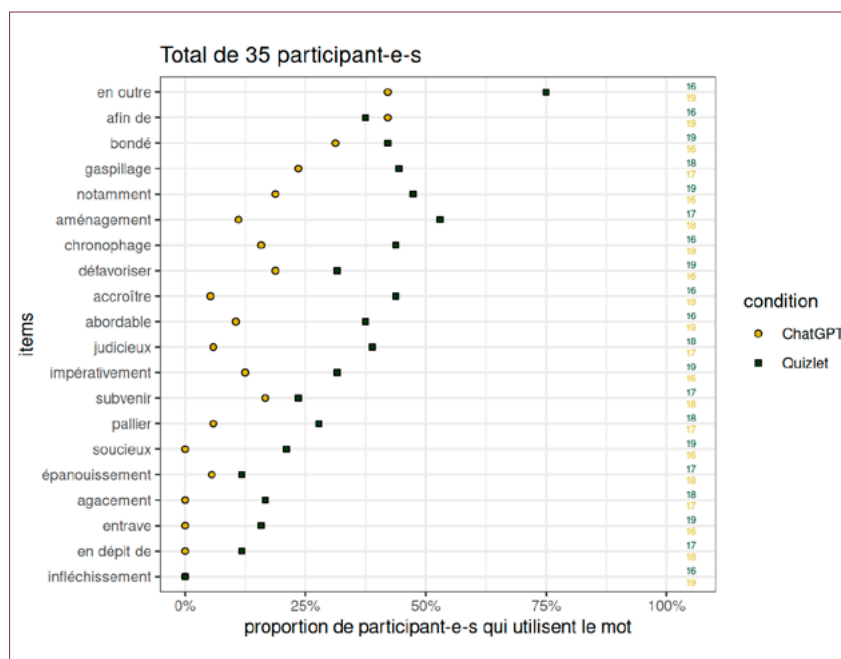


Figure 1

Proportion d'utilisation par mot et par condition (groupe contrôle pas compris)

la rédaction d'un texte, il est plus probable de se rappeler d'un mot appris avec Quizlet et que son emploi sera également plus correct.

Mots retenus après une semaine

Lors de la passation du post-test, après une semaine, chaque participant-e a dû répondre à des questions posées sur chaque mot étudié. En nous basant sur l'idée de la profondeur des connaissances lexicales (Wesche & Paribakht 1996), chaque mot (p. ex. <pallier>) a été présenté aux participant-e-s qui devaient indiquer / formuler une réponse selon les options suivantes :

- Je n'ai encore jamais vu ce mot.
- J'ai déjà rencontré ce mot mais je ne sais pas / plus ce qu'il signifie.
- J'ai déjà rencontré ce mot et il signifie ... (entrez un synonyme, une traduction, une brève explication dans le champ qui s'affichera)

Je sais comment utiliser ce mot correctement dans une phrase en français.

Lorsque la troisième option était choisie, un champ additionnel à compléter s'affichait. Ensuite, le dernier champ s'affichait afin que les participant-e-s puissent entrer une phrase (colonne 'bonus' de la Figure 2). Les réponses entrées dans ces deux derniers champs ont été évaluées manuellement par notre équipe.

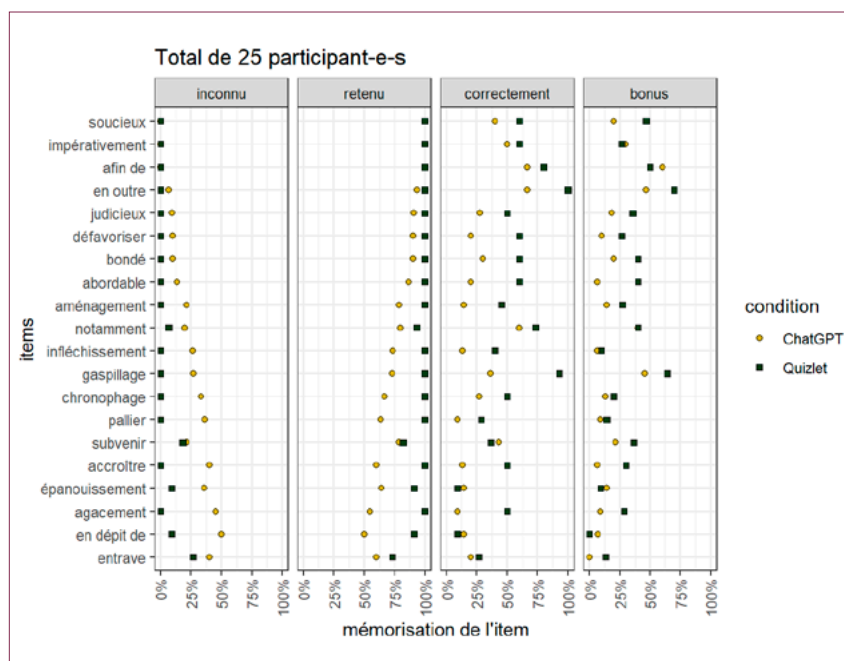


Figure 2
Résultats du post-test après une semaine en fonction des conditions.
Souvenir du sens des mots et utilisation correcte du mot dans une phrase (colonne 'bonus').

Sur la base de ces données, nous pouvons comparer le niveau de rétention en mémoire des items dans les deux conditions. En utilisant Quizlet, les participant-e-s se rappellent en moyenne d'environ 50 % des mots, pour seulement 30 % dans la condition ChatGPT. La figure 2 distingue plusieurs niveaux de connaissances lexicales, selon que les participant-e-s se rappellent du mot (mais peut-être seulement de manière incorrecte); se rappellent du mot de manière correcte (les participant-e-s ont été capables de donner un synonyme, une traduction, etc.); savent utiliser le mot correctement en contexte.

Les résultats du groupe ayant utilisé les deux méthodes pour les mots ne sont pas donnés ici. Sans surprise, le taux de rétention est bien plus élevé dans cette condition par rapport aux deux conditions comparées dans la Figure 2.

À nouveau, la modélisation statistique, que ce soit pour le niveau de mémorisation le plus bas (les participant-e-s ne se rappellent pas vs. se rappellent du mot) ou pour un niveau plus élevé (aucun souvenir ou souvenir incorrect vs. souvenir correct du mot) tel que décrit ci-dessus, indique des résultats significativement meilleurs dans la condition Quizlet ($p < 0.05$, augmentation des chances logarithmiques de 4.18 pour le premier

niveau, i.e. aucun souvenir vs. souvenir; $p < 0.001$, augmentation des chances logarithmiques de 1.43 pour le deuxième niveau, i.e. aucun souvenir ou souvenir incorrect vs. souvenir correct du mot).

Discussion

Cette exploration de l'utilisation d'un agent conversationnel à des fins pédagogiques est une première tentative qui présente des limites importantes.

Lors de l'observation du travail individuel ainsi que lors du débriefing avec les participant-e-s, nous avons constaté que les apprenant-e-s qui travaillaient avec ChatGPT étaient confronté-e-s à une activité complexe nécessitant le traitement d'une grande quantité d'informations. Les apprenant-e-s qui disposaient d'un niveau de lecture intermédiaire en français passaient ainsi beaucoup de temps à lire et à essayer de comprendre l'output de l'outil. Notre paradigme expérimental qui prévoyait d'employer un nombre de mots pour une passation estimée à une heure, a probablement surestimé la quantité d'informations que ces apprenant-e-s étaient en mesure de traiter en langue étrangère. La plus grande richesse informationnelle qui pourrait potentiellement améliorer l'apprentissage selon les théories de la charge d'implication ('involvement load') et de la profondeur de traitement a été en quelque sorte noyée dans la surcharge d'informations et dans la complexité de la condition ChatGPT.

Les meilleurs résultats observés dans la condition Quizlet (bien moins riche en informations) confirme le principe sous-jacent à certaines théories de l'apprentissage lexical qui postulent qu'une possible compétition entre différentes dimensions informationnelles péjorerait l'apprentissage (Barcroft 2015). L'effet potentiellement délétère d'un surplus d'informations expliquerait aussi d'autres observations que nous avons réalisées dans une étude récente à propos de l'emploi contre-productif de fiches répondant à une modalité plurilingue (plutôt que bilingue) pour l'apprentissage de nouveaux mots (Deuber & Berthele 2024). Dans la condition ChatGPT, un éventuel effet positif dû à l'output idiomatique en langue cible illustrant l'emploi des mots en contexte serait donc contrecarré par une surcharge informationnelle.

« L'emploi d'agents conversationnels dans une perspective de développement et d'approfondissement des connaissances lexicales présente un potentiel didactique certain. »

Nous restons néanmoins convaincus que l'emploi d'agents conversationnels dans une perspective de développement et d'approfondissement des connaissances lexicales présente un potentiel didactique certain, ce qu'ont confirmé certain-e-s participant-e-s à l'étude. Le protocole expérimental employé lors de cette dernière était cependant trop rigide et la quantité d'information à traiter dans un laps de temps imposé et relativement court était trop importante.

Plusieurs pistes pourraient être poursuivies pour de futures explorations. Afin d'augmenter la validité écologique de la tâche, une plus grande liberté de choix pourrait être donnée aux apprenant-e-s en ce qui concerne les prompts qu'ils entrent dans l'agent conversationnel. Ceci rendra bien évidemment l'analyse plus compliquée puisque le comportement des apprenant-e-s ne sera très certainement pas uniforme. Cependant, des enquêtes plus qualitatives et observationnelles sur l'utilisation d'agents conversationnels pourraient mettre en exergue l'emploi par les apprenant-e-s de stratégies interactives fructueuses et innovantes – dont on pourrait, dans un deuxième temps, à nouveau tester les apports en suivant un paradigme expérimental plus contrôlé.

Les agents conversationnels pourraient bien évidemment être mis à profit dans

une perspective acquisitionnelle au travers d'autres types d'activités. L'apprenant-e peut, par exemple, se servir de l'output de ChatGPT pour le comparer à ses propres productions, s'en servir à des fins de 'correction' ou en vue de la réécriture d'un précédent brouillon. Par une telle procédure, l'apprenant-e renouvellerait ainsi la comparaison entre la production rédigée de manière autonome et la version corrigée par l'IA.

Ces scénarios (et bien d'autres) partent évidemment du principe que les apprenant-e-s ont véritablement envie d'apprendre à mieux formuler et pas simplement à produire rapidement du texte dans la langue cible. Dans ce dernier cas, la disponibilité de ChatGPT, DeepL et autre GoogleTranslate rendrait l'apprentissage de compétences rédactionnelles dans une langue étrangère superflu.

Références bibliographiques

Barcroft, J. (2015). Specificity in Type of Processing and Learning: The TOPRA Model. In *Lexical Input Processing and Vocabulary Learning* (pp. 57–68). John Benjamins Publishing Company. <https://benjamins.com/catalog/Illt.43.05ch5>

Craik, F., & Tulving, E. (1975). Depth of Processing and the Retention of Words in Episodic Memory. *Journal of Experimental Psychology: General*, 104, 268–294. <https://doi.org/10.1037/0096-3445.104.3.268>

Deuber, F., & Berthele, R. (2024). Vocabulary card learning: Testing multilingual translations and increased semantic processing. *Language Teaching Research*, 13621688241254311. <https://doi.org/10.1177/13621688241254311>

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How Effective Are Intentional Vocabulary-Learning Activities? A Meta-Analysis. *The Modern Language Journal*, 104(4), 715–738. <https://doi.org/10.1111/modl.12671>

Wesche, M., & Paribakht, S. T. (1996). Assessing Second Language Vocabulary Knowledge: Depth Versus Breadth. *Canadian Modern Language Review* 53, 13–40.

Yanagisawa, A., & Webb, S. (2022). Involvement load hypothesis plus: Creating an improved predictive model of incidental vocabulary learning. *Studies in Second Language Acquisition*, 44(5), 1279–1308. <https://doi.org/10.1017/S0272263121000577>

Remerciements : cette étude a été conçue et menée lors d'un séminaire que j'offre dans le cadre du Master Didactique des Langues Étrangères de l'Université de Fribourg. Ce projet collaboratif a été mené à bien par notre groupe au cours du semestre de printemps 2023. Je tiens à remercier Benjamin Abt, Laura Berger, Jordy Lorenzo, Livia Rose, Eva Roski, Noémie Schalbetter et Lisa Schneider pour leur collaboration. Merci aux collègues de différentes haute écoles (Oriana Sassone PHZH, Elisabeth Reiser Bello-Zago UNIFR, Simone Ganguillet PHBE) et leurs étudiant-e-s d'avoir participé à la récolte de données. Un grand merci aussi à Nicolas Félix pour sa relecture critique d'une première version de cet article.

UN CLASSIFICATEUR DE CICÉRONIANITÉ POUR TRAVAILLER SUR LES IDIOMES LATINS

In diesem Artikel schlagen wir vor, einen Cicero-Klassifikator zu verwenden, d. h. eine künstliche Intelligenz (KI), die darauf trainiert ist, einen von Cicero verfassten Text zu erkennen, um Lernenden dabei zu helfen, lateinische Idiome zu erfassen und sich anzueignen um sich dadurch besser ausdrücken zu können. Das System erlaubt es auch, der Übung des lateinischen Aufsatzes eine spielerische Facette hinzuzufügen und den Schülerinnen und Schülern einige Besonderheiten der lateinischen Sprache bewusst zu machen, sowie die Relevanz ihrer Kenntnis. Da der Klassifikator per se fehlbar ist, bieten die Aktivitäten außerdem die Gelegenheit, mit den Schülerinnen und Schülern über Autoritätsfragen im Zusammenhang mit den Produktionen der KI zu diskutieren. Der Klassifikator ist kostenlos auf Deutsch, Französisch und Englisch unter <https://latin-ia.hepl.ch/classifier> verfügbar.

● Damien Cavaleri
| UER-MI, Haute École
Pédagogique Vaud
Micha Hersch
| UER-MI, Haute École
Pédagogique Vaud
Antje Kolde
| UER-MI, Haute École
Pédagogique Vaud

En classe de latin, le travail en communication, notamment en expression, est souvent laissé de côté. Cela est dû en partie aux habitudes d'enseignement et au manque de locuteur·e·s natif·ve·s ou de niveau C2. Il en résulte que l'activité d'expression présente des difficultés supplémentaires en latin par rapport à une langue moderne. En effet, non seulement les élèves et les enseignant·e·s sont confronté·e·s aux difficultés inhérentes à l'expression, respectivement à la correction, dans une langue tierce, mais surtout il peut exister une certaine appréhension vis-à-vis d'une activité peu habituelle. Pourtant ce travail d'expression relève d'une importance cruciale dans le processus d'acquisition de la langue (Canale & Swain, 1980 ; Swain, 1985 ; Nation, 2007). En ce sens, il nous a semblé utile de développer des moyens pour faciliter ce genre d'activité en classe. Dans ce contexte, un « classificateur de cicéronianité » — l'IA dont il est question dans cet article — a été développé dans une double optique : d'une part, faciliter l'expression en latin, c'est-à-dire, réduire la difficulté perçue de

l'activité et, d'autre part, favoriser la prise de contrôle par les élèves de la langue et de ses idiomes.

Le classificateur de cicéronianité est un petit modèle de classification textuelle. Il tourne dans un navigateur web et consomme donc peu de ressources en comparaison avec les IA conversationnelles actuellement en vogue. Il a été entraîné sur un corpus de textes latins à différencier des séquences écrites par Cicéron de celles d'autres auteurs latins. Pour l'utiliser, il suffit de lui fournir une phrase ou, mieux, un paragraphe latin. En retour, l'utilisateur reçoit un score en pourcentage. Ce « score de cicéronianité » représente, selon le modèle, une mesure de similarité de la séquence latine fournie au style de Cicéron. Des séquences trop courtes ne fournissant pas assez d'information à analyser, le modèle indique également un degré de fiabilité du score produit (fiable, moyennement fiable, non fiable). Le présent article décrit dans sa première partie comment cet outil a été utilisé dans diverses classes de latin en

Grâce à l'usage pratique de tournures et de constructions idiomatiques (par exemple, l'ablatif absolu ou la construction infinitive), les élèves non seulement s'emparent de la langue latine, mais surtout se construisent une sorte de boîte à outils d'expressions latines.

Suisse romande pour travailler les notions d'idiomes en latin, en s'appuyant sur ceux que présentent les textes de Cicéron — les élèves devaient en effet reformuler des phrases dans le but de tromper le classificateur et de se faire passer pour Cicéron. Dans sa seconde partie, il en analyse quelques apports que l'expérience aura mis en évidence.

Le modèle est librement disponible sur le site internet du projet : <https://latin-ia.hepl.ch/classifier>.

À l'usage

En classe, les élèves reçoivent donc une phrase ou un paragraphe latin écrit par un auteur contemporain, par ChatGPT ou par un auteur latin et ont la mission d'en faire monter le score de cicéronianité en recourant à des idiomes latins, tels que l'ablatif absolu, ou des spécificités du style de Cicéron, comme l'interpellation des juges. Pour que l'activité se déroule bien, il est important que l'enseignant-e ait au préalable introduit la notion d'idiome et certaines caractéristiques du style de Cicéron. En fonction de la phrase initiale, la focale du processus de reformulation change.

Par exemple, une phrase de ChatGPT ou d'un-e auteur-e contemporain-e aura tendance à intégrer des tournures non idiomatiques en latin (mais idiomatiques dans la langue-culture principale de celui-elle qui l'a formulée). Dans ce cas, la tâche des élèves sera essentiellement de « latiniser » la phrase en y intégrant des idiomes latins. À l'inverse, une phrase d'un auteur latin culturellement et chronologiquement plus proche de Cicéron permettra de se concentrer sur l'usage de figures de rhétorique ou d'autres effets de style cicéroniens.

Du point de vue des élèves, la tâche consiste à « tromper » le classificateur, c'est-à-dire à intégrer des idiomes latins, souvent non transposables en français, et à faire monter le score de cicéronianité. Lors de l'activité, chaque élève a accès au classificateur et peut obtenir sans délai un score pour sa composition. Cela lui permet de tenter différentes combinaisons de mots, de varier le lexique, d'insérer des tournures de phrases « bizarres », c'est-à-dire non transposables en français, vues en lecture, d'intégrer des figures de style et, ainsi, par tâtonnement, de progressivement faire monter le score de cicéronianité. Les figures 1 et 2 présentent des exemples de reformulation.



Damien Cavaleri,
UER-MI, Haute École
Pédagogique Vaud,
Lausanne, Suisse

Damien Cavaleri est chargé de recherche à la Haute École Pédagogique Vaud, spécialisé dans l'application de l'intelligence artificielle à la didactique du latin.



Micha Hersch,
UER-MI, Haute École
Pédagogique Vaud,
Lausanne, Suisse

Micha Hersch est professeur associé en didactique de l'informatique à la Haute École Pédagogique Vaud. Il s'intéresse entre autres à l'introduction de l'intelligence artificielle dans les écoles.



Antje Kolde,
UER-LC, Haute École
Pédagogique Vaud,
Lausanne, Suisse

Antje Kolde est professeure ordinaire en didactique des langues anciennes à la Haute École Pédagogique Vaud

Figure 1

Exemple de variation sur le lexique;
utilisation de l'enclitique "-que" au
lieu du "et" transposable en français.

Texte latin	Score de cicéronianité
Scipio et Hannibal imperatores bello clari, pace invisi fuerunt.	39.5%
Scipio Hannibal que imperatores bello clari, pace invisi fuerunt.	50.5%
<i>Scipion et Hannibal, généraux célèbres en temps de guerre, furent détestés en temps de paix.</i>	

Figure 2

Exemple de variation sur la syntaxe;
utilisation de la construction
infinitive au lieu de la structure
transposée de l'anglais.

Texte latin	Score de cicéronianité
Nicolaus Copernicus fuit astronomus et mathematicus Polonus. Qui primus argumentum proposuit de mundo heliocentrico, id est in quo Sol est centrum sistematis solaris. (ChatGPT)	16.7%
[...] Qui primus argumentum proposuit esse Solem in medio mundo situm.	58.9%
<i>Nicolas Copernic fut un astronome et mathématicien polonais. Le premier il proposa la théorie de monde héliocentrique, c'est-à-dire dans lequel le soleil est au centre du système solaire.</i>	

Au fil des exercices, grâce à l'usage pratique de tournures et de constructions idiomatiques (par exemple, l'ablatif absolu ou la construction infinitive), les élèves non seulement s'emparent de la langue latine, mais surtout se construisent une sorte de boîte à outils d'expressions latines. Ces éléments contribuent ensuite à rendre plus facile le travail de lecture et de traduction et aident à donner du sens aux structures grammaticales étudiées en classe.

Apports

Des séquences pédagogiques impliquant ce dispositif ont été testées dans des classes de différents niveaux de latin en Suisse Romande et les retours des élèves et des enseignant-e-s ont été recueillis. Il apparaît que la rétroaction immédiate et quantitative fournie par l'IA a un important effet motivationnel sur les élèves qui s'engagent davantage dans l'activité de production langagière. Le fait de devoir activement choisir entre différentes tournures de phrases afin de tromper la machine engage les élèves à dépasser les notions d'écrire juste ou faux, à investir les notions d'idiome et de style et d'en comprendre les ressorts. L'usage de ce

dispositif, notamment parce qu'il rompt avec les activités classiques des cours de latin, permet également de rendre saillantes ces notions qui pourront ensuite être mobilisées dans la suite du cours, pour une compréhension plus fine des textes étudiés.

En plus des apprentissages propres au latin, ce dispositif permet d'aborder avec les élèves des questions très actuelles autour de l'IA. Par exemple, le fait d'indiquer un niveau de fiabilité du résultat de l'IA aide à mettre en évidence cet aspect-là et de questionner la fiabilité d'autres IA auxquelles sont confronté-e-s les élèves. On peut aussi questionner la légitimité de cet outil d'IA, le classificateur, à servir de référence dans une langue ne comportant plus de locuteur-e natif-ve. Enfin, avec des élèves plus âgé-e-s, on peut discuter de la performativité des IA, c'est-à-dire leur propension à définir une caractéristique sous couvert de la mesurer. Ainsi, bien que la pertinence de la notion de «cicéronianité» quantifiée d'une séquence latine puisse être débattue, le fait de mettre à disposition une IA qui l'évalue est une manière d'imposer l'existence de cette notion et de contrôler, dans une certaine mesure, ce qu'elle recouvre. Ces mêmes logiques sont

En plus des apprentissages propres au latin, ce dispositif permet d'aborder avec les élèves des questions très actuelles autour de l'IA.

à l'œuvre dans divers domaines d'utilisation de l'IA, par exemple lorsque la vidéo-surveillance algorithmique définit un comportement « normal » ou avec des IA évaluant le risque de récidive dans un contexte judiciaire.

Ces éléments illustrent comment le détour par l'étude d'une langue ancienne permet d'enrichir la réflexion des élèves sur des enjeux contemporains.

Bibliographie

Canale, M., & Swain, M. (1980). Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1-47.

Nation, I. S. P. (2007). The Four Strands, *Innovation in Language Learning and Teaching*, 1 (1), 2-13.

Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In S. Gass, & C. Madden (Eds.), *Input in Second Language Acquisition*. Rowley, MA: Newbury House, pp. 235-253.

SEARCH 'A2, LIEBESLEBEN, 500 WORDS'. TEACHERS' USE OF AI-BASED TOOLS FOR THE GFL READING CLASS

Gegenstand des Beitrags ist der Leseunterricht in der Fremdsprache und die Rolle von KI-basierten Werkzeugen für Lehrpersonen. Der Fokus liegt auf einer Textsuche für authentische, niveaugerechte Lesetexte, besprochen werden zudem weitere KI-basierte Werkzeuge wie Textanalyse und Bilderzeugung. Diskutiert wird, wie Lehrpersonen diese Tools erfolgreich für die Vermittlung von Lesekompetenz in Deutsch als Fremdsprache (DaF) einsetzen können. Die im Projekt „LATILL – Level-Adequate Texts in Language Learning“ (2022-2025) entwickelte Textsuche ist auf einer Online-Plattform für Lehrpersonen verfügbar. Ein didaktisches Beispiel veranschaulicht den Einsatz der genannten Tools zur Vorbereitung von handlungsorientiertem, binnendifferenziertem DaF-Leseunterricht. Ziel des Beitrags ist, Unterrichtenden Anregungen dafür zu geben, wie KI-basierte Werkzeuge zur Planung und Vorbereitung von fremdsprachlichem Leseunterricht eingesetzt werden können.

● Anne Raveling
| Universität Wien



Anne Raveling is a researcher at the Department of German Studies at the University of Vienna. Her work

focuses on foreign language didactics, especially reading didactics, and teacher education, with a further research focus on language learning through art.

Teaching Reading in German as a Foreign Language (GFL)

Without question, the reading literacy of learners, and therefore the teaching of reading competencies, is important not only in the first language but also in a foreign language. In this article, thoughts about the teaching of reading in GFL are presented, discussing the action-oriented teaching approach and other didactic aspects. Subsequently, the LATILL (Level-Adequate Texts in Language Learning) project and its main ideas are introduced, before presenting an example of how to prepare an authentic reading text with the use of AI-based tools.

A teacher can influence learners' successful acquisition of reading literacy by offering a foreign language reading class that is action-oriented and internally differentiated. The action-oriented approach, as it is described in the CEFR (Trim et al., 2001) and CEFR Companion Volume (North et al., 2018), is associated with a focus on learners as users of

language and as social agents. This results in learner-oriented, task-based language teaching and learning, enabling learners to use the foreign language for their purposes within the "defined mission" (North, 2023, p. 8).

Equally important is the teaching principle of internal differentiation: The term covers all didactical measures that allow teachers to deal with the individual differences between learners (Buck & Hensch, 2023). As school classes with heterogeneous groups of learners are the norm, it is important to offer every learner an opportunity to acquire reading literacy. Because reading in a foreign language can be frustrating for learners, especially when a text is too difficult and/or contains too many unknown words, the focus is on presenting learning resources and scaffolds on different levels. One solution is to provide level-adequate authentic reading texts. How the LATILL platform enables this will be described later in this article. Foreign language (FL) learners can profit from reading authentic texts that are close to their language level.

One advantage of authentic texts is that they take up current discourses or refer to up-to-date topics. Therefore, it might be easier to motivate learners to read if they know that the text is authentic, perhaps written by a peer in a country where the foreign language is officially used. A further advantage of authentic texts is the richness of context and language. Authentic texts show a greater variation of words and grammatical structures than didactical, artificial texts.

In addition to the action-oriented teaching approach for a GFL reading class, that employs internal differentiation and offers level-adequate texts, learners can be supported in basic reading skills by enhancing their reading fluency. Especially at lower language levels, this can be achieved by fostering word recognition. The aim is to transfer (new) frequent words into the word memory in such a way that they can be recognized 'at a glance' without having to decipher them letter by letter (Rosebrock et al., 2011). Therefore, practicing the accuracy of the necessary decoding skills should also be part of reading instruction.

Another aspect to include in GFL reading instruction concerns higher-order skills: the knowledge and use of reading strategies for dealing with texts in a foreign language. This may include reading strategies such as thinking about the topic before reading (elaborating) or guessing unknown words from context (inferring). If such strategies are lacking when learners are confronted with a new reading text, the reading goal may not be achieved. That is why teachers explain and model reading strategies and provide learners with the opportunity to practice, evaluate, and improve their strategy use (Chamot Uhl & O'Malley, 1994).

In an ideal GFL reading class, learners are suitably equipped with the necessary reading strategies for dealing with an authentic, up-to-date reading text that meets the individual learner's language level. Reading the text is of relevance to the learners and motivates them to act in the FL. A GFL reading class that follows these didactic principles contributes to learners' experience of reading as a worthwhile process that helps develop their reading literacy in a foreign language. The AI-based tools of LATILL take into account all these aspects and didactic principles.

The LATILL project

The LATILL project aims at supporting teachers in GFL reading instruction, but also in their ability to identify and modify digital resources. Therefore, in the LATILL project, a level-adequate text search for authentic texts in German was created and combined with further AI-based tools. Together with GFL teachers, we tried to find ways how to best use these tools for preparing foreign language reading classes. During the period of a whole school year, thirteen teachers in Ukraine and Spain tested an online platform that was adapted according to their needs and ideas². The main tool is a text search for authentic, level-adequate reading texts accompanied by further tools such as text analysis, image generation, and automatic translation. Furthermore, didactic materials are provided. The LATILL platform has been developed in a joint interdisciplinary effort. The partners are e-learning experts who developed the platform and other tools (GRIAL, University of Salamanca), computational linguists (University of Tübingen) who created the text corpus and developed a CEFR-level classification supported by partners from the ÖSD (Vienna), and GFL (University of Czernivtsi, University of Vienna) who implemented teacher training and tested the platform³.

Within the project, the emphasis of the teacher training was on finding new and useful ways of using the platform. This is why the LATILL reading didactics focused on lesson preparations. On the one hand, teachers are experiencing rising expectations concerning their pedagogical, didactical, as well as digital knowledge. On the other hand, these expectations are accompanied by limited time to prepare their teaching due to challenging curricula and large classes and a lot of bureaucracy. Time is therefore a crucial resource in planning and preparing for the upcoming class, an important part of a teacher's work. For this reason, lesson preparation offers a good opportunity to explore how AI-based tools can help teachers to work time-efficiently and yet in the learner's interest to provide better GFL classes.

For the project's teacher training, reading instruction was developed which was conceptualized for and applied in a webinar series (Raveling & Schramm, 2023).

- 1 "LATILL – Level-Adequate Texts in Language Learning", Erasmus+ KA2-Project 2021-1-AT01-KA220-SCH-000029604 <https://erasmus-plus.ec.europa.eu/projects/search/details/2021-1-AT01-KA220-SCH-000029604>, project website <https://latill.eu/>. Co-funded by the European Union. Views and opinions expressed are those of the author and do not necessarily reflect those of the European Union, which cannot be held responsible for them.
- 2 A big thank you to all Ukrainian and Spanish teachers who participated in our project during these difficult times.
- 3 The project members are: Universität Wien (AT): Univ.-Prof. Dr. Karen Schramm, Anne Raveling, Daniele Polizio, Associate Researcher: Dr. Martina Kienberger (Universidad de Granada); Universidad de Salamanca/GRIAL (ESP): Dr. Roberto Therón, Dr. Alicia García-Holgado, Dr. Andrea Vázquez-Ingelmo, Nastaran Shoeibi, Ass. R.: Dr. María Mar Soliño Pazó; Eberhard-Karls-Universität Tübingen (DE): Univ.-Prof. Dr. Detmar Meurers, Denise Loefflad, Benedikt Beuttlér; Chernivtsi National University Yuriy Fedkovych (UKR): Dr. Tetiana Koropatnitska, Dr. Olena Mudra, Prof. Dr. Bohdana Labinska (former project member); Verein Österreichisches Sprachdiplom Deutsch (AT): Dr. Manuela Glaboniat, Dr. Christina Glinik, Melanie Weiss; for more details and further publications see project website <https://latill.eu>.

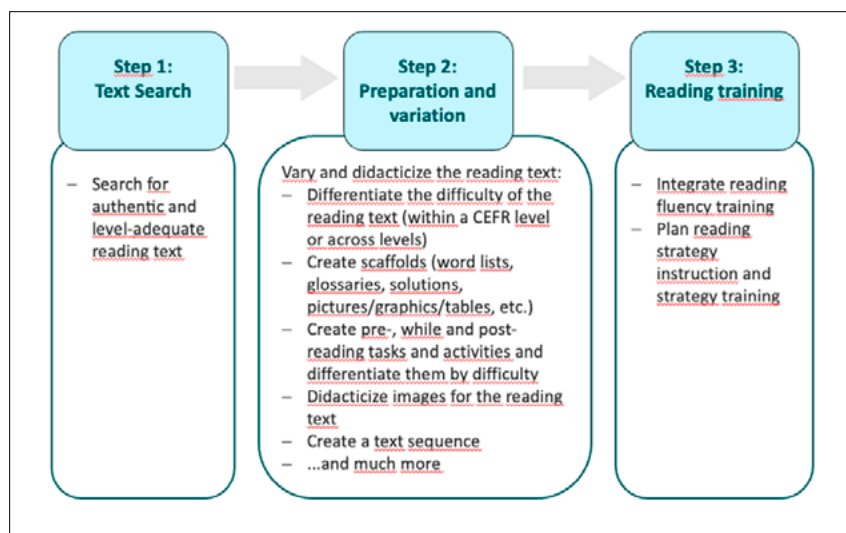


Figure 1

LATILL reading didactics: Lesson planning and preparation step-by-step

Themenverteilung

Altag, Familie und Gesellschaft: 2

Filter

Liebesleben

Search for:

☒ All words ☐ Any word ☐ Exact phrase

Select Publisher

⌕ A2 |

Select Thema

Anzahl der Wörter

Min: 0 - Max: 500

Figure 2

LATILL text search in corpus with keyword, CEFR-level and number of words

Heast wieso hast du keinen Freund

Text info | Text level | Collections

Select a CEFR level to highlight its corresponding words in the text.

A1 A2 B1 B2 C1 C2

Die Welt hört kurz auf sich zu drehen, wenn mir mal wieder diese Frage gestellt wird. Seid ihr alle deppert? Am Sonntag verkündet mir mein Vater: „Hi, Mama und ich fragen uns, wieso du eigentlich keinen Freund findest...“ Ist das dein Ernst, Papa? Haben meine Eltern keine anderen Sorgen, als sich zu fragen, wieso ihre 21-jährige Tochter nicht vergeben ist? Am Montag besuche ich ein befreundetes, serbisches Ehepaar. „Jetzt, wo du die Uni abgeschlossen hast, wird es Zeit, an Familie zu denken. Du hast einen Freund, oder?“, werde ich von ihnen begutachtet. Naja, sie haben ja eh zwei unverheiratete Söhne. Ich lasse bei zwei guten Freundinnen meinen Groll über diese nervige Frage aus, und sie trösten mich: „Keine Sorge, es wird schon der Richtige kommen.“ Fun-Fact: Die eine ist glücklich verheiratet und die andere frisch verliebt. Wer sagt, dass ich mir Sorgen mache? Ganz ehrlich: Es ist niemand mehr fix zam! Bist doch eh hübsch ich bin nicht unglücklich, weil ich keinen Freund habe. Ich mache all die Sachen, die ich will, und genieße mein Leben. Ahja: Und so viel beschäftigte ich mich gar nicht damit, wieso, weshalb, warum ich Single bin. Ist halt so, passt halt grade nicht. Aber wenn alle zwei Tage irgendein verliebter Depp mich fragt wieso, weshalb, warum, werde ich grantig. Die Standard-Meldung ist dann: „Du bist doch hübsch, wird wohl nicht so schwer sein, wen zu finden.“ Irgendwer ist besser als niemand Diese sinnbefreite Aussage, sich irgendwen zu suchen, ist fast so dumm wie die Frage nach meinem Beziehungsstatus. Wollen Menschen, die mich gern haben, wirklich, dass ich mir einfach irgendwen „zulege“? Wenn ja, dann vielen Dank. Die glücklich Verliebten verstehen nicht, dass auf den Straßen Wiens nicht nur heiratswillige, gutaussiehende, smarte Männer rumrennen. Glaubt mir, ich weiß wovon ich spreche. Besser allein, als einen, der mir erklärt, er würde jedes Wochenende besoffen unter einer Parkbank aufwachen. Oder der Typ, der mir nach unserem vierten Date von seinem Kind und seiner Freundin erzählt – er beteuert immer wieder, wie gern er mich hat. Klischee, ich weiß, aber so schaut die Singlerealität aus, meine lieben, besorgten Freunde! Jeder will dir sagen, was du tun musst, um nicht mehr Single zu sein. Was stimmt nicht mit dir? ...hat mein Bruder mich letztens gefragt, der nur so nebenbei auch Single ist. Also ich glaube eigentlich, dass alles mit mir stimmt. Nur wenn ich die ganze Zeit bedrängt werde, an meinem Liebesleben zu arbeiten, fängt es bei mir auch an zu rattern. Jetzt mal im Ernst: Kann man an seinem Liebesleben arbeiten? Und vor allem: Will ich das? Ich kann natürlich bei Tinder restriktiv nach Kandidaten suchen, jedem Vollidioten, der mich angräbt, eine Chance geben. Oder durch die Telefonliste meiner Verflorenen stöbern und mir überlegen, ob vielleicht jemand von denen doch alle Tassen beieinander hatte. Nein, danke. Da lass ich mir lieber einreden, dass was mit mir nicht stimmt. Es gibt auch Menschen wie mich, die nicht jeden Tag Strategien entwerfen, um ihr Dasein nicht als alte Jungfer zu fristen.

Figure 3

LATILL text search offering words marked according to CEFR level classification

The goal of the teacher training was to discuss the topic of GFL reading didactics under certain aspects and to intensify the reflective use of the LATILL platform, especially the level-adequate text search, and further AI-based tools (Raveling & Koropatnitska, 2024). The teachers met during five webinars and additionally during informal group meetings within the international community of practice. This is where they shared their ideas and challenges in using the LATILL platform as well as prepared and used authentic reading texts from the LATILL text corpus in their GFL classes. According to the LATILL reading instructional design, the suggested steps for planning and preparing a GFL reading class with the LATILL platform are the following: [Figure 1].

The first step is to select an authentic, level-adequate reading text. For this purpose, a text search can be conducted looking for a CEFR level and/or key-word(s). The LATILL text search in the corpus (“discover texts”) allows not only to choose a CEFR level from “A1” to “B2 and above” but also between one or

all keywords or an exact phrase to be present in the text. It is also possible to apply additional filters concerning text length or filters for grammatical features, e.g. average sentence length or dependent clauses. For even more authentic and up-to-date reading texts, the LATILL platform additionally offers a “web search” that prioritizes didactically relevant sources and shows ten texts per search that are classified in one of the CEFR levels from “A1” to “B2 and above”. After the search, a CEFR level can be chosen so that only level-adequate texts are displayed. It is important to note that texts are not inherently easy or difficult, but the level of difficulty depends on the language learner’s knowledge and skills combined with specific features of a text, e.g., vocabulary used, as well as features external to the text, e.g., reading goal, learning task, time for reading, and many others. Nevertheless, the ‘level’ of a learning resource is a crucial factor for teachers, often determining whether the learning material is chosen and widely requested.

Tools for teachers

In the following, the planning and preparation of a GFL reading class using AI-based tools will be described. This is of course only one of many possibilities how AI-based tools can be used for preparing an authentic reading text. This example is generalizable to foreign language classes other than GFL, as the didactic principles apply to any instructed learning setting in foreign or second language teaching.

As planning should always start with the learners, a hypothetical but typical GFL class is taken as a starting point, e.g. a third year of secondary school, where the learners are around fifteen years old, and have been learning German as a second foreign language after English for two years at this point. The learners’ language level is expected to reach B1 by the end of the school year, but their language competencies and reading skills vary from the beginnings of A2 to B1.

The appropriate CEFR ‘can do’ descriptors from “overall reading competence” and “reading for information/argument”, which set the broad learning objectives, read in slightly modified form as follows:

Overall reading competence

- B1 Can read an *online column* on an interest-related subject with a satisfactory level of comprehension.
- A2 Can understand a short, simplified *online column* on a familiar matter that consists of frequent everyday language.

Reading for information/argument

- B1 Can recognize significant points and understand most information in an *online column* on a familiar subject of interest.
- A2 Can understand the main points and partly identify specific information in a short *online column* dealing with an everyday subject

(cf. North et al., 2018, pp. 60–63, italics added by the author)

Considering the topic of this exemplary reading class, the idea is that at this age, relationships are one of the interests of learners. So “love life” is in the top ten of their GFL topic wish list. Thus, a search is conducted on the LATILL platform for an authentic, level-adequate reading text with the keyword “Liebesleben”, setting the filter “A2” for CEFR level, and the filter “500” for the number of words. [Figure 2]

The corpus search offers five texts. To get a broader choice of texts, the CEFR levels “B1” and “B2 and above” are added as filters. One text seems to particularly suit the topic and supports the broad reading goals, however, it is too long. Since the displayed text is marked with a “cc by” license next to the named and linked source, it can be adapted for teaching purposes. After the text search, the second step, “Preparation and variation”, begins. As the reading text is too long and the LATILL “text info”-feature shows several words on CEFR level B1 and some on B2, the text needs to be simplified. [Figure 3]

The text is copied and pasted into the free Chat GPT-Model 3 in the German version, using a prompt defining the role and task⁴. It must be noted that the chatbot cannot reliably tell which words in the German language are associated with which CEFR level. Therefore, the tool’s outcome is to be controlled and assessed by the teacher, e.g. through the use of a text analysis tool. Pasting the text into the LATILL analysis tool it confirms that the text is classified on

⁴ Prompt: “Als Linguist vereinfachst du einen Text für den Fremdsprachenunterricht. Die schweren Wörter sollen ersetzt und der Text halb so lang werden.”, Chat GPT 3, free version, July 17 2024

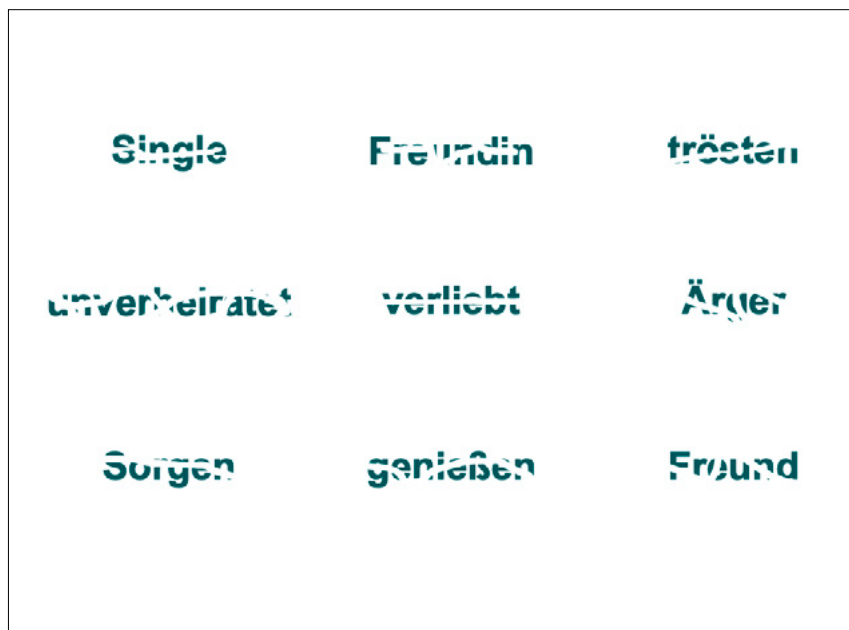


Figure 4
Exercise with "broken words" to train reading fluency

level B1. Several features of the original text are retained, including dialect and colloquial vocabulary, so that the simplified text still offers some advantages of the original authentic text. To offer learners help when needed, the chatbot is asked to create a word list⁵. The produced list is too long and refers to the original text version, so it needs to be manually shortened and finalized. For each of the three paragraphs, the chatbot is asked to create a subtitle⁶. The outcome serves as a helpful inspiration for the creation of the final subtitles that are shorter and more specific. Taking into account the lower level learners within the heterogeneous group, a simpler text version is created by shortening each paragraph manually. Its level A2 is confirmed by the LATILL text analysis tool. In this way, the reading text is differentiated across CEFR levels.

All this of course takes time when done for the first time, especially as it involves two different tools. Nevertheless, with a little practice and applying tested prompts these steps can be quickly completed. As the reading class will be structured in a pre-, while- and post-reading phase with internally differentiated activities, it is now time to create the tasks. The

action-oriented reading goal "Create and submit an online post" can be split into the following can-do statements "Can read and understand an online column on the specific topic of 'Single Life' and can write an online post as response expressing their own opinion" serves as a guideline for task creation, as the task in Figure 5 below shows.

To start the lesson and to recall learners' existing language knowledge, two pictures will be shown, one referring to a "single" person and one to a "married couple". The LATILL image generation generates two images based on the highlighted words in the text. These are added to the text collection and can be downloaded as a Word file.

Finally, the third step of integrating "reading training" is taken: using a standard text program the main words from the word list are designed in such a way that they can be used as an exercise for reading fluency [Figure 4].

According to the learning objectives mentioned above and with the inclusion of all developed learning resources, the task for this scenario could be formulated as show in figure 5.

- 5 Prompt: "Erstelle zum Text eine Wortliste mit den schwierigen Wörtern, aufgeteilt in Verben und Nomen. Schreibe die Nomen mit bestimmten Artikeln auf. Schreibe Verben mit der Perfekt-Form auf. Schreibe darunter Adjektive. Schreibe zu jedem Wort den Beispiel-Satz aus dem Text auf.", Chat GPT 3, free version, July 17 2024
- 6 Prompt: "Füge zu jedem der drei Textabsätze eine kurze, unterhaltsame Überschrift ein.", Chat GPT 3, free version, July 17 2024
- 7 For the formal revision of the text (spelling and grammar check), the free online version of DeepL Write was used. The revised version was partially adopted by the author. <https://www.deepl.com/de/write>, Aug. 16, 2024

Besser Single oder zu zweit?

Was ist deine Meinung zu dem Thema? Am Ende der Woche kannst du selbst einen Online-Post veröffentlichen. Aber halt: Starten wir am Anfang. Lies heute die *Online-Kolumne* einer jungen Journalistin aus Wien (Österreich).

1. Lies die Überschrift. Welche Wörter sind neu und was bedeuten sie? Besprecht alle Wörter.
2. Überlegt zusammen: Was könnte in dem Text stehen?
3. Wähle eine Text-Version (A2 oder B1). Lies den ersten Text-Abschnitt. Was denken die Freunde? Was denkt die Kolumnen-Schreiberin?
4. Lies den zweiten Text-Abschnitt. Warum kann es schwer sein, den „richtigen“ Menschen zu finden?
5. Lies den dritten Text-Abschnitt. Warum schreibt die Journalistin „Nein, danke.“ Was meint sie?
6. Lies den ganzen Text noch einmal. Versuche beim Lesen, unbekannte Wörter mithilfe des Textes zu verstehen: Lies den ganzen Satz oder auch den Satz davor und danach.
7. Besprecht zusammen die wichtigsten Wörter, die du immer noch nicht verstehst.
8. Sprecht über die Frage: „Besser Single oder zu zweit?“ - Was denkt die Autorin? Und was denkst du?

Ausklang: Worttraining. Wer erkennt das Wort zuerst?

Hausaufgabe: Worttraining: Sieh die „kaputten“ Wörter an und versuche sie zu lesen. Mache das morgen wieder und noch ein weiteres Mal vor der nächsten DaF-Stunde.

Figure 5

Learning task for action-oriented and differentiated reading about "Single Life"

This example aimed to show how an originally authentic text, that meets the individual learners' language levels, can offer rich linguistic input and content while at the same time enabling the learners to develop their language skills enjoyably.

Finally, it is important to mention that ethical aspects and certain limitations in the use of AI-based tools have to be considered. They concern data protection, copyrights, and environmental issues. Also, AI-based tools show various biases, for example when creating images or using a certain language register.

Conclusion

The article suggested how an action-oriented and internally differentiated GFL reading class can be created with the support of AI-based tools. The example presented in detail can be transferred to foreign language classes other than GFL. The use of AI-based tools can help teachers adapt their learning materials to the learners and thus enhance adaptive learning, offering useful support to GFL teachers. However, it is necessary to judge any of the AI-based tools' outcomes by its linguistic quality, in addition to

other factors. In the end, it is up to teachers' linguistic knowledge and didactical know-how to choose the level-adequate text that is most suitable for their learners. But once teachers find a good way of using new tools, they can build a repertoire of habits and prompts that save capacities and time or deliver inspiration, while improving the offered learning materials for class. In establishing this way of lesson preparation, a high teaching standard can be maintained while staying flexible in terms of content and learners interests⁷.

References

- Buck, K., & Hensch, J. (2023). Binnendifferenzierung. Herausforderungen und Potentiale. *Binnendifferenzierung. Fremdsprache Deutsch* 69/2023, 3–7.
- Chamot Uhl, A., & O'Malley, M. J. (1994). *The CALLA Handbook. Implementing the Cognitive Academic Language Learning Approach*. Addison Wesley.
- Glaboniat, M., Müller, M., Schmitz, H., Rusch, P., & Wertenschlag, L. (2002). *Profile deutsch. Gemeinsamer europäischer Referenzrahmen. Lernzielbestimmungen, Kannbeschreibungen, Kommunikative Mittel, Niveau A1-A2-B1-B2*. Langenscheidt.
- Grabe, W., & Yamashita, J. (2022). *Reading in a second language: Moving from theory to practice* (2nd ed.). Cambridge University Press.
- Kienberger, M., & Schramm, K. (eds.). (2023). *Lesedidaktik Deutsch als Fremdsprache: Aktuelle Entwicklungen und Ansätze*. Peter Lang. <https://doi.org/10.3726/b21207>
- North, B. (2023). The CEFR companion volume and the action-oriented approach. *Italiano LinguaDue* 2, 1–23. <https://doi.org/10.54103/2037-3597/19566>
- North, B., Goodler, T., & Piccardo, E. (2018). *Common European Framework of Reference for Languages: Learning, teaching, assessment. Companion volume with new descriptors* (Council of Europe). www.coe.int/lang-cefr
- Raveling, A., & Koropatnitska, T. (2024). *DaF-Leseunterricht gemeinsam entwickeln. Webinarserie für eine internationale Lehrer:innen-Gemeinschaft*. <https://latill.eu>
- Raveling, A., & Schramm, K. (2023). *DaF-Leseunterricht gemeinsam entwickeln. Konzeptpapier zur Webinarserie für DaF-Lehrer:innen im Rahmen des LATILL-Projekts (2022-2025)*. <https://latill.eu>
- Rosebrock, C., Nix, D., Rieckmann, C., & Gold, A. (2011). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe*. Kallmeyer Klett.
- Trim, J., Coste, D., & North, B. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Council of Europe, Ed.). www.coe.int/lang-cefr

UN ESTUDIO SOBRE EL USO DE CHATGPT PARA RESOLVER TAREAS ESCRITAS EN LA CLASE DE ELE

This study examines how six B1-level adult learners of Spanish use ChatGPT to complete a monologic task. The results indicate that while students generally follow the structure suggested by GPT, they also adapt its recommendations to suit both the context and their language proficiency. Additionally, individual variation in responses is observed, reflecting students' personal engagement with the communicative situation. The findings suggest that the six students critically utilize AI tools in foreign language learning, highlighting the importance of supporting and encouraging this habit in educational settings.

● Nicola Brocca
| University
of Innsbruck
Rebeca Iniesta
Jiménez
| University
of Innsbruck

1. Introducción

Utilizar modelos grandes de lenguaje (LLMs, *large language models* en inglés) para escribir en una lengua extranjera es una práctica cada vez más habitual en la vida cotidiana (por ejemplo, para corregir un texto breve en una lengua extranjera). Las opiniones de los expertos en la enseñanza de lenguas extranjeras ante esta evolución son diversas: si bien los profesores consideran útiles estas herramientas, les preocupa que los alumnos confíen excesivamente en ellas (Ohashi, 2022; Berthele & Udry, 2023b) y afirman que los aprendientes no pueden desarrollar su competencia comunicativa si delegan todo el trabajo en la máquina y no realizan ellos mismos un esfuerzo cognitivo (Berthele & Udry, 2023a). Otros son de la opinión de que los LLMs no se pueden ignorar (Poole & Polio, 2023) y los profesores deberían integrarlos en el aula de idioma extranjero para ayudar a los alumnos a realizar tareas comunicativas auténticas.

Este estudio de caso investiga cómo estudiantes universitarios de español como lengua extranjera de nivel B1 utilizan un LLM para resolver tareas comunicativas y cómo esto influye en sus producciones escritas desde el punto de vista de la cortesía lingüística. En concreto, se examinan los resultados de 6 estudiantes que han producido dos tareas monológicas (petición y cancelación) utilizando GPT-4 (OpenAI, 2024). Dada su naturaleza exploratoria, la presente investigación se basa en una muestra pequeña de datos y no pretende ofrecer resultados generalizables. Más bien, representa un primer paso hacia una mejor comprensión de la enseñanza del español como segunda lengua utilizando IA, así como una posible fuente de sugerencias para los profesionales.

Los estudios científicos hasta la fecha se centran predominantemente en el inglés, descuidando el uso de los LLMs para la enseñanza de otras lenguas como el español, con pocas excepciones (p. ej., Fredholm, 2019). Algunas propuestas

didácticas para la clase de ELE plantean actividades para utilizar la IA en un entorno de enseñanza, siguiendo un protocolo específico para su uso (Arriagada, 2023; Hagenhof y Leienbach, 2020). En nuestro estudio examinamos específicamente cómo los estudiantes utilizan los LLM de forma espontánea cuando deben completar una tarea.

Por otra parte, este estudio explora el impacto del uso de los LLM sobre la cortesía lingüística. Usaremos el término *cortesía lingüística* para referirnos a la cortesía incorporada en el discurso que puede plasmarse en el uso de términos específicos de tratamiento, honoríficos, expresiones formulaicas convencionales (“gracias”, “perdón”) y varios recursos lingüísticos, como los empleados para mitigar la fuerza ilocutiva directa de una petición (p. ej. el uso de la expresión “por favor” o el uso del condicional) o para reducir los efectos negativos de un rechazo (Félix-Bresdefer, 2024).

Este estudio plantea la siguiente pregunta de investigación: (P) ¿Cómo realizan los estudiantes una tarea monológica desde el punto de vista de la cortesía lingüística cuando utilizan un LLM?

De manera más precisa, esta cuestión se abordará respondiendo a las siguientes preguntas:

P1: ¿Cómo utilizan los estudiantes GPT para producir sus textos?

P2: ¿Qué sugerencias de GPT incorporan los estudiantes en sus textos finales?

2. Revisión de la literatura

En el debate actual sobre la integración de herramientas basadas en inteligencia artificial en la enseñanza de idiomas, a menudo se asume que los estudiantes podrían usar estas herramientas de manera acrítica, lo cual podría tener efectos adversos en el proceso de adquisición del idioma (Berthele & Udry, 2023a). Muchos estudios, por lo tanto, abogan por realizar actividades que fomenten la interacción activa con los resultados generados por la IA —como releer, reinterpretar o modificar las sugerencias de la IA— reinvirtiendo así estas herramientas como recursos de apoyo en el aprendizaje. Por

ejemplo, Hagenhof y Leienbach (2020) proponen un ejercicio práctico en el que los estudiantes comparan traducciones producidas por dos herramientas de traducción basadas en IA y eligen la opción más precisa o adecuada al contexto. Arriagada (2023) destaca además el valor de las aplicaciones de traducción de voz para fomentar la conciencia pragmática intercultural en la clase de ELE, promoviendo el uso de herramientas de IA bajo un protocolo específico. Aunque en estos trabajos las interacciones de los estudiantes con LLMs ocurren en contextos espontáneos y no estructurados, los estudios ofrecen directrices muy específicas para los estudiantes.

Además, algunos estudios sugieren que la producción de formas de cortesía en el lenguaje por parte de los estudiantes no siempre está directamente influenciada por el input o la instrucción explícita. Por ejemplo, Siegal (1996) observó que algunas mujeres occidentales, a pesar de comprender las normas de cortesía en japonés, optaron por no ceñirse completamente a estas convenciones por no ajustarse a su identidad. Asimismo, los hablantes nativos pueden experimentar incomodidad cuando los hablantes no nativos se ajustan demasiado a las normas nativas, percibiéndolo como algo excesivamente formal o incluso pretencioso (Blum-Kulka, 1991: 269).

Basándonos en este contexto, planteamos la hipótesis de que, al redactar mensajes personales, los estudiantes pueden optar activamente por no replicar de forma literal las sugerencias generadas por la IA, lo que indica una interacción más compleja entre la asistencia de la IA y la autonomía del estudiante en el uso del idioma.

3. Metodología

3.1 Participantes

En este estudio participaron 6 personas matriculadas en un curso de español de nivel intermedio (B1) en una universidad en Austria. Todos los participantes eran hablantes nativos de alemán o poseían un dominio del idioma equiparable al de estos. Las edades de los participantes oscilaban entre los 21 y los 31 años. En las semanas previas a la realización de las tareas asignadas, los participantes recibieron clases sobre cómo expresar peticiones corteses utilizando el condicional,



Nicola Brocca is a post-doctoral fellow at the University of Innsbruck. His research interests include pragmatics in the context of speech-acts, TBLT and LLMs. nicola.brocca@uibk.ac.at



Rebeca Iniesta Jiménez is a doctoral candidate at the University of Innsbruck. Her research interests include the use of AI in high school Spanish classes. rebeca.iniesta-jimenez@uibk.ac.at

el uso del subjuntivo para expresar deseos y la redacción de mensajes breves. También recibieron input de saludos y expresiones formulaicas convencionales (“gracias”, “disculpa”) a través del libro de texto y de la práctica en clase.

3.2 Recolección de datos y análisis

A los participantes se les pidió que resolvieran en línea las siguientes tareas (Brocca 2021; Cortés Velásquez, Nuzzo 2022):

- Una cancelación (informal)

Un/a amigo/a te ha invitado a cenar esta noche y has aceptado la invitación. En el último momento, tienes que cancelar la cita porque tienes otro compromiso. Escribe un mensaje por WhatsApp a tu amigo/a comunicando tu ausencia.

- Una petición (formal)

Hoy tienes que enviar tus deberes al/a la profesor/a de español. No has tenido tiempo porque has tenido que prepararte para otro examen. Escribe un e-mail a tu profesor/a de español solicitando más tiempo.

Los participantes realizaron las tareas en un formulario en línea de Google. En las instrucciones de la tarea se indicaba a los estudiantes que utilizaran GPT para prepararla. Para monitorizar este proceso, en el formulario se pedía a los estudiantes que copiaran y pegaran el *prompt* que habían insertado en GPT, así como la respuesta de la máquina. En la página siguiente del cuestionario se pedía a los estudiantes que escribieran sus mensajes finales.

Para el análisis de las estrategias de apropiación de las sugerencias de GPT, nos basamos en un cotejo entre el *prompt*, la respuesta de GPT y el mensaje final. Para el análisis de las tareas, utilizamos un

análisis de los actos de habla (Trosborg, 1995); este análisis permite examinar los movimientos de apoyo, como la justificación de la petición o de la cancelación, los movimientos de reparación y los marcadores de cortesía.

4. Resultados

Para responder a la P1, observaremos cómo formularon los estudiantes el *prompt* en GPT. Los estudiantes utilizaron dos estrategias distintas: i) pedir a GPT que resolviera la tarea desde cero y ii) utilizar GPT para mejorar el borrador del mensaje formulado por ellos mismos. La mayoría (4 de 6) utilizó el enunciado de la tarea como *prompt*, insertándolo en GPT. Solo dos estudiantes utilizaron GPT para mejorar el borrador de su mensaje.

Para responder a la P2, analizaremos ahora cómo los estudiantes pasaron de la propuesta de GPT a la versión final. A la hora de escribir el mensaje final (producción), los alumnos emplearon dos estrategias diferentes: iii) adaptar la respuesta de GPT o iv) ignorar prácticamente la respuesta de GPT. Al escribir el mensaje final, 9 mensajes de GPT fueron adaptados y tres sirvieron únicamente para tomar prestados algunos elementos léxicos y fraseológicos.

Las adaptaciones consistieron básicamente en ajustes menores al acto principal (petición o cancelación) siguiendo la estrategia de GPT, e incluyeron frecuentemente movimientos de apoyo sugeridos por GPT. Los estudiantes tendieron a adaptar elementos léxicos y morfosintácticos para ajustarlos a su nivel de competencia. En el siguiente ejemplo, S2 expresó la petición en indicativo en lugar de en condicional, como sugería GPT, mientras que S4 reemplazó palabras menos comunes por vocabulario más

básico. S4 también añadió un marcador de cortesía (“por favor”).

S2 (GPT): “¿Podría concederme unos días más para entregarlos?”

S2 (Estudiante): “¿Me *puede* dar unos días más para entregarlos?”

S4 (GPT): “Necesito unos días adicionales para completar las tareas adecuadamente.”

S4 (Estudiante): “Necesito dos días *más, por favor*, para completar los deberes.”

En algunos casos, las adaptaciones sirvieron para hacer el texto más adecuado al contexto, como añadir el nombre real del destinatario en el saludo o elegir un saludo más informal para el profesor, como «Hola», en lugar de «Estimado».

En raras ocasiones, algunos movimientos de apoyo sugeridos por GPT fueron incorporados en las versiones finales de los estudiantes, típicamente expresiones de gratitud o de arrepentimiento. Por ejemplo:

S1 (GPT): “¡Me puede hacer un gran favor!”

S4 (GPT): “Agradezco de antemano su comprensión y consideración.”

S2 (GPT): “Lamento mucho avisarte en el último momento.”

Aunque la adaptación es la estrategia más común, dos estudiantes personalizaron más sus textos, tomando solo algunos elementos de GPT. Esta estrategia es evidente comparando la siguiente sugerencia de GPT y el mensaje final del estudiante:

S3 (GPT): “Lo siento mucho, pero me ha surgido un compromiso de última hora y no podré ir a la cena de esta noche. 🙏 ¿Podemos quedar otro día? Mil disculpas por avisar tan tarde.”

Tabla 1

Distribución de estrategias utilizadas para resolver la tarea utilizando GPT

		Preparación	
		Pedir a GPT resolver la tarea desde cero	Introducir un borrador y pedir a GPT que lo corrija y mejore
Producción	Copiar la respuesta de GPT	–	–
	Adaptar la respuesta de GPT	S2, S4, S5 (petición)	S1, S6
	Tomar elementos de la respuesta de GPT	S3, S4 (cancelación)	--

S3 (Estudiante): “Hola, Seb. Me sabe fatal, pero tengo que cancelar la cena de esta noche contigo. ¿Piensas que podemos quedar otro día? Disculpas por avisar más tarde. 🙄”

En el mensaje de S3, se añade un saludo y el nombre del destinatario, mientras que la expresión inicial de disculpa, “lo siento mucho,” se reemplaza por la expresión más idiomática “me sabe fatal,” que intensifica la disculpa. Aunque se omite la explicación sugerida por GPT, la oferta de reprogramar (“¿Podemos quedar otro día?”) se suaviza mediante una estructura incrustada (“¿Piensas que...?”). Se añade también un emoticono de cara de mono, que introduce un tono de humor auto-crítico y agrega un toque personal más allá de la carita triste sugerida por GPT.

La tabla 1 muestra la distribución de los estudiantes según las estrategias que utilizaron.

5. Discusión

El análisis indica que los textos finales de los estudiantes están bastante influenciados por las sugerencias de GPT. Sin embargo, tanto las estrategias de adaptación como de selección indican que los estudiantes interactúan críticamente con las sugerencias de GPT, prefiriendo expresiones que pueden producir de forma independiente dentro de su nivel de competencia lingüística. Además, los estudiantes en ocasiones excluyen expresiones que no se alinean con sus preferencias de cortesía, introducen elementos como el nombre del destinatario o utilizan el saludo “Hola” en lugar del más formal “Estimado”. Aunque la estructura de los mensajes es en su mayoría similar, las adaptaciones tienden a reforzar una mayor informalidad entre los estudiantes y el destinatario. Otro hallazgo es la variabilidad personal en las estrategias utilizadas por los estudiantes: por ejemplo, S3 mostró una creación de tareas divergente y más autónoma, mientras que los otros estudiantes siguieron de manera más sistemática las sugerencias de GPT.

Podemos proponer dos hipótesis para explicar este fenómeno: copiar un texto se percibe como algo éticamente incorrecto, especialmente en un contexto didáctico. Otra hipótesis está relacionada con el tipo de tarea: escribir un mensaje personal.

Los estudiantes creen que aceptar las sugerencias de GPT hace que el mensaje no sea auténtico o creíble, por lo que prefieren ceñirse a la versión que escribirían espontáneamente. Esto podría explicar el rechazo de los estudiantes a las fórmulas de cortesía sugeridas por GPT, ya que, aunque correctas, las consideran alejadas de su forma de expresarse (Siegal, 1996).

Bibliografía

- Arriagada, M.** (2023). Pragmática práctica para una clase intercultural II/II: Límites y posibilidades de los traductores automáticos de voz para el aprendizaje de los actos de habla en español-alemán. *Hispanorama*, 180(2), 8-11.
- Berthele, R. & Udry I.** (2023a). Welche fremdsprachlichen Lernziele der Schulen können mit digitalen Tools aber ohne Lernen erreicht werden? Simply playing the ostrich will not help in the long run. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 28.1. doi: 10.48694/zif.3519 (recuperado el 8 de agosto de 2024).
- Berthele, R. & Udry I.** (2023b). Digitale Übersetzungsprogramme und Online Wörterbücher im Fremdsprachenunterricht: Erkenntnisse aus einer Umfrage bei Lehrpersonen und Lernenden der obligatorischen, post-obligatorischen und tertiären Bildung. *Linguistik Online*, 120.2, 145-167. doi: 10.13092/lo.120.9720.
- Blum-Kulka, S.** (1991). Interlanguage Pragmatics: The Case of Requests. In R. Phillipson et al. (eds.), *Foreign/second Language Pedagogy Research*. Clevedon: Multilingual Matters, 255-72.
- Brocca, N.** (2021). LADDER: Un corpus di scritture digitali per l'insegnamento della pragmatica in L2. Un esempio di analisi di disdetta in WhatsApp. *Italiano LinguaDue*, 13(1), 241-259.
- Félix-Brasdefer, J. C.** (2024). Politeness 1. *Pragmatics at Indiana University*. <https://pragmatics.indiana.edu/politeness/index.html> (recuperado el 6 de agosto de 2024).
- Fredholm, K.** (2019). Effects of Google translate on lexical diversity: vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de las Lenguas*, 13(26), 98-117.
- Hagenhoff, T. & Leienbach, S.** (2022). Übersetzungstools im Spanischunterricht - ein Plädoyer für einen kritischen Umgang. *Hispanorama*, 72-82.
- Ohashi, L.** (2022). The use of machine translation in L2 education: Japanese university teachers' views and practices. En: B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouèsny, & M. J. Whelpton (eds.): *Intelligent CALL, granular systems and learner data: short papers from EUROCALL 2022*, 308-314.
- OpenAI.** 2024. ChatGPT (Version GPT-4o) [Computer software] <https://openai.com/> (recuperado el 6 de agosto de 2024).
- Poole, F. J. y Polio. C.** (2024). From sci-fi to the classroom. Implications of AI in task-based writing. *TASK Vol. 3* (2), 243-272. <https://doi.org/10.1075/task.00022.poo>.
- Siegal, M.** (1996). The Role of Learner Subjectivity in Second Language Sociolinguistic Competence: Western Women Learning Japanese. *Applied Linguistics*, 17(3), 356-382. <https://doi.org/10.1093/applin/17.3.356>
- Trosborg, A.** (1995). *Interlanguage pragmatics. Requests, Complaints and Apologies*. Berlin: Mouton de Gruyter.

INTEGRATION KI-BASIERTER SCHREIBTOOLS IM FREMDSPRACHENUNTERRICHT: CHANCEN UND HERAUSFORDERUNGEN FÜR LERNENDE AUF NIEDRIGEM SPRACHNIVEAU

Current studies point to uncertainty regarding the use of *Machine Translation* (MT) in classrooms and show inconsistent results on the necessary prior knowledge in foreign languages for meaningful use of MT in teaching. However, there is consensus that successful educational use of MT can only be achieved if learners possess adequate *MT-Literacy*. This means understanding the technology, mastering its applications, and knowing how to integrate it sensibly into their learning process. Given that translation machines are used in many classrooms, this paper emphasizes the need to develop *MT-Literacy* from the early stages of language learning. Based on findings from two projects at PHBern, an approach for integrating AI tools into writing tasks for lower secondary education is outlined.

● Brigitte Reber
| PHBern
Meike Raaflaub
| PHBern



Dr. Brigitte Reber
ist Dozentin für
Fachdidaktik Englisch
an der PHBern.



Meike Raaflaub
ist Dozentin für
Fachdidaktik Englisch
an der PHBern.

Einleitung

Angesichts des weit verbreiteten Einsatzes von KI-basierten Schreibtools wie DeepL auch unter Schüler*innen der Volksschulstufe (Raaflaub & Reber 2022) stellt sich die Frage, ob und unter welchen Bedingungen diese Tools im Fremdsprachenunterricht mit Lernenden auf niedrigem Sprachniveau eingesetzt werden sollen. Schüler*innen und Lehrpersonen stehen der Integration solcher Tools ambivalent gegenüber (Deng & Yu 2022; Raaflaub & Reber 2022; Perrin et al. 2022). Lehrpersonen argumentieren beispielsweise, dass Lernende *Machine Translation* (MT) zum Schreiben von Texten nutzen würden, anstatt sich selbst anzustrengen, und verbieten deshalb deren Nutzung oder beschränken diese auf Einzelwortübersetzungen (Raaflaub & Reber 2022). Im Rahmen dieses Artikels werden didaktische Ansätze zur erfolgreichen Integration von KI-basierten Schreibtools in den Unterricht präsentiert und diskutiert.

KI-basierte Schreibtools im Unterricht: Chancen, Herausforderungen und didaktische Ansätze

Die Gründe für den Einsatz von KI-basierten Schreibtools sind vielfältig. Wichtige Faktoren sind einfache Zugänglichkeit, Schnelligkeit und ständige Verfügbarkeit (vgl. Literaturübersicht zu MT in Jolley & Maimone 2022). Gerade für Lernende, die beim Schreiben in der Fremdsprache Unsicherheiten und Ängste zeigen, können Schreibtools hilfreiche Unterstützung bieten. MT ermöglicht es zum Beispiel, lexikalisch vielfältigere und präzisere Texte zu verfassen (Jolley & Maimone 2022). In der Unterrichtspraxis wird MT hauptsächlich zum Nachschlagen einzelner Wörter genutzt (Udry & Berthele 2023; Perrin et al. 2022). Moderne Übersetzungstools bieten jedoch Funktionen, die über die Einzelwortübersetzung hinausgehen, wie das Aufzeigen von Alternativen auf Wort- und Satzebene, die Möglichkeit des Rückübersetzens oder

das Anhören von Wörtern, Sätzen oder ganzen Texten.

Bisher fehlen klare Aussagen über den Zusammenhang zwischen Spracherwerb und dem Einsatz von Schreibtools sowie welches Sprachniveau erforderlich ist, um diese als Instrument für das Sprachenlernen effektiv nutzen zu können. Einige Wissenschaftler*innen knüpfen zum Beispiel den erfolgreichen Einsatz von MT an ein bestimmtes Sprachkompetenzniveau (Fredholm 2019; Carré et al. 2022). So argumentieren Carré et al. (2022), dass ein Sprachniveau von B1 oder B2 des Gemeinsamen Europäischen Referenzrahmens (Council of Europe 2001) notwendig ist, damit Lernende den Output der Maschine adäquat einschätzen und erfolgreich für ihre eigene Sprachproduktion nutzen können. Klimova et al. (2023) halten in ihrer Überblicksstudie dagegen fest, dass Lernende mit niedrigem Sprachniveau dank dem Einsatz von MT eine effektivere Kommunikation in der Zielsprache erreichen können. Sie kommen mit lexikalischen Elementen in Berührung, denen sie sonst nicht begegnet wären, was den inzidentellen Spracherwerb unterstützen kann. Weiter wird argumentiert, dass MT Schüler*innen auf tieferem Sprachniveau unterstützen kann, eigenständig Sprachwissen zu generieren (Reinhardt 2023: 269).

Gerade bei Lernenden auf niedrigem Fremdsprachenniveau kann ein potenziell sprachförderlicher Einsatz von KI-basierten Schreibtools nur gelingen, wenn sich die Schüler*innen aktiv mit dem sprachlichen Input in die Maschine und deren Output auseinandersetzen. Dazu müssen Lernende über ausreichende Kenntnisse (z.B. MT-Literacy) verfügen (O'Brien & Ehrensberger-Dow 2020; Reinhardt 2023). Das bedeutet, dass sie Wissen zur Funktionsweise und zu Nutzungsmöglichkeiten und Grenzen der Tools aufbauen. Für die Unterrichtspraxis ist es sinnvoll, mit den Lernenden über den lernförderlichen Einsatz zu reflektieren, anstatt deren Verwendung als "Schummelei" abzutun oder den Einsatz auf bestimmte Funktionen wie die Einzelwortübersetzung zu beschränken (Raaflaub & Reber 2022). Dabei sollte betont werden, dass die Tools als Ressource für sprachliche Ideen oder Möglichkeiten und nicht als Garant für perfekte Texte gesehen werden müssen (Pym et al. 2013). Nur so werden die Schüler*innen verstehen, dass sie sich

Gerade bei Lernenden auf niedrigem Fremdsprachenniveau kann ein potenziell sprachförderlicher Einsatz von KI-basierten Schreibtools nur gelingen, wenn sich die Schüler*innen aktiv mit dem sprachlichen Input in die Maschine und deren Output auseinandersetzen.

mit dem maschinell generierten Output auseinandersetzen müssen, indem sie gezielt all ihre eigenen sprachlichen Ressourcen und die verschiedenen Funktionen von Schreibtools nutzen.

Integration von KI-basierten Schreibtools in fremdsprachliche Aufgaben

Um besser zu verstehen, wie ein kompetenter und lernförderlicher Umgang mit KI-basierten Schreibtools im Fremdsprachenunterricht anzulegen ist, wurden an der PHBern zwei Entwicklungs- und Forschungsprojekte zu Übersetzungsmaschinen im Englischunterricht durchgeführt. Im ersten Projekt wurden zwei komplexe Aufgabenstellungen für die Jahrgangsstufe 8 (Alter 13-14 Jahre) entwickelt, die den Gebrauch von Übersetzungsmaschinen explizit erlaubten. Diese Aufgaben wurden in zwei Klassen erprobt. Durch Beobachtungen und Interviews mit den Lehrpersonen wurde untersucht, wie die Schüler*innen die Aufgaben bearbeiteten. Die Ergebnisse zeigen, dass sowohl Lehrpersonen als auch Schüler*innen viel Unterstützung benötigten, um MT sinnvoll in den Schreibprozess zu integrieren.

Auf dieser Grundlage wurde in Zusammenarbeit mit einem anderen Projekt (Hofmann 2024) ein Set an Strategien im Umgang mit MT entwickelt und in Expertengesprächen validiert. Folgende Grundüberlegungen (Raaflaub & Reber 2022) waren dabei leitend:

Ein*e MT-kompetente*r Lernende*r sollte

- a) die Voraussetzungen für den Einsatz von MT verstehen (z. B. gut strukturierter und korrekt geschriebener Eingabetext),

- b) verschiedene Techniken für den Einsatz von Übersetzungsmaschinen kennen (z. B. Auswahl von Wortvarianten zur Anpassung und Vereinfachung eines übersetzten Textes, Wissen über die Verwendung der Aussprachefunktion),
- c) entscheiden können, an welchem Punkt im Aufgabenprozess der Einsatz von MT hilfreich sein könnte (z. B. Verstehen von Aufgabenstellung, Überprüfung, ob ein Text für ein bestimmtes Publikum geeignet ist).

Im Anschluss wurde in einem zweiten Projekt eine Unterrichtseinheit aus den obligatorischen Unterrichtsmaterialien von etwa 15 Lektionen so angepasst, dass der Gebrauch von MT im Rahmen bestehender oder angepasster Aufgaben sinnvoll war und die Strategien explizit vermittelt werden konnten. Mittels Reflexionsbögen (N = 89 -112) und Fokusgruppeninterviews (N = 15) vor, während und nach der Einführung der Strategien wurde das MT-Nutzungsverhalten der Schüler*innen untersucht. Diese empirischen Untersuchungen bilden die Grundlage für die folgenden didaktischen Implikationen. Dabei ist zu betonen, dass zur Zeit der Umsetzung der beiden Projekte nur reine Übersetzungstools zur Verfügung standen und Aussagen zu KI-Tools wie DeepL Write oder ChatGPT auf weiteren Überlegungen beruhen.

In beiden Projekten zeigte sich, dass (Schreib-) Aufgaben komplex und prozessorientiert sein müssen, damit sie sinnvoll (auch) mit KI-basierten Schreibtools bearbeitet werden können und eine vielfältige Auseinandersetzung mit Sprache ermöglichen. Zur Illustration des Prinzips der Aufgabenkomplexität soll folgende Aufgabeninstruktion herangezogen werden:

Schreibe sechs Sätze über dein Lieblingslied.

Das Hauptziel der Aufgabe scheint daraus zu bestehen, eine bestimmte Anzahl von Sätzen zu produzieren. KI-basierte Schreibtools können hier relativ einfach einen adäquaten Text liefern, ohne dass sich die Lernenden vertieft mit ihrem In- und Output beschäftigen müssen. Um dies zu vermeiden, müssen digital unterstützte Schreibaufgaben kognitiv herausfordernd respektive komplex sein. Hiermit gewinnen bestehende Forderungen an eine gute Schreibaufgabe wie der

Lebensweltbezug, der kommunikative Kontext und das Verfassen eines spezifischen Produktes mit einer klaren (Text-) Funktion (Hallet 2014) an Bedeutung. Folgende Anpassungen könnten dementsprechend bei der vorher aufgeführten Aufgabeninstruktion gemacht werden:

Verfasse eine Musikkrezension für einen Musikblog über eines deiner Lieblingslieder. Schreibe, welchen anderen Jugendlichen diese Musik ebenfalls gefallen könnte und begründe deine Empfehlung.

Eine weitere wichtige Komponente ist die Bewusstmachung der spezifischen Konventionen eines Genres (Hallet 2016). Eine gute Musikkrezension sollte informativ und überzeugend sein, sich an Jugendliche richten und typische Merkmale wie eine Einleitung, eine Musikanalyse und eine Empfehlung enthalten. Es geht dabei nicht um die reine Produktion von Sprache, sondern um die Erfüllung genregerechter Kriterien, was kritisches Denken und die Fähigkeit erfordert, Informationen auszuwählen, zu strukturieren und adressatengerecht einem konkreten Publikum zu präsentieren. Es kann also nicht einfach der "erstbeste" Output verwendet werden; vielmehr muss er evaluiert und gezielt an die generischen Anforderungen angepasst werden, beispielsweise durch die Auswahl eines für Mitschüler*innen verständlichen Wortschatzes.

Während der Bearbeitung der Schreibaufgaben muss zudem ein Fokus auf den Schreibprozess gelegt werden. Dieser wird als dynamischer, zyklischer Prozess beschrieben, der in verschiedenen Phasen abläuft (Surkamp & Viebrok 2018: 124f.). Die Teilprozesse umfassen:

- Ziele setzen
- Ideen generieren
- Informationen ordnen
- Passende Sprache wählen
- Einen Entwurf verfassen
- Lesen und überprüfen
- Überarbeiten und editieren

(Flower & Hayes 1981)

In jedem dieser Teilprozesse können KI-basierte Schreibtools gezielt zur Schreibunterstützung eingesetzt werden. Wichtig ist, dass den Schüler*innen deutlich gemacht wird, dass sie der Maschine die relevanten Informationen im Input bereitstellen und den maschinell generierten Output sprachlich sowie inhaltlich evaluieren und genregerecht

anpassen müssen. Auch Lernende auf niedrigem Sprachniveau können überlegen, was sie mitteilen möchten, bei Bedarf der Maschine einen passenden Textgenerierungsauftrag erteilen und mithilfe bestimmter Techniken wie der Rückübersetzung oder der Auswahl von Wortvarianten den Output der Maschine kritisch beurteilen.

Da KI-basierte Tools formal korrekte Texte produzieren, verschiebt sich der Fokus des Lernens vermehrt weg von der eigenen Sprachproduktion und hin zu Überlegungen bezüglich Input in die Maschine und zur Evaluation von vorgeschlagenen sprachlichen Bausteinen. Eine solche Auseinandersetzung mit Sprache deckt viele Aspekte der in den meisten Lehrplänen verankerten "Language Awareness" oder "Bewusstheit für Sprache" ab (Association for Language Awareness; D-EDK 2014).

Unterrichtspraktische Instrumente

Als Hilfsmittel zur Bewältigung solcher prozessorientierter Schreibaufgaben sollen im Folgenden zwei unterrichtspraktische Instrumente vorgestellt werden: Ein Entscheidungsbaum zur Strukturierung des Schreibprozesses sowie ein Set an Strategien zur Förderung von MT-Literacy.

a) Entscheidungsbaum

Der Entscheidungsbaum kann dazu dienen, verschiedene Möglichkeiten des (digitalen) Schreibens zu visualisieren und somit individuelle Entscheidungsprozesse unterstützen (vgl. Abbildung 1). Da der Fokus auf dem Umgang mit spezifischen KI-basierten Schreibtools liegt, setzt der Entscheidungsbaum bei den Teilprozessen "Passende Sprache wählen" (und folgende) an. Je nach individuellen Faktoren der Lernenden, wie z.B. Sprachniveau, wird für das Verfassen eines ersten Entwurfs ein Übersetzungstool verwendet, um mithilfe schulsprachlicher Textbausteine adressatengerechte zielsprachliche Textelemente zu generieren. Oder das Verfassen eines ersten Entwurfs geschieht bereits in der Zielsprache und ein KI-Schreibtool wird nur zur Optimierung des Textes eingesetzt. Als dritte Möglichkeit kann ohne (digitale) Unterstützung gearbeitet werden. Der Entwurf wird (mit Hilfe digitaler Tools) überarbeitet. Um den Austausch über Sprache zu fördern

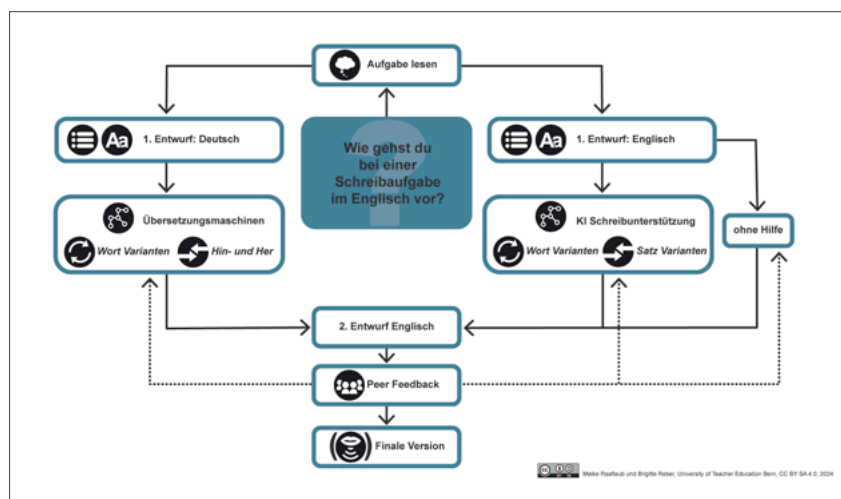


Abbildung 1

Entscheidungsbaum für digital gestützte Schreibaufgaben.



Abbildung 2

Strategien zur Entwicklung von MT-Literacy.

und die Bedeutung der Adressatengerechtigkeit zu betonen, empfiehlt sich für den Teilschritt "Lesen und überprüfen" ein Peer Feedback zu Punkten wie Nachvollziehbarkeit der Aussagen oder zielgruppengerechte Sprache.

b) Strategien

Zur Förderung von MT-Literacy mit Lernenden auf tieferem Fremdsprachenniveau wurde ein Set an Strategien entwickelt (vgl. Abbildung 2), welche auch im Entscheidungsbaum durch entsprechende Icons symbolisiert sind:

Die Strategien Nr. 3, 5 und 7 thematisieren grundlegende Voraussetzungen, um MT sinnvoll einsetzen zu können. In Strategie Nr. 4 wird die Problematik der Einzelwortübersetzungen aufgegriffen. Strategien Nr. 1, 2 und 6 führen in

verschiedene Techniken ein, welche gerade Schüler*innen auf niedrigem Fremdsprachenniveau bei der Textgestaltung und -überarbeitung unterstützen können. Als Strategie und zu entwickelnde Kompetenz zugleich kann Strategie Nr. 8 betrachtet werden. Wie hier am Beispiel von MT gezeigt, müssen Schüler*innen dazu befähigt werden, KI-basierte Schreibtools so zu nutzen, dass sie zwar Unterstützung im Schreibprozess erhalten, die Verantwortung über den Text jedoch bei den Lernenden bleibt.

Schlussbemerkung

Während die Integration von KI-basierten Schreibtools in den Fremdsprachenunterricht den Lernenden helfen kann, präzisere und lexikalisch vielfältigere Texte zu

verfassen, besteht die Gefahr, dass Schüler*innen auf niedrigem Sprachniveau diese Tools nutzen, ohne sich aktiv mit der Fremdsprache auseinanderzusetzen. Um dies zu verhindern, ist es entscheidend, dass Lernende über ausreichende Kenntnisse zu den Tools verfügen und die Technologie reflektiert einsetzen. Dies erfordert eine gezielte didaktische Begleitung und die Entwicklung von Unterrichtssettings, die eine aktive Auseinandersetzung mit dem Input in die Maschine und deren Output fördern. Nur so kann sichergestellt werden, dass KI-basierte Schreibtools als unterstützende Ressource genutzt werden, die den Spracherwerb fördert.

Weitere didaktische Hinweise und Informationen sind auf dieser [Projektwebseite](#) (Raaflaub & Reber) zusammengestellt.

Da KI-basierte Tools formal korrekte Texte produzieren, verschiebt sich der Fokus des Lernens vermehrt weg von der eigenen Sprachproduktion und hin zu Überlegungen bezüglich Input in die Maschine und zur Evaluation von vorgeschlagenen sprachlichen Bausteinen.

Literatur

Association for Language Awareness.

Definition of Language Awareness. *Association for Language Awareness: About*. (https://www.languageawareness.org/?page_id=48) (Accessed August 13, 2023.)

Carré, Alice & Kenny, Dorothy & Rossi, Caroline & Sánchez- Gijón, Pilar & Torres-Hostench, Olga. (2022). Machine translation for language learners. In Kenny, Dorothy (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*, 187–207. Berlin: Language Science Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

D-EDK, Deutschschweizer Erziehungsdirektoren-Konferenz. (2014). Lehrplan 21. *Lehrplan21*. (<http://www.lehrplan21.ch/>) (Accessed February 2, 2022.)

Deng, Xijie & Yu, Zhonggen. (2022). A Systematic Review of Machine–Translation–Assisted Language Learning for Sustainable Education. *Sustainability* 14(13). 7598.

Flower, Linda & Hayes, John R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*. National Council of Teachers of English 32(4). 365–387.

Fredholm, Kent. (2019). Efectos del traductor de Google sobre la diversidad léxica: el desarrollo de vocabulario entre estudiantes de español como lengua extranjera: Effects of Google translate on lexical diversity: vocabulary development among learners of Spanish as a foreign language. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas* 13(26). 98–117.

Hallet, Wolfgang. (2014). Die komplexe Kompetenzaufgabe. Fremdsprachige Diskursfähigkeit als kulturelle Teilhabe und Unterrichtspraxis. In Hallet, Wolfgang & Krämer, Ulrich (eds.), *Kompetenzaufgaben im Englischunterricht: Grundlagen und Unterrichtsbeispiele*, 8–19. 2. Auflage. Seelze: Klett, Kallmeyer.

Hallet, Wolfgang. (2016). *Genres im fremdsprachlichen und bilingualen Unterricht: Formen und Muster der sprachlichen Interaktion*. 1. Auflage. Seelze: Klett/Kallmeyer.

Hofmann, Katrin. (Accepted/In press). Hard to assist? Hard to resist? FONMT in foreign language learning: Instructors' perspectives on learners' rationale. *CALICO Journal*.

Jolley, Jason R. & Maimone, Luciane. (2022). Thirty Years of Machine Translation in Language Teaching and Learning: A Review of the Literature. *L2 Journal* 14(1). 26–44.

Klimova, Blanka & Pikhart, Marcel & Benites, Alice Delorme & Lehr, Caroline & Sanchez-Stockhammer, Christina. (2023). Neural machine translation in foreign language teaching and learning: a systematic review. *Education and Information Technologies* 28(1). 663–682.

O'Brien, Sharon & Ehrensberger-Dow, Maureen. (2020). MT Literacy – A cognitive view. *Translation, Cognition & Behavior*. John Benjamins Publishing Company 3(2). 145–164.

Perrin, Raphaël & Diederich, Catherine & Wild, Sandrine & Grimm, Linda. (2022). Überzeugungen von Lernenden der Sekundarstufe I zur Nützlichkeit der Übersetzungstools DeepL und LEO für die Wortschatzarbeit im Fremdsprachenunterricht. *Babylonia Journal of Language Education* 1. 38–41.

Pym, Anthony & Malmkjaer, Kirsten & Gutierrez-Colon Plana, Mar. (2013). *Translation and language learning – The role of translation in the teaching of languages in the European Union – A study*. Publications Office of the European Union: European Commission: Directorate-General for Translation.

Raaflaub, Meike & Reber, Brigitte. (2022). Übersetzungsmaschinen im Englischunterricht des Zyklus 3. *Babylonia Journal of Language Education* 1. 42–49.

Raaflaub, Meike & Reber, Brigitte. Übersetzungsmaschinen im Fremdsprachenunterricht. *Arbeitsgruppe Digitalität in den Fremdsprachen*. (<https://ttim.phbern.ch/uebersetzungsmaschinen/>)

Reinhardt, Janina. (2023). Sprachbewusstheit 2.0: Digitalised Language Awareness. *PraxisForschungLehrer*innenBildung. Zeitschrift für Schul- und Professionsentwicklung. (PFLB)* 5(3). 258–281.

Surkamp, Carola & Viebrock, Britta (eds.). (2018). *Teaching English as a foreign language: an introduction* (Lehrbuch). Stuttgart: J. B. Metzler Verlag.

Udry, Isabelle & Berthele, Raphael. (2023). Digitale Übersetzungsprogramme und Online-Wörterbücher im Fremdsprachenunterricht: Erkenntnisse aus einer Umfrage bei Lehrpersonen und Lernenden der obligatorischen, post-obligatorischen und tertiären Bildung. *Linguistik Online* 120(2). 145–167.

EXPLORING THE POTENTIAL AND CHALLENGES OF CHATGPT IN LANGUAGE EDUCATION: AN INVESTIGATION INTO STUDENT TEACHERS' CRITICAL THINKING, INTERACTIONS AND TASK-BASED ACTIVITY DESIGN

Tema

L'introduzione di ChatGPT ha destato alcune preoccupazioni in numerosi ambiti, tra cui quello dell'educazione linguistica. Pertanto, è importante che studenti ed educatori acquisiscano le competenze necessarie per usare in modo appropriato gli strumenti di intelligenza artificiale (IA) nell'insegnamento e nell'apprendimento. Sebbene alcune ricerche abbiano esplorato le potenzialità ed i limiti di ChatGPT, oggi gli studi riguardanti la sua applicazione nell'attività didattica e il suo impatto sulle capacità di pensiero critico sono ancora esigui. Questa ricerca esplorativa indaga l'utilità di ChatGPT nel realizzare delle attività di *Task-Based Language Learning* (TBLL), il suo impatto sul pensiero critico degli studenti aspiranti insegnanti e l'interazione dei partecipanti con lo strumento. Attraverso un approccio *mixed method*, i dati raccolti evidenziano che i partecipanti sono stati in grado di creare delle attività di TBLL di buona qualità, risparmiando del tempo con ChatGPT. Tuttavia, circa la metà dei soggetti coinvolti nella ricerca ha copiato buona parte degli output forniti dal modello e sono emerse delle difficoltà nell'interazione con ChatGPT. Sulla base dei risultati ottenuti, sono state elaborate alcune soluzioni e proposte per studi futuri in merito all'uso di ChatGPT nell'educazione linguistica.

● Corinne Ca' Foscari
Baradel
| University of Venice



Corinne Baradel holds a Master's degree in Language Sciences from Ca' Foscari University of Venice and a Master's degree in Teacher Development for Foreign Language Education from Universitat Autònoma de Barcelona. Her research interests focus on the integration of AI-driven tools and digital technologies in language education.

Introduction

Recently, integrating Artificial Intelligence (AI) technologies into daily life has transformed multiple sectors, including education. Current AI tools have enormous potential to enhance learning and teaching, make education globally accessible and tailor learning to students' needs. Among these tools, ChatGPT (OpenAI, 2022) has attracted significant attention from researchers and academics for its capabilities. In fact, this Large Language Model (LLM) chatbot can serve as a virtual assistant, generating human-like conversations, answering questions and clarifying complex concepts (Kasneci et al., 2023).

The educational potential of AI tools, however, is accompanied by security risks and ethical concerns. These technologies often collect sensitive data, necessitating responsible data management. Moreover, biases inherent in training data can lead to errors, while concerns about plagiarism, unethical behaviour, diminishing

critical thinking and creativity also arise. Given ChatGPT's emergence and impact on education, this research¹ aims at investigating the following two research questions amidst university learners: to what extent is the usage of ChatGPT useful for student teachers to design Task-Based Language Learning (TBLL) activities (Willis, 1996) addressed to high-school language students, in terms of planning time, quality of the final products and development of their critical thinking? What are the strategies used by student teachers when interacting with ChatGPT to create a TBLL activity?

Literature review

ChatGPT is an AI chatbot (OpenAI, 2022) trained on extensive text and able to generate human-like responses in multiple languages (García-Peñalvo, 2023). Its outputs, however, are strictly influenced by the quality of user prompts (OpenAI, 2022; Jalil et al., 2023). Despite its potential, the AI chatbot may generate biased

¹ This paper is based on the master's thesis by the researcher (Baradel, 2023). Some sections of the text may replicate the original content from her master's thesis.

content or provide incorrect information due to the nature of its pre-training (OpenAI, 2022; Jalil et al., 2023; Rudolph et al., 2023).

In the field of language education, researchers (Kasneci et al., 2023) asserted that, on the one hand, ChatGPT can be a valuable resource for students throughout their language learning process. In fact, the AI tool may improve both writing and reading abilities by assisting in grammatical and syntactic corrections, summarising complex texts, creating quizzes, organising thoughts during the writing process and providing feedback. On the other hand, in the teaching practices, the chatbot may aid teachers in organising lesson plans and creating engaging materials. Therefore, teachers could save both efforts and time. Furthermore, empirical studies (Ali et al., 2023; Muñoz et al., 2023) demonstrated ChatGPT's efficacy in increasing student engagement and motivation, particularly in subjects like languages.

While ChatGPT offers numerous benefits in language education, various challenges in this sector need consideration. According to Kasneci et al. (2023), one concern is the difficulty in distinguishing AI-generated texts from human writing, which may undermine academic integrity. Hence, using ChatGPT in class may boost plagiarism and unethical behaviour (Lin et al., 2023). A further risk is developing an overreliance on ChatGPT, neglecting to use critical thinking abilities (Kasneci et al., 2023), defined as the combination of higher-order thinking skills (deduction, analysis, reasoning, problem solving and evaluation) (Bellaera et al., 2021). These capacities allow human beings to make reasonable decisions, interpret and analyse large amounts of information rationally and objectively (Thornhill-Miller et al., 2023).

Given the aforementioned concerns, educators should integrate ChatGPT responsibly into educational environments by teaching students to critically evaluate ChatGPT's outputs in order to avoid possible false information or bias (Bitzenbauer, 2023). Moreover, teachers are advised to elaborate higher-order assignments that emphasise skills ChatGPT lacks, such as providing emotional and real-world practical situations (Liu & Wei, 2024).

Current AI tools have enormous potential to enhance learning and teaching experiences, make education globally accessible and tailor learning to students' needs.

Besides ChatGPT and its implications on language education, for this research it is important to introduce the Task-Based Language Learning (TBLL) method proposed by Willis (1996). This framework for language education focuses on using meaningful tasks as the core of the learning process and includes three phases. In the pre-task, learners are introduced to the topic and related vocabulary, often participating in activities that activate their prior knowledge. During the task cycle, students in groups perform, plan and report the task and their results. Finally, the post-task refers to the language focus, in which learners not only analyse and reflect on the language used throughout the task, but also practice certain language features to enhance their understanding and accuracy.

Methodology

Considering both potentials and limits of ChatGPT in language education, an exploratory study was conducted to comprehend the usefulness of ChatGPT in designing TBLL activities (Willis, 1996) in terms of planning time and quality, the effects on student teachers' critical thinking skills and the interaction strategies between ChatGPT and student teachers. The current investigation involved 39 voluntary postgraduate participants (referred to as *student teachers*) studying the subject Instructional Design for Language Education. The participants in groups or individually had to design a TBLL activity with the usage of ChatGPT. It is important to underline that before this assignment, the student teachers attended a lesson where the TBLL model (Willis, 1996) was deeply explained and ChatGPT was briefly introduced.

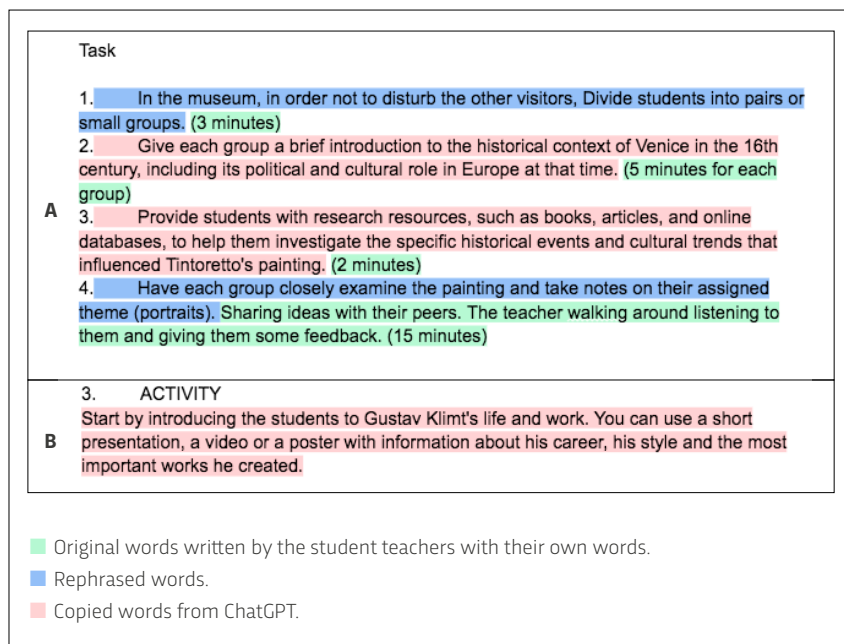


Figure 1

Illustration of original, copied, and rephrased words from two TBLL activity samples.

Note. Reprinted from *Interaction, Design, and Assessment: An Exploratory Study on ChatGPT in Language Education* by Baradel (2023). <http://hdl.handle.net/10579/25711>.

The assignment required designing a TBLL activity (Willis, 1996) based on a specific theme (food education, selfies and portraits or anything related to sustainability) for 16-to-18-year-old students with a B1 English level. The task had to be set in a museum, focussed on speaking skills and include learning aims, a detailed description of the phases, exercises, materials and ICT tools used.

In April 2023, data was collected anonymously through a mixed-method approach (Dörnyei, 2007), including 33 student teacher online surveys, 2 student teacher online interviews, 39 student teacher interactions with ChatGPT (based on the GPT-3.5 model) and a total of 14 final TBLL activities. A custom ChatGPT webpage was created to gather the chats between the student teachers and AI chatbot, enabling the researcher to directly analyse their interactions, minimising any self-report bias.

The participants were expected first to interact with the chatbot, second, to upload the TBLL activity on Moodle and finally, to fill in the online questionnaire. In the end, the participants were invited for an interview. The evaluation criterion

for the TBLL activity required the participants to simply publish it on Moodle. If the activity adhered to the TBLL structure (Willis, 1996), the student teachers would obtain 0.3 points to add to their final grade.

After gathering the data, the information was organized to address the research questions. The quantitative findings from the online survey were integrated with qualitative insights into open-ended questions and online interviews. Content analysis (Dörnyei, 2007; Evans, 2017) was used to examine the responses to the open-ended questions in the online questionnaire. In order to compare student teachers' interactions with ChatGPT to their final TBLL activities, a quantitative content analysis using several categories (original, copied, rephrased and copied-exercise words) was conducted and word counts for each were calculated. To identify the copied sequences, each conversation and related TBLL activity were analysed by employing the *Similarity Texter* tool (Kalaidopoulou & Weber-Wulff, 2016) (see an example in Figure 1). No further tools were used to verify whether original words were copied from other sources. Regarding interactions between student teachers and ChatGPT, a further content analysis (Dörnyei, 2007; Evans, 2017) was performed categorising the different types of prompts provided by the participants to ChatGPT.

Research findings

The results consisted in examining to what extent ChatGPT was useful for student teachers to design TBLL activities in terms of planning time, quality of the final products and development of student teachers' critical thinking skills. It is important to mention that some student teachers (39.4%) did not know ChatGPT and the majority (78.8%) did not use the AI chatbot before this research. The first findings demonstrated that all student teachers (100%) considered ChatGPT as a time-saving AI tool to create TBLL activities, especially in the creation of exercises (84.8%), brainstorming ideas (72.7%) and structuring the task (66.7%).

The second outcomes suggest that most student teachers (97%) believed they had created a better or same quality TBLL activity thanks to ChatGPT. In fact, student

teachers' interviews revealed that original ideas for their classes came up and their self-confidence in teaching abilities increased when utilising ChatGPT.

The third results reveal a mixed experience among student teachers when using ChatGPT. A majority (61.3%) reported not encountering any difficulties or limitations with the AI tool. In contrast, 38.7% did experience challenges. Specifically, only 25.8% of all student teachers emphasised the importance of critically analysing ChatGPT's outputs and expressed dissatisfaction with the answers provided by the chatbot. Participants noted that they often needed to rewrite certain parts and that the AI sometimes hallucinated links or provided inaccurate information about museums. Furthermore, 19.4% encountered difficulties in formulating prompts to obtain meaningful answers or suggestions, particularly during the initial phase of use. In general, these outcomes may raise concerns about an excessive reliance on ChatGPT. In fact, researchers (Cooper, 2023; Jalil et al., 2023; Rudolph et al., 2023) underlined that human judgment is necessary when using ChatGPT to interpret results and make ethical decisions, ensuring the AI tool is applied appropriately.

Indeed, the overall findings indicate a pressing need to organise trainings for student teachers in proper ChatGPT use, outlining its limits to enhance critical thinking skills. This recommendation aligns with previous studies (Kasneci et al., 2023), supporting continuous training for teachers and students to keep updated on ChatGPT's latest advancements.

Furthermore, the comparison between student teachers' interactions with ChatGPT and their related final TBLL activities (Figure 2) reveals that, on the one hand, 6 out of 14 TBLL activities present minimal copying from ChatGPT (less than 14% of copied words). In fact, checking one by one those activities, it can be seen that the student teachers analysed, interpreted, elaborated, and modified ChatGPT's responses, demonstrating critical thinking skills. Moreover, in the two interviews, the student teachers explained that each member studied the structure and the characteristics of TBLL model before using ChatGPT. Then, they discussed the content and elements to include in their activity. Finally, the

The overall findings indicate a pressing need to organise training courses for student teachers in the proper use of ChatGPT, outlining its limits to enhance their critical thinking skills.

group members interacted with ChatGPT and collaboratively reflected on each of its responses.

On the other hand, 6 out of 14 TBLL activities contain a high percentage (more than 45%) of copied text. This result may suggest a mix of psychological, pedagogical and practical factors that could explain the observed reliance on copied words. First, student teachers possibly faced time constraints, causing them to prefer the quicker approach to copy words directly. Second, the participants might lack confidence in their teaching abilities or knowledge about the TBLL methodology (Willis, 1996), which led to an increased tendency to trust ChatGPT's responses and opt for copying to avoid mistakes. Third, the instructions of how to perform the task could have been misunderstood. Fourth, a few student teachers may be primarily focussed on getting some points easily for their final grade in the university subject. Consequently, they copied ChatGPT's answers without applying critical thinking skills.

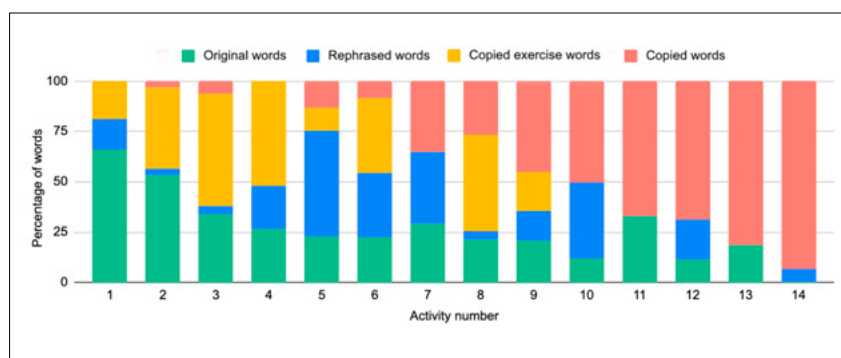


Figure 2

Bar chart displaying the percentages of original, rephrased, copied exercises and copied words from the student teacher ChatGPT conversations in the 14 TBLL activities. Note. The analysis was performed solely using the Similarity Texter tool (Kalaidopoulou & Weber-Wulff, 2016).

These final outcomes may indicate that ChatGPT offers valuable ideas for designing TBLL tasks and demonstrates the potential to encourage student teachers to apply critical thinking when using the AI tool, as also indicated by Jiang et al. (2024). For student teachers who did not effectively use their critical thinking skills, these findings underline the need to improve assessment methods to foster intrinsic motivation and address plagiarism. This aligns with the recommendations of some scholars (Liu & Wei, 2024; Rudolph et al., 2023), advocating for assignments that require higher-order thinking and creativity.

The second research question aimed to examine student teachers' interactions with ChatGPT in designing TBLL activities, particularly exploring self-developed strategies without prior training. Outcomes reveal various interaction methods, including requests for general information (16.7%), material creation (8.3%), idea brainstorming (54.2%) and learning objective formulation (25%). Most student teachers asked for precise items, but nearly a third (29.1%) tried to generate the entire task with a single input and prompt quality often lacked critical details, affecting ChatGPT's accurate outputs. According to the scientists Fulford and Ng (2023), inputs have to follow some principles: they should be clear, detailed and include delimiters, when necessary. In addition, the prompts should be fragmented, giving ChatGPT time to process. This means that the correct design of a TBLL activity strongly depended on the detailed student teachers' inputs given to ChatGPT. Moreover, some requests

exceeded ChatGPT's capabilities at that time like requesting for valid internet links, for instance, the student teachers asked: *Could you suggest a brief introductory video about Frida Kahlo?, Can you please link a website where I can find more information about the ongoing exhibition in Padua you mentioned?*, underlining the need for a training in effective interactions and understanding the chatbot's limitations.

Conclusions

Overall, the current exploratory study demonstrates that the student teachers could save time and create equivalent or better-quality TBLL activities using ChatGPT. Moreover, the chatbot was considered a useful AI tool, capable of giving a feeling of self-confidence in their teaching abilities. However, some participants seemed to fail to recognise ChatGPT's limitations, with about half copying outputs directly, indicating possible misunderstanding of assignment instructions or inadequate assessing methods for the subject. Although the student teachers employed various strategies to interact with ChatGPT, some participants faced challenges of writing prompts. To address these issues, the researcher recommends a training course to teach proper interactions with ChatGPT, comprehending its limits, avoiding plagiarism and boosting participants' critical thinking skills. Furthermore, reassessing assignment methods to emphasise critical thinking, problem solving and creativity is strongly advised.

However, the current investigation presents some limitations due to the specific sample and context for data collection; therefore, it would be worthwhile to replicate the study with a broader range of classes from different years or assigning other activities beyond TBLL. Furthermore, future research could delve deeper into a comparative analysis of the TBLL activities designed by student teachers and their application of critical thinking skills before and after a training in ChatGPT usage. Additionally, empirical investigations into ChatGPT's effectiveness in promoting student teachers' self-confidence would be of considerable interest.

For student teachers who did not effectively use their critical thinking skills, these findings underline the need to improve assessment methods to foster intrinsic motivation and address plagiarism.

Bibliography

- Ali, J., Shamsan, M., Hezam, T., & Mohammed, A.** (2023). Impact of ChatGPT on Learning Motivation: Teachers and Students' Voices. *Journal of English Studies in Arabia Felix*, 2(1), 41-49. <https://doi.org/10.56540/jesaf.v2i1.51>.
- Baradel, C.** (2023). *Interaction, Design, and Assessment: An Exploratory Study on ChatGPT in Language Education* [Master's Thesis, Università Ca' Foscari Venezia]. DSpace. <http://hdl.handle.net/10579/25711>.
- Bellaera, L., Weinstein-Jones, Y., Ilie, S., & Baker, S.** (2021). Critical thinking in practice: The priorities and practices of instructors teaching in higher education. *Thinking Skills and Creativity*, 41, 1-16. <https://doi.org/10.1016/j.tsc.2021.100856>.
- Bitzenbauer, P.** (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, 15(3), 1-10. <https://doi.org/10.30935/cedtech/13176>.
- Cooper, G.** (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, 32(3), 444-452. <https://doi.org/10.1007/s10956-023-10039-y>.
- Dörnyei, Z.** (2007). *Research Methods in Applied Linguistics*. Oxford University Press.
- Evans, M.** (2017). Analysing Qualitative Data. In E. Wilson (Ed.), *School-based Research: A guide for education students*. Sage publications.
- Fulford, I., & Ng, A.** (2023, July). *ChatGPT Prompt Engineering for Developers* [Online course]. DeepLearning.AI. <https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/>.
- García-Peñalvo, F. J.** (2023). The perception of Artificial Intelligence in educational contexts after the launch of ChatGPT: Disruption or Panic? *Education in the Knowledge Society (EKS)*, 24. <https://doi.org/10.14201/eks.31279>.
- Jalil, S., Rafi, S., LaToza, T., Moran, K., & Lam, W.** (2023). ChatGPT and Software Testing Education: Promises & Perils. *2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*, 4130-4137. DOI: 10.1109/ICSTW58534.2023.00078.
- Kalaidopoulou, S., & Weber-Wulff, D.** (2016). Similarity Texter [Computer software]. Hochschule für Technik und Wirtschaft Berlin. <https://people.f4.htw-berlin.de/~weberwu/simtexter/app.html>.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., ... Kasneci, G.** (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 1-9. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Jiang, J., Li, J., & Chen, C. H.** (2024). Enhancing Critical Thinking Skills with ChatGPT-Powered Activities in Chinese Language Classrooms. *International Journal of Chinese Language Teaching*, 5(1), 47-73. <https://doi.org/10.46451/ijclt.20240105>.
- Lin, S. M., Chung, H. H., Chung, F. L., & Lan, Y. J.** (2023). Concerns About Using ChatGPT in Education. In Y. M. Huang & T. Rocha (Eds.), *Innovative Technologies and Learning. ICITL 2023. Lecture Notes in Computer Science*, vol. 14099. Springer, Cham. https://doi.org/10.1007/978-3-031-40113-8_4.
- Liu, Y., & Wei, L.** (2024). A Study on Students' Critical Thinking Cultivation in English Teaching in the ChatGPT Era. In I. A. Khan et al. (Eds.), *Proceedings of the 2024 2nd International Conference on Language, Innovative Education and Cultural Communication (CLEC 2024)*, Advances in Social Science, Education and Humanities Research 853. Atlantis Press. https://doi.org/10.2991/978-2-38476-263-7_10.
- Muñoz, S. A. S., Gayoso, G. G., Huambo, A. C., Tapia, R. D. C., Incaluque, J. L., Aguila, O. E. P., Cajamarca, J. C. R., Acevedo, J. E. R., Rivera, H. V. H., & Arias-González J. L.** (2023). Examining the Impacts of ChatGPT on Student Motivation and Engagement. *Social Space*, 23(1), 1-27.
- OpenAI.** (2022, November 30). Introducing ChatGPT. OpenAI. <https://openai.com/index/chatgpt/>.
- Rudolph, J., Tan, S., & Tan, S.** (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, 6(1), 1-22. <https://doi.org/10.37074/jalt.2023.6.1.9>.
- Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., Hayek, S. E., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T.** (2023). Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *Journal of Intelligence*, 11(3), 54. 1-32. <https://doi.org/10.3390/jintelligence11030054>.
- Willis, J.** (1996). *A framework for task-based learning*. Longman.

ENLISTING CHATGPT TO ENHANCE LANGUAGE TEACHER TRAINING

L'intégration de l'IA, en particulier du ChatGPT, dans la formation des enseignants de langues évolue rapidement et présente à la fois des opportunités et des défis. Cet article examine comment le ChatGPT peut être utilisé comme outil dans la formation des enseignants d'anglais langue étrangère (EFL) pour améliorer et évaluer de manière formative l'apprentissage des enseignants. Par le biais d'une pratique exploratoire en classe, l'étude examine comment ChatGPT soutient le développement de connaissances didactiques théoriques et pratiques et de compétences d'auto-réflexion critique. L'article détaille les procédures d'utilisation de ChatGPT, évalue son alignement sur les principes didactiques et discute des avantages et des limites de l'IA dans le développement professionnel. Il conclut que si le ChatGPT offre de précieuses possibilités d'auto-évaluation personnalisée et un retour d'information en temps réel, son intégration nécessite une gestion prudente afin d'équilibrer le soutien de l'IA avec l'apport humain essentiel.

● Holli Schaubert
| University of Geneva



Holli Schaubert, Ed.D., teaches EFL Didactics and Research Methods in Foreign Language teaching in the IUFE, at the University of Geneva. Her teaching and research are conducted as related activities.

Introduction

The integration of AI in language teacher education has rapidly gained our attention sparking a considerable incentive to explore, understand, and harness the new opportunities it offers. While the novelty accompanies intense enthusiasm, many educators are scrambling to keep pace with the flood of anecdotal accounts and research-driven claims highlighting its innovative and harmful potential. Endorsements and cautionary words dot the current literature (Moorhouse, 2024; Moorhouse & Kohnke, 2024; Karan & Angadi, 2023; Kohnke et al., 2023; Sok & Heng, 2023; Mai et al., 2024). Notwithstanding divided claims, the preoccupation with the role AI can play in language teacher preparation has led to a wave of exploratory in-class practices to which this article seeks to contribute. For the most part, those initiatives and the case described here currently represent ancillary synchronous teacher training events and tools rather than didactic cornerstones as we learn how to

navigate the possibilities, risks, and challenges AI presents. As part of the current exploratory wave, this article provides a practice-driven overview illustrating how ChatGPT serves as a teacher training tool to deliver innovative in-class learning opportunities that support, enhance, and reinforce the fundamental theoretical principles and practical frameworks of professional development in English as a foreign language (EFL) didactics (Bodong et al., 2023; Johnson & Golombek, 2020; Starkey, 2020). A theory-driven practical example of L2 listening is presented (selected because it was the focus of the didactics class on the day ChatGPT was used) alongside procedural details, followed by a discussion of how ChatGPT supports, benefits and enhances EFL teacher learning in light of course objectives.

Didactics course objectives and the integration of ChatGPT

These objectives emphasize informed instruction-driven skills, reflective practices, and professional development guided by the EPOSTL (The European Portfolio for Student Teachers of Languages). Several overarching objectives drive the EFL didactics course and the teacher education approach used to initiate trainees into the field with the support of ChatGPT:

- Develop an informed didactic perspective to design coherent, logically sequenced EFL lessons that respond to diverse learner needs and the cantonal curriculum.
- Apply various teaching methodologies and strategies effectively
- Develop a theory-driven practical understanding of how to optimize the teaching of the four language skills: reading, **listening**, speaking, and writing, and the subsystems of grammar, vocabulary, and pronunciation within the larger interactive and communicative frameworks resting on the can-do descriptors in the EPOSTL.
- Develop the ability to critically reflect on and assess one's teaching and knowledge against the backdrop of the theoretical principles examined in the didactics course through active and autonomous learning affordances that prompt trainees to test, revisit, and reinforce their understanding of the didactics course content.

The integration of ChatGPT is specifically aligned with those objectives and is intended to:

1. Evaluate whether the prompt requests entered into ChatGPT generated teaching principles and notions commensurate with the theoretical frameworks and practical models examined in the didactics course.
2. Facilitate the use of ChatGPT as a tool for collaborative problem-solving and lesson brainstorming among teachers.
3. Prompt trainees to question ChatGPT's capacity to support their teaching objectives, as well as the learning objectives and affordances they intend for their learners.
4. Formatively evaluate trainee concept development around teaching L2 listening.

The role AI can play in language teacher preparation has led to a wave of exploratory in-class practices

After an informal in-class survey of over 25 trainee teachers enrolled in two EFL didactics courses, the majority admitted to using both the free and paid versions of ChatGPT to generate lesson plans, materials, activities, and prompts for their teaching practicum. Both the free and paid versions of ChatGPT generate output text based on user input prompts; however, only the paid option, ChatGPT 4.0, provides more detailed, nuanced, and accurate responses. While many trainees claimed to be capable of generating those teaching resources themselves (as required by their teacher educators and practicum mentors), they admitted relying on ChatGPT because of the enormous time-saving advantages it provided. However, they likewise revealed their uncertainty about whether the output generated by either version of ChatGPT systematically resonated with the theoretical principles discussed in the didactics courses. Motivated by these admissions, a decision was made to formally introduce ChatGPT into the didactics course as a formative teacher and self-evaluation tool to determine whether trainees were using suitable prompt requests to generate desired outputs from ChatGPT and whether they could determine how well those outputs aligned with the principles and practices of the theoretical frameworks examined in the course.

Procedures used for Integrating ChatGPT:

The procedures for integrating ChatGPT sought to actively engage trainees in critically evaluating: a) ChatGPT's output in comparison to their own professional knowledge; b) the suitability of the prompts; and c) the presence or absence

of the theory-driven practical considerations discussed in the didactics course. For each ChatGPT session, the following procedure was implemented:

1. **Solicit a Prompt Request:** Students offered a prompt request for ChatGPT to simulate an activity or set of questions linked to the theme of the didactics class. Typically, the first volunteered prompt was entered, although often more than one prompt is required to generate text that aligns closely with the learning objectives/output the trainees seek.
2. **Enter the Prompt:** The teacher educator entered the prompt into ChatGPT on the screen, while some students did the same in parallel on their own laptops.
3. **Review the Output:** The group reviewed the output generated by ChatGPT within a short timeframe individually or in tandem with classmates.
4. **Evaluate Alignment:** The group collectively evaluated whether the generated output aligned with the principles, models, and frameworks discussed prior to the simulation. Trainees began to offer their thoughts on the output's link to the theory. Some trainees took notes to add to their current understanding of L2 listening.
5. **Identify Elements and Suggest Modifications:** Trainees were asked to identify elements that were present or absent in the output and to suggest how it might be modified in light of their contextual challenges and relevant theory. We also briefly discussed the need to be precise

and adjust input prompts, but did not explore that idea further in this context.

6. **Evaluate Understanding:** Trainees used this event to evaluate their understanding of the didactic concepts by assessing whether they could reasonably explain whether the ChatGPT output reliably resonated with the theory. This is when requests for clarification of their understanding emerged, and when they could confirm or refute their comprehension.

Example 1. Prelistening

In this output example, ChatGPT is enlisted to simulate elements considered crucial for effective pre-listening, allowing trainees to evaluate the output's alignment with theoretical principles examined and suggest modifications as needed.

Summary of the theory discussed in the didactics class on L2 listening

In the didactics course, we explore examples of listening sequences that emphasize pre- during and post-listening events that activate schema and encourage the learner to circle back to confirm or refute content generated during the pre-listening phase. Our examination of L2 listening prioritizes pre-listening because of the way it focuses learner attention and provides a purpose for listening and prepares students to monitor their comprehension. We examine how the teacher can demonstrate the development of listening skills and strategies via pre-listening questions and predictions, front-loading and/or scanning of key vocabulary, and the schema activation of content via visuals, anecdotal references, discussion, or semantic mapping. These focal points are set against the backdrop of the principles of second language acquisition and cognitive psychology as events that scaffold learners' comprehension by activating their prior knowledge and arming them with strategies that they can apply independently in follow-up listening activities (Goh & Vandergrift, 2021). It is through this framework for a listening sequence that students in the didactics course were asked to evaluate ChatGPT output pre-listening to compare it with their own didactic knowledge and understanding.

The integration of ChatGPT into EFL teacher training reveals both innovative opportunities and noteworthy risks.

ChatGPT prompt suggested by a trainee teacher:

what should you include in prelistening for EFL students?

ChatGPT output:

Prelisting activities for EFL (English as a Foreign Language) students are crucial for setting the stage for successful comprehension and engagement with the listening material. These activities should aim to activate prior knowledge, build context, and prepare students for the specific content and language they will encounter. Here are some key components to include in prelistening activities:

1. **Activate Prior Knowledge:** Begin by activating students' prior knowledge about the topic of the listening task. This could involve discussing related vocabulary, concepts, or personal experiences.
2. **Predictions:** Encourage students to make predictions about the content of the listening material based on the title, topic, or any visual aids provided. This helps set a purpose for listening and primes students to anticipate key information.
3. **Vocabulary Preview:** Introduce or review any important vocabulary or expressions that students may encounter in the listening. This could involve pre-teaching keywords, discussing their meanings, or providing context to aid understanding.
4. **Listening Task Instructions:** Clearly explain the task or tasks students will complete while listening. This could include specific questions to answer, information to note down, or gaps to fill in a summary.
5. **Schema Activation:** Help students activate their background knowledge related to the listening topic. This might involve brainstorming related ideas, discussing relevant cultural or contextual information, or making connections to previous learning.
6. **Listening Strategies:** Teach and review effective listening strategies with students, such as predicting, summarizing, and focusing on key information. Encourage students to employ these strategies while listening to enhance their comprehension.

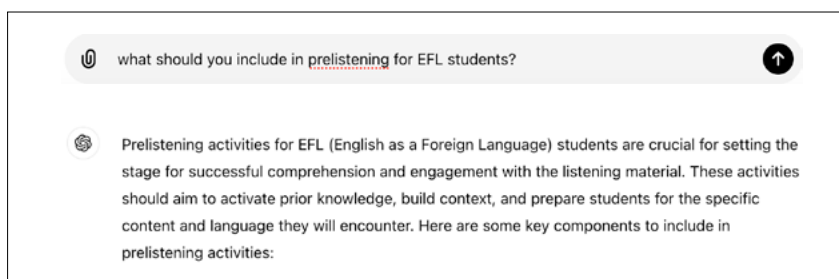


Figure 1

ChatGPT prompt-response interaction

7. **Modeling:** Model the listening task or a similar task yourself to demonstrate effective listening strategies and provide an example of how to approach the activity.
8. **Visual Aids:** Use visual aids such as pictures, diagrams, or videos to support comprehension and provide context for the listening material.
9. **Interactive Activities:** Engage students in interactive activities related to the listening topic, such as discussions, role-plays, or games. This helps activate their interest and engagement before listening.
10. **Review:** Quickly review any relevant grammar or language points that students may need to understand the listening material. This could include reviewing verb tenses, sentence structures, or pronunciation patterns.

By including these elements in your pre-listening activities, you can effectively prepare EFL students to engage with and comprehend the listening material.

Trainee evaluation of the output

Elements aligned with principles discussed in the didactics course

- ✓ aligned with top-down schema activation through interaction, images, and personalization of the topic, which allows learners to connect on some level with the content;
- ✓ aligned with providing learners with a focus for listening through predictions and task instructions so they listen actively;
- ✓ aligned with bottom-up schema activation through frontloading of vocabulary and grammar so that their encounter with the language is not a first exposure;
- ✓ aligned with enabling listening strategy application through teacher demonstration

Elements deemed missing from ChatGPT output

- Even though the output refers to language points and pronunciation patterns, trainees identified word stress as an integral part of pre-listening and believed it should have been specifically mentioned.
- Pre-listening questions were also deemed missing from the output and trainees pointed out that learners can either generate their own questions or can preview the questions in the coursebook.
- Circling back in the post-listening to answer these questions allows the learners to check their comprehension of the content. It also sets a purpose for the listening event and creates a possible instructional sequence that helps learners avoid getting stuck in bottom-up processing.

Perspectives on the benefits to teacher development and the attainment of course objectives

This experience helped trainee teachers strengthen their conceptual knowledge through active engagement in teacher learning, the fundamental objective of the didactics course. They found that using ChatGPT to test themselves on L2 listening principles and practices in the didactics course was beneficial for the following reasons:

- ✓ Trainees recognized whether they could confidently evaluate the resonance of the output with their understanding of the range of events and elements associated with prelistening's role in developing learners' listening skills.
- ✓ Trainees admitted that the opportunity to question one's understanding of theory reinforced and consolidated their professional knowledge.
- ✓ Trainees declared that evaluating the output created space to address any gaps in their understanding, particularly if the output featured information that they were not previously aware of or had not considered as part of their self-evaluation, reflection, and examination.
- ✓ Trainees claimed that generating prompts empowered them to determine the validity and reliability of ChatGPT output and recognize

whether human expertise was necessary for a critical analysis of the output.

- ✓ Trainees admitted that the process reinforced teacher autonomy and the reflective elements promoted in the course.
- ✓ Trainees agreed that the speed of the process made ChatGPT a highly valuable and easily accessible teacher learning and self-evaluation tool.
- ✓ Trainees appreciated revisiting and reinforcing theoretical concepts individually and collaboratively
- ✓ Trainees held that it confirmed their understanding of key concepts in L2 listening

To the teacher educator, the ChatGPT prompted process exposed areas of misunderstanding about the practical elements involved in activating pre-listening. These misunderstandings included how much time to spend on the pre-listening phase, what role they and their learners play in the pre-listening phase, and what combinations of pre-listening events to target to set a purpose and focus for the listening (activating schema about the content, front-loading vocabulary and word stress, student-generated pre-listening questions and predictions). Trainees expressed their appreciation for the immediate teacher-educator feedback while recognizing the need for additional repetition and reinforcement to enhance their professional knowledge and improve their lesson planning for L2 listening instruction. What is also noteworthy is how ChatGPT personalized and scaffolded the learning process by revealing the trainees' individual needs for follow-up clarification on pre-listening (Mai et al., 2024). Each trainee interacted with the output individually, signaling to the rest of the class whether they could independently and confidently evaluate it from the perspective of the L2 listening principles examined in the didactics course. That further necessitated an appeal to the teacher educator's input for confirmation beyond what ChatGPT and their classmates offered.

While these perspectives highlight how compatible the ChatGPT teacher learning event was with the course objectives of developing professional knowledge through active engagement involving critical self-reflection and evaluation, they likewise expose ChatGPT's

limitations as an independent, complete, and fundamental source of conceptual input and suggest the continued need for teacher educator oversight and backup. Notwithstanding the need for human intervention, ChatGPT offers a meaningful, scaffolded, and active way to refine the teacher learning experience that is not only easy to integrate but is on the training tip of one's fingertips. ChatGPT's real-time tutorial function provides trainees with opportunities to revisit and adjust their thinking en route to deepening their understanding of key language teacher fundamentals that may not yet be part of their professional repertoire or are at least developing. It should not be viewed as a competitor but rather as an ally.

Conclusions and Implications for Language Teacher Training

The integration of ChatGPT into EFL teacher training reveals both innovative opportunities and noteworthy risks that underscore the need for a **balanced** approach aligned with course objectives and teacher educator expertise. In practical terms, it promotes engagement with core language teaching principles that encourage trainees to confirm, question, reflect upon, and expand their professional knowledge. These form part of ChatGPT's unique combination of teacher training elements that enhance teacher learning through instantly adaptive, scaffolded support, synchronous formative feedback, and personalized learning mediated by the teacher educator. Despite its appeal as a novel teacher training tool, over-reliance on AI risks undermining or eliminating critical reflection, evaluation, and independent decision-making, essential skills for informed lesson preparation and instruction. Teacher educators and trainees should first be prepared to critically assess and challenge the output and, second, question whether the prompts entered need adjusting. Both require expertise, as this case illustrates. Drawing a direct link between what we teach trainees about pre-listening, the ChatGPT output omitted the important role of word stress in acclimating learners to the words they encounter, of learner-generated questions in providing a focus and a purpose for listening and of revisiting those questions during the post-listening phase to

confirm or refute comprehension. Additionally, while ChatGPT personalizes learning, it may unintentionally limit valuable interactions with the expertise of mentors and peers.

In conclusion, although ChatGPT can reinforce didactic concepts and provide real-time feedback, it must be managed carefully. Guided by the teacher educator's expertise, trainees should critically evaluate AI-generated responses against expert pedagogical knowledge and classroom realities, adjusting prompts to generate reliable output while dismissing ineffective ones. As AI evolves, its role in teacher training should complement traditional methods to ensure the human element remains central to education.

References

- Bodong Chen, Xinran Zhu, & Fernando Díaz del Castillo H.** (2023). Integrating generative AI in knowledge building. *Computers and Education. Artificial Intelligence*, Volume 5.
- Newby, D., Allan, R., Fenner, A. B., Jones, B., Komorowska, H., & Soghikyan, K.** (2007). *European Portfolio for Student Teachers of Languages: A reflection tool for language (EPOSTL)*. European Council on Modern Languages. Graz, Austria: Council of Europe.
- Harari, Y.N.** (2019). *21 Lessons for the 21st Century*. London: Vintage Publications.
- Goh, C.C.M., & Vandergrift, L.** (2021). *Teaching and Learning Second Language Listening: Metacognition in Action* (2nd ed.). Routledge. <https://doi.org/10.4324/9780429287749>
- Johnson, K.E., & Golombek, P.R.** (2020). Informing and transforming language teacher education pedagogy. *Language Teaching Research*, 24 (1) (2020), pp. 116-127.
- Karan, B., & Angadi, G. R.** (2023). Potential Risks of Artificial Intelligence Integration into School Education: A Systematic Review. *Bulletin of Science, Technology & Society*, 43(3-4), 67-85. <https://doi.org/10.1177/02704676231224705>
- Kohnke, L., Moorhouse, B.L., & Zou, D.** (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54 (2), pp. 537-550.
- Moorhouse, B.L. & Kohnke, L.** (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators, *System*, Volume 122.
- Mai, D.T.T., & Da, C.V., & Hahn, N.V.** (2024). The use of ChatGPT in Teaching and Learning: A systematic review through SWOT analysis approach. *Frontiers in Education*, 9. <https://doi.org/10.3389/educ.2024.1328769>
- Moorhouse, B.L.** (2024). Beginning and first-year language teachers' readiness for the generative AI age, *Computers and Education: Artificial Intelligence*, Volume 6.
- Sok, Sarin and Heng, Kimkong.** (2023). ChatGPT for Education and Research: A Review of the Benefits and Risks. <http://dx.doi.org/10.2139/ssrn.4378735>
- Starkey, L.** (2020). A review of research exploring teacher preparation for the digital age. *Cambridge Journal of Education*, 50 (1), pp. 37-56.

USING ARTIFICIAL INTELLIGENCE FOR ENHANCED LESSON PLANNING AND MATERIAL DEVELOPMENT: TRANSFORMING TEACHING PRACTICES

Cette étude visait à démontrer comment les enseignants en formation utilisent les outils basés sur l'intelligence artificielle (IA) dans leurs pratiques pédagogiques. L'étude a porté sur soixante-huit enseignants d'anglais en formation initiale, en troisième année au département d'enseignement des langues étrangères d'une université d'État. L'analyse s'est concentrée sur les outils d'intelligence artificielle utilisés par les participants dans leurs plans de cours lors du micro-enseignement dans le cadre du cours obligatoire « Teaching Language Skills » (enseignement des compétences linguistiques). Les résultats indiquent que l'outil d'IA a aidé les participants à générer du contenu, notamment en leur permettant de lancer des idées, de développer du matériel et des activités attrayants et de gagner du temps pour créer et simplifier le matériel. Cependant, plusieurs questions ont été soulevées concernant la qualité du contenu et l'éthique.

● Ferit Kılıçkaya
| Burdur Mehmet
Akif Ersoy University,
Türkiye



Ferit Kılıçkaya is a Professor of English Language Teaching (ELT) in the Department of Foreign Language

Education at Burdur Mehmet Akif Ersoy University, Türkiye. His research primarily explores the benefits and challenges of integrating technology into language teaching and learning, with a particular focus on its implications for language assessment.

Introduction

The rapid developments and the emergence of new tools in artificial intelligence (AI) have begun to reshape various sectors, especially education. AI tools are integrated into educational practices and support administrative tasks, enhancing classroom engagement and language learning practices (Ahn et al., 2024; Bonner et al., 2023; Kılıçkaya, 2023). There has also been a growing public interest and increase in AI use in teaching and learning (Cardona et al., 2023; Walton Family Foundation, 2024). In addition to language learning, AI has the potential in language teaching to help teachers design lessons based on the needs and levels of the learners. This short article explores how pre-service language teachers used AI tools as a source of material development in their lesson planning in a teaching methodology course.

AI integration

There is a plethora of research on AI in education, which investigates how AI tools can support teaching and learning, such as producing texts regarding lesson planning and assessment (e.g., Ahn et al., 2024; Kic-Drgas & Kılıçkaya, 2024). The expansion and the rapid emergence of AI tools with various features have led to integrating these tools into language teaching and learning in various aspects, such as writing, chatbots, and machine translation (e.g., Godwin-Jones et al., 2024; Hockly, 2024). It is not an exaggeration to say that there is now an AI-based tool for almost any teacher and learner-related task in the field, from creating images to lesson planning and generation (e.g., Kılıçkaya, 2023). In line with the expansion of AI and AI tools, there is a growing body of research on the potential and uses of AI that can help teachers with tasks such as creating materials that will allow them to focus more on instructional design and student engagement (Dizon, 2024). For example,

It is not an exaggeration to say that there is now an AI-based tool for almost any teacher and learner-related task in the field, from creating images to lesson planning and generation.

Hong (2023) discusses how ChatGPT can be used as a tool for learning and to support learning by practicing the four skills in a language and support teaching, such as creating lesson plans and exercises. In addition to these promising benefits, it is also highlighted that the drawbacks of using AI tools such as ChatGPT include plagiarism and the lack of accuracy of the outputs (Kohnke et al., 2023; Sparks et al. (2024).

The current study focuses on how pre-service language teachers benefited from AI tools in their lesson plans for micro-teaching in a required skills course at a state university.

Methodology

The research design of the study consisted of a content analysis of the lesson plans as well as reflective journals and peer feedback created by the participants of the Teaching Language Skills II course for pre-service language teachers, who were enrolled in the Department of Foreign Language Education at a state university in Türkiye. In this course, the participants were encouraged, but not forced, to use AI tools to create materials as input in their lesson plans, which they would use for micro-teaching.

Participants

The participants of the study were 68 pre-service English teachers in their third year at the Department of Foreign Language Education. Of the participants, 29 were male, while the rest were female. The participants' ages were between 21 and 25.

Data collection

At the beginning of the semester, the participants were introduced to the course, during which they would discuss various ways to teach grammar, reading, and writing and practice these language components and skills via micro-teaching sessions. The participants were also asked to form groups of 4 (or 3, exceptionally, subject to the lecturer's approval) to prepare 35-minute microteaching in the skills and language components presented and discussed in the course. The participants were also required to submit the revised and improved version of their lesson plan during the final exam based on the feedback provided by their

classmates and the lecturer. As for the policy on AI use, the researcher informed the participants that their use of AI tools (e.g., ChatGPT, Magic School and Twee) was permitted and encouraged, but not mandatory, in their lesson planning for the following activities: Generating ideas, editing, and improving ideas; drafting outlines to organize their thoughts; preparing and simplifying texts for classroom activities; creating audio-visual aids; grammar and style checking. The participants were also asked to indicate their work and the output created by any AI tool.

The microteaching sessions aimed to provide the participants with a practical setting for implementing lesson plans, allowing both the lecturer and the participants to observe the effectiveness of their lesson plans and the tools they used in practice. Following these sessions, students were asked to submit reflective journals detailing their own experiences with lesson planning and AI as well as peer feedback, which provided qualitative data on the perceived benefits and limitations of AI in lesson planning.

Data analysis

The data, consisting of the lesson plans, peer feedback forms based on the microteaching and reflective journals, were subject to content analysis in order to investigate the participants' use of AI tools in lesson planning and the potential benefits and challenge. The data was manually classified according to the codes and the categories by following the method by Saldaña (2009).

Benefits regarding content creation uses made	Benefits regarding efficiency and time management	Challenges and limitations
<ul style="list-style-type: none"> – Brainstorming of lesson activities: <ul style="list-style-type: none"> • Role play activities for EFL classrooms • Culturally relevant reading materials • as leveled texts and resources (e.g., Diffit) – AI tools (e.g., Dall-E, Suno) were used to create interactive visuals and multimedia resources. – AI-based routine tasks (e.g., grammar and language check, formatting). 	<ul style="list-style-type: none"> – Easy to use – Saving time and effort to allow for reflection on pedagogical aspects 	<ul style="list-style-type: none"> – Quality of Output: AI-generated content must often be checked for accuracy and content. – Cost: Some advanced features require a paid (expensive) version. – Ethical Concerns: Issues of originality and potential plagiarism regarding AI-generated content. – Lack of clear guidelines and sufficient support regarding the use of AI, academic integrity

Table 1

Summary of the outcomes of the content analysis

Findings and Discussion

The content analysis of the lesson plans, the feedback and the reflective journals produced during the *Teaching Language Skills* course yielded two key categories of benefits as well as a set of potential challenges of the use of AI tools used for lesson planning (Table 1).

The table summarizes the outcome of the content analysis deploying three evaluation criteria: 1. Benefits regarding content creation: uses made, 2. Benefits regarding efficiency and time management, and 3. Challenges and limitations associated with integrating artificial intelligence (AI) tools in lesson plans. The use of tools such as ChatGPT exemplified how AI facilitated brainstorming and content creation by enriching lesson plans and educational resources, e.g., role-play scenarios. It was also observed that the role of AI in content creation extended beyond text production. Tools such as DALL-E and Suno enabled the production of interactive visuals and multimedia resources such as songs for learner engagement. Most participants used AI tools like Diffit to benefit from differentiated instruction by providing leveled texts and resources for learners' varying proficiency levels. In addition to content creation, another significant benefit was the efficiency and time management afforded by AI tools. The user-friendly nature of these applications simplified routine educational tasks, such as grammar checking and language improvement. By automating these basic,

yet time-consuming tasks and creating exercises, AI saved participants time and effort.

However, adopting AI in these lesson plans also had its challenges and limitations. One of the primary concerns was the need for the participants to verify the appropriateness and quality of AI-generated content. While these tools could generate content such as exercises and simplified texts, they did not always guarantee appropriate and quality content, necessitating careful oversight. Furthermore, the cost associated with advanced AI features posed a barrier to access, as some functionalities required the paid versions. Ethical considerations also played a critical role in the discussion surrounding the participants' activities, such as the issues of originality and plagiarism regarding AI-generated content.

The findings indicated that the most notable benefit of integrating AI in lesson plans was to generate ideas and materials for activities. Participants frequently reported that AI tools, particularly ChatGPT, helped them create lesson activities and exercises. One example included the AI-assisted development of role-play activities for micro-teaching, which simulated real-life communication challenges in EFL classrooms, enabling students to practice their speaking and problem-solving skills within a controlled environment. The following prompt was used to create a role-play:

Create a role-play scenario for A2-level English learners. The scenario should involve one learner playing a lost traveler at an airport who needs to ask for directions and another playing an airport staff member who gives simple directions and asks for information about the flight. The dialogue should include polite questions, basic directions (e.g., 'turn left,' 'go straight'), and simple responses suitable for A2-level learners." ChatGPT produced the following output:

Role-play Example:

"At the Airport – Asking for Directions"

Scenario:

You are at the airport and cannot find the gate for your flight. You need to

ask an airport staff member for directions. The staff member will give directions and ask questions about your flight.

Roles:

1. Student A (Traveler): You are lost in the airport and need to find your gate. You speak basic English and need help from the airport staff.
2. Student B (Airport Staff): You work at the airport information desk. You need to ask travelers for flight details (e.g., gate number and flight time) and give them simple directions.

Instructions:

- Student A should ask polite questions like: “Excuse me, where is Gate 12?” or “Can you help me find my flight to London?”
- Student B should respond with clear, simple directions: “Go straight and turn left.” They should also ask for more information: “What time is your flight?” or “Which airline are you flying with?”

Similarly, AI was used to create linguistically relevant reading materials tailored to students’ interests, backgrounds, and levels in the lesson plan. One participant in his reflective journal indicated that he could level the linguistically complex text to the learners’ level (A2 as indicated in the lesson plan) and added that by using the AI tool (Diffit), he and his group members quickly adapted the text of the reading passage that they had taken from a coursebook and customized questions and vocabulary by checking the vocabulary and grammar.

The journals and lesson plans showed that the participants employed tools like Dall-E, Tweek, and MagicSchool to generate input for learners, such as interactive visuals and quizzes, to be incorporated into lesson plans. The participants believed that this use of AI could enrich their lesson content. In one of the lesson plans, for example, the participants used AI-generated images to present various grammar rules to offer concrete representations of abstract concepts, which would otherwise not have been found on the internet or would have required much time. These examples and specific uses of AI in lesson plans and reflective journals underscore the role of AI in producing materials for language teaching

and customizing these materials for different levels. However, teachers’ expertise was considered important in checking the content, and these tools cannot replace the critical posture of teachers.

Using AI tools in lesson planning also provided several advantages in checking the content for grammar and language enhancement. In this way, participants stated that they obtained more time for pedagogical reflection and refinement of their lesson plans, such as selecting appropriate teaching strategies and exercises. This was found particularly valuable given the time constraints of the course. The following prompt was used to create a vocabulary activity:

Create a matching activity for A1-level learners in an EFL classroom. On the left side, list personality adjectives such as “punctual,” “angry,” and “helpful.” On the right side, provide clear and straightforward definitions for each adjective. Make sure the definitions are accessible to A1-level learners. The task should challenge students to match each adjective with its correct definition. Ensure that the definitions are distinct enough to avoid confusion and that the language aligns with their proficiency level. Include about 8-10 adjectives in total.

Based on this prompt, the following exercise was created.

Matching Activity:

Instructions: Match each word on the left with its correct definition on the right. Draw a line to connect the word and its meaning.

Adjectives	Definitions
1. Punctual	a) Nice and easy to talk to
2. Angry	b) Feels nervous when talking to new people
3. Helpful	c) Not afraid, ready to do dangerous things
4. Kind	d) Makes people laugh
5. Lazy	e) Does not want to work or do anything
6. Funny	f) Nice and caring
7. Quiet	g) Likes to help others
8. Friendly	h) Feeling very mad
9. Brave	i) Doesn't talk much
10. Shy	j) Always on time

Despite these benefits, using AI output in lesson planning also presented several challenges, as indicated by the participants. One limitation was the quality of the output generated by AI tools, which, according to the participants, often lacked depth and required content checking. This issue highlighted that AI tools could help the participants in materials design in various ways, such as text creation and exercise; however, it would not replace the need for teachers' critical thinking and expertise to check the appropriateness of the content.

The participants also voiced ethical concerns regarding the originality of the output generated by AI tools and potential plagiarism, as AI tools might benefit from the information available on the Internet. They were concerned about the authorship and originality of the created activities and lesson plans. These considerations emphasized the need for clear guidelines and policies set by the institutions on the responsible use of AI in educational settings. They underscored the importance of maintaining academic integrity.

Conclusion and suggestions

This study may indicate the potential to revolutionize lesson planning by offering pre-service teachers innovative tools to enhance their instructional practices. **As suggested in the current study, AI can support teachers in generating creative ideas, developing engaging materials, and saving time on creating and simplifying materials. However, integrating AI in lesson planning and materials creation also requires careful consideration of content as well as of ethical aspects.** AI-generated content must be carefully checked and critically assessed, which may require that both learners and teachers receive the necessary support (cfr. e.g., Baradel; Alloatti & Montemarano; Cotelli Kureth et al.; Ferris et al., this issue). Further research can be conducted on how access to AI-based tools (or the lack thereof) might affect the digital divide among learners. Although there are AI-based tools that offer free access to certain features, in the same classroom some learners might have access to paid versions with full features, while others may only have limited access.

Declaration of the use of generative AI

The manuscript has been proofread for language issues as well as the organization of the paper using Grammarly and ChatGPT.

However, teachers' expertise was considered important in checking the content, and these tools cannot replace the critical posture of teachers.

References

- Ahn, J., Lee, J., & Son, M.** (2024). *ChatGPT in ELT: disruptor? Or well-trained teaching assistant?* *ELT Journal*, 78(3), 345–355. <https://doi.org/10.1093/elt/ccae017>
- Bonner, E., Lege, R., & Frazier, E.** (2023). Large language model-based artificial intelligence in the language classroom: Practical ideas for teaching. *Teaching English with Technology*, 23(1), 23–41. <https://doi.org/10.56297/BKAM1691/WIE01749>
- Cardona, M. A., Rodríguez, R. J., & Ishmael, K.** (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. U.S. Department of Education, Office of Educational Technology. <https://www2.ed.gov/documents/ai-report/ai-report.pdf>
- Dizon, G.** (2024). ChatGPT as a tool for self-directed foreign language learning. *Innovation in Language Learning and Teaching*. <https://doi.org/10.1080/17501229.2024.2413406>
- Godwin-Jones, R., O'Neill, E., & Ranalli, J.** (2024). Integrating AI tools into instructed second language acquisition. In C. A. Chapelle, G. H. Beckett, & J. Ranalli (Eds.), *Exploring artificial intelligence in applied linguistics* (9–23). Iowa State University Digital Press. <https://www.iastatedigitalpress.com/plugins/books/154/>
- Hockly, N.** (2024). *Nicky Hockly's 30 essentials for using artificial intelligence*. Cambridge University Press. <https://doi.org/10.1017/9781009804509>
- Hong, W. C. H.** (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1), 37–45. <https://jeti.thewsu.org/index.php/cieti/article/view/103/64>
- Kılıçkaya, F.** (2023, December). Navigating the future: A list of AI tools for language teaching and learning. *APACALL Newsletter*, 27, 8–10. https://www.apacall.org/news/APACALL_Newsletter27_2023.pdf
- Kic-Drgas, J., & Kılıçkaya, F.** (2024). Using artificial intelligence (AI) to create language exam questions: A case study. *XLinguae*, 17(1), 20–33. <https://doi.org/10.18355/XL.2024.1701.02>
- Kohnke, L., Moorhouse, B. L., & Zou, D.** (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Saldaña, J.** (2009). *The coding manual for qualitative researchers*. Sage.
- Sparks, J. R., Ober, T. M., Tenison, C., Arslan, B., Roll, I., Deane, P., Zapata-Rivera, Z., Gooch, R. M., & O'Reilly, T.** (2024). *Opportunities and challenges for assessing digital and AI literacies*. ETS. <https://www.ets.org/content/dam/ets-org/pdfs/rd/highlights-measuring-digital-literacies-in-the-age-of-ai.pdf>
- Walton Family Foundation.** (2024). *AI chatbots in schools: Findings from a Poll of K-12 teachers, students, parents, and colleague undergraduates*. <https://www.waltonfamilyfoundation.org/learning/the-value-of-ai-in-todays-classrooms>

COMMENT DÉVELOPPER LA LITTÉRACIE DIGITALE DES ENSEIGNANT-ES ET DES APPRENANT-ES ? CONCLUSIONS DU PROJET 'DIGITAL LITERACY IN UNIVERSITY CONTEXTS'

This article shares the results of the project 'Digital Literacy in University Contexts' which aims to develop the digital literacy of teachers and students at Swiss institutions of higher education. These conclusions illustrate an AI literacy model (Cardon et al. 2023) and allow us to suggest the following steps to develop, for ourselves and our students, application, authenticity, agency and responsibility in our interactions with GenAI tools. It is important to:

- discuss AI with peers and students
- be aware of and present information on: technical details that have an impact on usage; the weaknesses of human posteditors;
- include tasks with AI in the classroom.

● Sara Cotelli Kureth
| UniNE (1)
Alice Delorme Benites
| ZHAW (2)
Caroline Lehr
| ZHAW (3)
Hasti Noghrechi
| UniNE (4)
Elizabeth Steele
| BFH (5)
Elana Summers
| ZHAW (6)

1. Introduction

Cet article présente les principaux résultats du projet swissuniversities PgB-8 'Digital Literacy in University Contexts' (2021-2024)¹ dont l'objectif est de développer la littéracie digitale dans les hautes écoles suisses grâce à diverses actions (enquête; interventions et ateliers; création de matériel pédagogique; recherches-actions). Initialement, nous avons travaillé sur la littéracie en traduction automatique (TA) (Machine translation literacy, voir Bowker & Buitrago-Ciro, 2019) et, dès 2023, nous avons élargi nos formations aux outils d'IA générative (IAGen) tels que les chatbots (ChatGPT, Gemini, Copilot). En effet, les versions

neuronales des outils de TA (par ex. DeepL) sont les précurseurs de l'IAGen et on dispose d'une abondante littérature scientifique sur les effets des textes produits par ces systèmes. L'inclusion de la TA en classe de langue a également fait l'objet de nombreuses expériences (Klimova et al., 2023) qui valident leur intérêt pédagogique. La TA neuronale utilise des technologies similaires aux autres outils d'IAGen, reposant également sur les grands modèles de langage (Large Language Models, LLMs) et l'architecture d'apprentissage des transformers (Benites et al., 2023). Ainsi, les outils d'IAGen comme de TA présentent des biais et des erreurs, ou hallucinations, et les premières expériences faites sur ChatGPT-3 tendent à montrer que le texte produit par ces systèmes présente le même manque de richesse syntaxique et lexicale, parfois nommé *machine translationese*, etc. que la TA (Kenny, 2022). Il est ainsi possible d'appliquer certaines des conclusions de notre projet et d'ouvrir la définition de la littératie en TA aux outils d'IAGen pour la création de texte. Il est même très

¹ Leading house Zürcher Hochschule für Angewandte Wissenschaften (ZHAW) et 3 hautes écoles partenaires: Université de Neuchâtel (UniNE), Berner Fachhochschule (BFH) et Pädagogische Hochschule Zürich (PHZ). Le projet est cofinancé par swissuniversities et les hautes écoles partenaires.

intéressant d'aborder les IAGen dans l'enseignement/apprentissage des langues au travers des leçons acquises sur les outils de TA, ce point de vue étant peu présent dans l'abondante littérature sur l'IA parue ces deux dernières années. C'est ce que nous proposons dans cet article après une brève définition de la littéracie en IA.

2. Littéracie en IA

Comme la littéracie en TA, la littéracie en IA (AI literacy) vise à promouvoir une utilisation critique, durable et raisonnée de ces outils. Plusieurs modèles et définitions existent, mais tous se réfèrent à un ensemble de connaissances, de compétences et d'expériences sur différentes facettes de l'IA (par ex. technologie, implications éthiques, etc.) qui permettent *in fine* un usage ciblé et une interaction pertinente avec l'IA (Pinski & Benlian, 2024: 9). Vu l'utilisation courante de l'IA-Gen en classe de langues pour la production de texte, nous mentionnons ici le modèle développé par Cardon et al. (2023) de littéracie en IAGen pour la communication, qui identifie quatre compétences majeures: l'**application** (connaissances technologiques et alignement de l'outil à la tâche), l'**authenticité** (communication authentique prenant comme standard la communication humaine), l'**agentivité** (maintien du pouvoir de décision et des compétences sans la machine) et la **responsabilité** (prise en charge du contenu et – nous ajoutons ce descripteur – utilisation transparente avec mention des apports de l'IA).

3. Différentes actions du projet

Pour mieux comprendre l'utilisation de la TA dans les hautes écoles suisses, nous avons mené une grande enquête entre 2021 et 2022 auprès des corps étudiant·in, enseignant et administratif. L'analyse de cette enquête (Delorme Benites et al., 2021), ainsi qu'une revue de la littérature autour de l'IAGen pour l'enseignement/apprentissage des langues, ont servi de base au développement d'ateliers pour les apprenant·es et pour les enseignant·es de langues, ainsi que d'autres plus ciblés pour le personnel enseignant et administratif et les doctorant·es. Ces ateliers, menés dans plusieurs hautes écoles suisses entre 2022 et 2024, ont été évalués par un questionnaire. Une séance de partage d'expériences entre formatrices et

Nous estimons en outre qu'il est crucial d'encourager les enseignant·es à discuter des outils d'IAGen avec leurs étudiant·es.

formateurs a également été organisée, mettant en lumière les points forts de ces interventions. Nous les présentons ci-dessous.

En parallèle des interventions, nous avons conduit plusieurs recherches-actions avec des apprenant·es de langue afin de tester des séquences didactiques (Cotelli Kureth & Noghrechi, 2024) et des protocoles visant à développer la littéracie en TA (Cotelli Kureth & Summers, 2023). Nous avons également développé un règlement sur l'utilisation de la TA dans les hautes écoles qui a été mise en place à la Haute école spécialisée bernoise².

Au terme du projet, notre site internet proposera, en libre accès, de courtes capsules vidéo en français, anglais et allemand. Elles informeront sur le fonctionnement des outils de TA, les interactions des utilisatrices et des utilisateurs avec la machine et sur les questions éthiques. Nous préparons aussi des exemples de séquences didactiques testées en classe pour intégrer les outils d'IAGen dans les cours: <https://www.zhaw.ch/en/linguistics/research/digital-literacy-in-university-contexts/>.

4. Développer la littéracie en IA: leçons de la TA

Nous revenons ici brièvement sur les points forts relevés dans notre enquête, les entretiens et les recherches-actions développés dans le cadre de ce projet, en les reliant aux compétences du modèle de Cardon et al. (2023), précisant et complétant ainsi celui-ci.

² Voici le lien vers la politique en français: https://www.bfh.ch/dam/jcr:121ab3db-069e-46a0-8a0b-5381e897043d/Policy%20zur%20Verwendung%20von%20maschinellem%20%C3%9Cbersetzung%20im%20Hochschulkontext_fr.pdf Elle existe également en allemand et en anglais.

Les 6 auteures ont collaboré lors du projet 'Digital Literacy in University Contexts'. Elles travaillent: (1) au Centre de langues de l'UniNE, comme directrice et (4) à l'ILCF comme chargée d'enseignement; à l'Institut für Übersetzen und Dolmetschen de la ZHAW, comme (2) directrice, (3) professeure et (6) doctorante; (5) à la BFH, comme responsable du Service spécialisé du bilinguisme et plurilinguisme.

La bonne facture syntaxique, grammaticale et orthographique du texte pousse la relectrice et le relecteur à faire confiance au message transmis, même si son sens commence à différer de celui du message d'origine.

4.1 Agentivité : s'exprimer sur ses peurs, ses doutes et ses émotions

Un des premiers points que nous retenons de notre expérience de trois ans d'ateliers est l'importance de la discussion. L'arrivée massive des outils d'IAGen dans les habitudes des apprenant-es a créé un fort sentiment d'insécurité auprès des enseignant-es qui se sentent souvent démunis-es. Les gros titres sensationnalistes annonçant régulièrement la fin des cours de langues (par ex. Benini, 2024) renforcent ce sentiment d'insécurité. Nos interventions ont souvent consisté à offrir aux enseignant-es un espace sûr pour exprimer leurs craintes, leurs doutes, leur enthousiasme et leur incompréhension. Nous avons toujours adopté une perspective positive sur les outils, tout en les désacralisant et en exposant leurs faiblesses (voir introduction). Nous avons parfois comparé ces nouvelles technologies à d'autres qui ont révolutionné la société (pensons à la calculatrice, à l'Internet ou aux MOOCs) et changé en profondeur l'enseignement sans pour autant mettre les enseignant-es au chômage. Cette approche a contribué à atténuer les craintes des participant-es en renforçant leur agentivité. Vu le manque de communication au sujet de la TA constaté dans notre enquête, nous estimons en outre qu'il est crucial d'encourager les enseignant-es à discuter des outils d'IAGen avec leurs étudiant-es. Au niveau universitaire, ces deux publics partagent des préoccupations et des émotions similaires face aux nouveaux outils d'IAGen (Lehr & Cotelli Kureth, 2024), bien que les enseignant-es mentionnent plus souvent leurs peurs devant le plagiat et la tricherie.

4.2. Application : informer sur la technologie pour éviter les usages peu judicieux

Suite aux résultats de notre enquête, nous avons constaté certains usages problématiques de la TA, notamment l'insertion

d'un mot unique. Comme ces machines fonctionnent – jusqu'à présent – à partir du contexte d'une phrase, elles ne produisent pas un résultat optimal avec des informations si peu contextualisées. Les usagères et usagers pensent souvent l'utilisation des outils de TA en termes de mots et non de phrases ou de textes, comme si rédiger équivalait à additionner des mots (Cotelli Kureth et al., 2023). En conséquence, en plus d'expliquer certains éléments technologiques, il est également crucial pour les enseignant-es de mieux décrire et entraîner les différentes étapes du processus d'écriture pour que les apprenant-es soient bien équipés-es à rédiger dans la L1 et la L2, avec ou sans outils d'IAGen.

Les outils d'IAGen reposent sur des technologies similaires à la TA et il est important que les utilisatrices et utilisateurs les connaissent pour savoir qu'il faut fournir à la machine un contexte suffisant pour qu'elle fonctionne de manière optimale. C'est le cas pour les générateurs de texte comme ChatGPT qui ont besoin d'un *prompt* détaillé pour produire un bon résultat. Souvent une ligne ne suffit pas et créer un *prompt* efficace demande du temps et de la réflexion.

4.3 Agentivité et authenticité : les faiblesses des utilisatrices et utilisateurs humain-es

La plupart des présentations destinées aux utilisatrices et utilisateurs de l'IAGen insistent sur les faiblesses de ces outils qui reposent sur des corpus gigantesques mais qui ne sont pas 'propres' et sur des algorithmes statistiques qui entraînent des problèmes (*machine translationese*, biais, etc.). Il est bien sûr très important pour l'utilisatrice et l'utilisateur d'avoir conscience de celles-ci mais dans notre recherche sur l'IAGen à partir de la TA, nous avons appris qu'il était tout aussi important de connaître le facteur

humain, un discours aujourd'hui encore peu présent et que nous avons donc placé au cœur de la seconde capsule vidéo du projet.

Citons d'abord l'exemple déjà présenté de l'utilisation sous-optimale de la TA comme simple dictionnaire qui souligne le minimalisme dont font parfois preuve les utilisatrices et les utilisateurs d'outils d'IAGen : une première faiblesse des relectrices et relecteurs humain-es. Notre enquête a aussi révélé l'importance qu'ils et elles donnent à la post-édition (relecture attentive et critique du texte produit par la machine) au détriment de la pré-édition (relecture attentive et critique du texte que l'on fournit à la machine) : seuls un quart des répondant-es pré-éditent, alors que les trois quarts post-éditent³. Ceci pose problème, car les études sur la post-édition de textes traduits par la TA montrent la difficulté pour les relectrices et relecteurs, même expérimentés-es, à repérer les fautes éventuelles (Loock & Léchauguette, 2021), également en raison d'une charge cognitive plus élevée (Kasper et al., 2023). On parle ainsi de la « fausse aisance » des textes générés par la TA (Martindale & Carpuat, 2018) : la bonne facture syntaxique, grammaticale et orthographique du texte pousse la relectrice et le relecteur à faire confiance au message transmis, même si son sens commence à différer de celui du message d'origine. On peut postuler qu'il en est de même avec les textes produits par ChatGPT, qu'il s'agisse de textes réécrits à partir d'originaux, de textes 'corrigés' ou entièrement rédigés par la machine⁴. Connaître ses propres faiblesses en tant que relectrices et relecteurs humain-es nous permet ainsi de développer des stratégies pour mitiger les risques posés par la fausse aisance et conserver le pouvoir de décision sur le contenu.

³ Pour la pré-édition : À la question « est-ce que vous faites des modifications sur le texte original avant de le mettre dans un système de TA ? », les réponses sont : non 74 %, oui 21 %, pas de réponse 5 % (n=2150). À la question « est-ce que vous faites des modifications sur le texte après avoir utilisé un système de TA ? », les réponses sont : non 18 %, oui 77 %, pas de réponse 5 % (n=2150).

⁴ On en voit des prémises dans l'expérience menée par Fyfe (2023).

4.4 Agentivité et responsabilité : connaître et expérimenter

Finalement, des recherches-actions menées au Centre de langues de l'Université de Neuchâtel ont montré une amélioration manifeste de la littéracie en TA des apprenant-es à la suite d'une intervention reprenant les points évoqués ci-dessus et quelques notions d'éthique (Cotelli Kureth & Summers, 2023). La présentation et l'introduction durant le semestre d'une tâche avec un outil d'IAGen en parallèle a permis aux participant-es d'encore mieux développer cette littéracie. Une dernière expérience (Cotelli Kureth et al., en préparation) nous a en outre permis de constater qu'il était important de faire confiance aux étudiant-es. Dans le cadre de l'évaluation des apprentissages, nous leur avons demandé de rédiger un texte à la maison en utilisant un code couleur pour marquer leur utilisation des outils (ChatGPT, TA, corpus bilingues et dictionnaires) et en détaillant leur processus d'écriture. La grille critériée d'évaluation comprenait une partie sur la littéracie en IA. Un questionnaire a ensuite permis de tester le degré de transparence avec laquelle la tâche avait été effectuée et a montré un fort engagement de la part des étudiant-es. De telles expérimentations sont cruciales pour que les apprenant-es développent une responsabilité face aux outils d'IA. Nous confirmons ainsi l'intuition de Pinski & Belian (2024) qui incluent l'expérience comme constitutive de la littéracie d'IAGen, au même titre que les compétences et les connaissances.

5. Conclusion

Notre projet n'est bien sûr qu'une première étape dans le développement d'une littéracie de l'IA auprès du personnel enseignant et étudiant des hautes écoles suisses. Nous retenons l'importance que revêtent la discussion, l'expérimentation et la transmission des bonnes informations, toujours étayées scientifiquement, pour développer l'application, l'authenticité, l'agentivité et la responsabilité dans nos interactions avec les outils d'IAGen et invitons les enseignant-es de langues à utiliser, parmi d'autres ressources disponibles, le matériel didactique (vidéos et tâches) que nous avons développé.

Références

- Benini, S.** (2024, 11 mars). Künstliche Intelligenz und Sprache: Fremdsprachen lernen? Braucht es bald nicht mehr! *Tages Anzeiger*. URL: <https://www.tagesanzeiger.ch/kuenstliche-intelligenz-fremdsprachen-lernen-ist-bald-passe-598312448467>
- Benites, F., Delorme Benites, A. & Anson, C. M.** (2023). Automated Text Generation and Summarization for Academic Writing. In: O. Kruse et al. (éds.), *Digital Writing Technologies in Higher Education. Theory, Research, and Practice*. Springer, pp. 279-301.
- Bowker, L. & Buitrago Ciro, J.** (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing.
- Cardon, P., Fleischmann, C., Artiz, J., Logemann, M. & Heidewald, J.** (2023). The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. *Business and Professional Communication Quarterly*, 86(3).
- Cotelli Kureth, S., Delorme Benites, A., H. Noghrechi, M. Haller & E. Steele** (2023). 'I looked it up in DeepL': Machine translation and digital tools. In: N. Froeliger et al. (éds.), *Human Translation and Natural Language Processing: Towards a New Consensus?* Edizioni Ca' Foscari, pp. 81-96.
- Cotelli Kureth, S. & Noghrechi, H.** (2024). Intégrer GoogleTranslate comme support d'enseignement au niveau débutant : Atelier pratique pour un public de réfugiés ukrainiennes. *Partages*, 1. <https://dx.doi.org/10.35562/partages.98>
- Cotelli Kureth, S., Paliot, E. & Zink, S.** (en préparation). Fostering transparency : A critical introduction of generative AI in students' assignments.
- Cotelli Kureth, S. & Summers, E.** (2023). Tackling the elephant in the language classroom: Introducing machine translation literacy in a Swiss language centre. *Language Learning in Higher Education*, 13(1), 213-230.
- Fyfe, P.** (2023). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*, 38, 1395-1405.
- Kasperé, R., Motiejūnienė, J., Patšienė, I., Patašius, M. & Horbačauskienė** (2023). Is machine translation a dim technology for its users? An eye tracking study. *Frontiers in Psychology*, 14, 1076379.
- Kenny, D.** (éd.) (2022). *Machine translation for everyone: Empowering users in the age of artificial intelligence*. Language Science Press.
- Klimova, B., Pikhart, M., Delorme Benites, A., Lehr, C. & Sanchez-Stockhammer, C.** (2023). Neural machine translation in foreign language teaching and learning: A systematic review. *Education and Information Technologies*, 28(1), 663-682.
- Lehr, C. & Cotelli Kureth, S.** (2024, avril). Intérêt, mépris, peur, soulagement et contentement : la palette des émotions élicitées par les outils d'IA générative. Présentation à la journée *Perspectives didactiques pour la traduction automatique dans l'apprentissage des langues*, Université de Lorraine, Nancy, le 5 avril 2024.
- Loock, R. & Léchauguet, S.** (2021). Machine translation literacy and undergraduate students in applied languages: Report on an exploratory study. *Revista Tradumàtica: tecnologies de la traducció*, 19, 204-225.
- Martindale, M., & Carpuat, M.** (2018). Fluency over adequacy: A pilot study in measuring user trust in imperfect MT. In *Proceedings of AMTA 2018*, vol. 1, pp. 13-25.
- Pinski, M, Benlian, AL.** (2024). AI literacy for users - A comprehensive review and future research directions of learning methods, components and effects. *Computers in Human Behavior: Artificial Humans*, 2, 100062.

ENGLISH QUEST : EXPLORATION DU POTENTIEL PÉDAGOGIQUE DES IA GÉNÉRATIVES DANS L'APPRENTISSAGE DES LANGUES

This article explores *English Quest*, an innovative video game for English language learning. Combining artificial intelligence, *Digital Game-Based Learning*, and *Interactive Fiction*, the game offers immersive interactions with virtual characters. The study examines its pedagogical potential, highlighting its support for stimulating, creative, and authentic learning. While promising, the game has limitations, including a high difficulty level and lack of flexibility. The article concludes by discussing improvement perspectives and *English Quest*'s potential impact on the future of language learning in the context of digital education.

● Guillaume Tschupp
| HEP-US
Erica Borloz
| HEP-US

Dans cet article, nous explorons *English Quest*, un jeu vidéo innovant développé par Gaëtan Marti de l'entreprise Immensive SA à Lausanne¹, conçu pour l'apprentissage de l'anglais à travers des interactions immersives avec des personnages virtuels. Cet article est le fruit d'une collaboration entre Guillaume Tschupp, expert en e-learning et éducation numérique, et Erica Borloz, experte en didactique de l'anglais, enrichie par des discussions régulières avec le créateur du jeu.

L'objectif de cet article est d'examiner le potentiel de *English Quest* en tant que prototype, montrant comment les jeux vidéo et les technologies d'intelligence artificielle (IA) peuvent transformer l'apprentissage des langues. Le jeu s'inscrit dans un cadre théorique solide, combinant les concepts de *Digital Game-Based Learning* (DGBL), où les mécanismes de jeu sont utilisés pour motiver les apprenants (Aguilera & Roock, 2022), et d'*Interactive Fiction* (IF), qui permet aux utilisateurs de s'immerger dans des récits narratifs via des interactions textuelles (Montfort, 2013).

En abordant cette exploration, nous mettons en lumière les avantages pédagogiques de l'intégration de l'IA dans l'enseignement des langues, tout en discutant des pistes d'amélioration pour transformer ce prototype en un outil d'apprentissage complet et adaptable à divers contextes éducatifs.

Présentation de *English Quest*

English Quest est un jeu vidéo conçu pour l'apprentissage de l'anglais à travers des scénarios interactifs et immersifs. Ce qui distingue particulièrement ce jeu, c'est l'utilisation de l'intelligence artificielle (IA) pour créer des interactions textuelles qui s'adaptent en temps réel aux réponses des apprenants. Les joueurs sont immergés dans des quêtes narratives où ils interagissent avec des personnages virtuels (Table 1), chaque interaction étant conçue pour favoriser l'apprentissage de la langue.

Le jeu fonctionne principalement autour de deux types de tâches : des enquêtes complexes qui nécessitent l'utilisation

¹ <https://www.immensive.com/fr/>

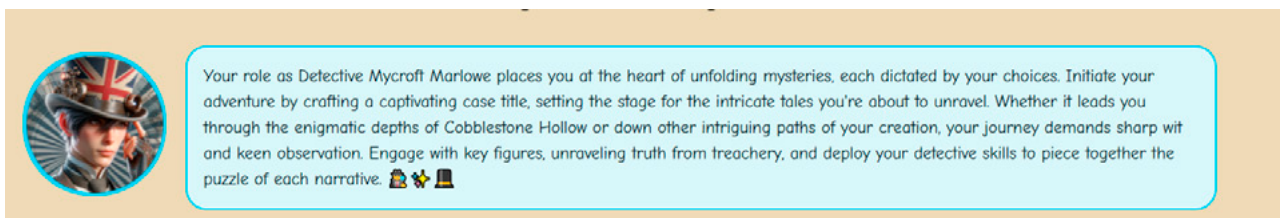


Table 1

Aperçu de l'interface principale où les apprenants interagissent avec les personnages virtuels. Chaque interaction est générée par l'intelligence artificielle, offrant une expérience d'apprentissage personnalisée.

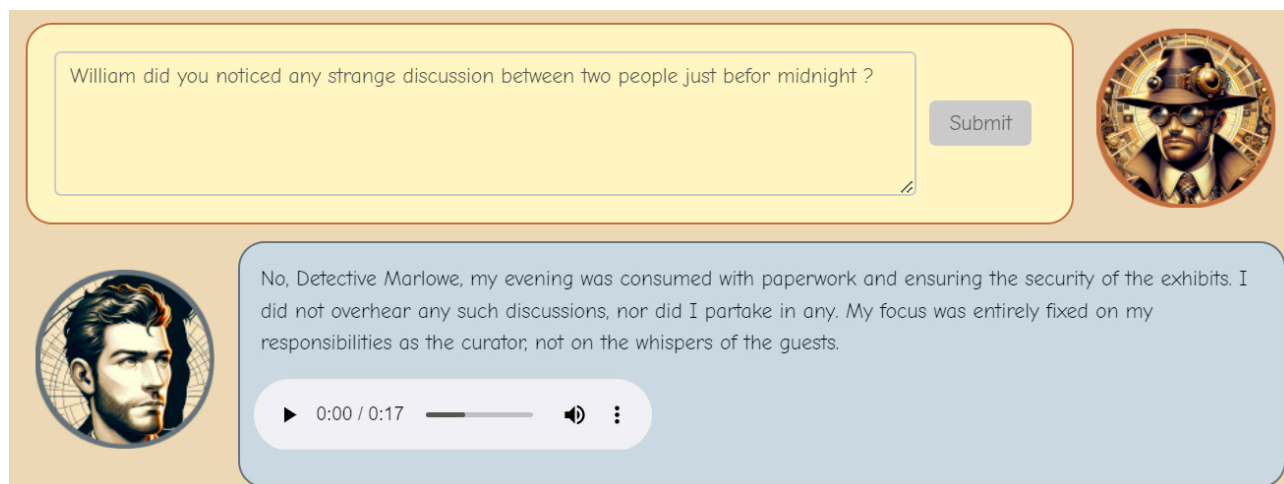


Table 2

Interactions personnalisées
avec texte et audio

active de la langue dans des contextes variés, et des missions annexes centrées sur des objectifs linguistiques spécifiques, tels que lire l'heure, dire les dates, ou converser sur des sujets quotidiens. Ces tâches ne sont pas simplement des exercices linguistiques isolés, mais sont intégrées dans un scénario global qui maintient l'engagement et la motivation des apprenants. La structure narrative du jeu, avec ses missions et objectifs, renforce l'immersion des apprenants, les plaçant dans des situations de communication authentiques.

Ce qui rend *English Quest* unique, c'est que les textes ne sont pas prédéfinis comme dans les supports papier ou les jeux narratifs classiques. Ici, les réponses des personnages sont générées en temps réel par l'IA, en fonction des entrées spécifiques de l'apprenant. Par exemple, lorsqu'un joueur pose une question précise, le personnage répond de manière adaptée tout en intégrant des nuances qui enrichissent l'interaction (Table 2). Dans une séquence du jeu, un personnage affirme quelque chose tout en dissimulant habilement une ruse qui sera révélée plus tard. Cette capacité de l'IA à créer des dialogues

dynamiques et narrativement cohérents en fonction des actions du joueur permet une immersion complète.

Cette personnalisation signifie que chaque interaction est unique et suit l'évolution des productions et des choix de l'apprenant. Même si le joueur commet des erreurs dans ses productions, l'IA est capable d'en comprendre le sens et de répondre de manière adéquate. Cela met l'accent sur la communication et l'interaction, rendant l'expérience très proche d'un échange avec un partenaire humain, comme un comédien improvisant en temps réel dans le cadre d'un scénario. *English Quest* ne se limite donc pas à enseigner une langue; il engage l'apprenant dans une narration interactive et adaptative, offrant une approche véritablement novatrice de l'apprentissage. Ce mécanisme encourage une pratique plus fluide et confiante de l'anglais, sans crainte de faire des erreurs qui bloqueraient la conversation.

En plus des interactions textuelles, le jeu propose également des conversations libres où les apprenants peuvent exprimer des idées personnelles et



Guillaume Tschupp : chargé d'enseignement en éducation numérique à la HEP-VS et responsable du support au e-learning. Il est également membre actif du LEAN et du groupe de R&D Lab C.I.T.E., ainsi que du groupe « A Roadmap for Open Education in Switzerland ».



Erica Borloz : chargée d'enseignement à la HEP-VS, didacticienne d'anglais et coordinatrice des langues étrangères. Elle est également chercheuse au sein du groupe R&D Lab C.I.T.E.

English Quest se distingue par sa capacité à s'adapter à l'apprenant en temps réel et à intégrer des tâches complexes et authentiques dans un format ludique.

émotionnelles. Ces conversations sont structurées de manière à être adaptées au niveau de chaque joueur, permettant ainsi une différenciation pédagogique. Cela permet non seulement de développer des compétences sémantiques et syntaxiques, mais aussi des compétences pragmatiques et sociolinguistiques essentielles à une communication efficace.

Les feedbacks fournis par l'IA sont détaillés et constructifs, soutenant les apprenants dans leur progression. Par exemple, lors de la rédaction d'un message durant une enquête, l'IA analyse de manière complexe la syntaxe utilisée et fournit des recommandations pour améliorer la précision linguistique (Table 3).

Analyse didactique et pédagogique

Du point de vue de la didactique des langues, *English Quest* se distingue par sa capacité à s'adapter à l'apprenant en temps réel et à intégrer des tâches complexes et authentiques dans un format ludique, ce qui répond à plusieurs enjeux pédagogiques essentiels. Selon Nissen (2019), un outil didactique doit s'adapter au niveau des apprenants, soutenir leur motivation, offrir un feedback pertinent, et un scénario pédagogique cohérent. *English Quest* répond à ces critères en proposant des scénarios variés où les apprenants sont invités à résoudre des enquêtes complexes ou à engager des conversations sur des sujets quotidiens, renforçant ainsi un apprentissage situé et contextualisé.

L'intégration de scénarios ouverts et engageants, comme le rappelle González-Lloret (2020), est essentielle pour favoriser l'apprentissage des langues. *English Quest* permet aux apprenants de participer

activement à des récits narratifs immersifs, facilitant le développement de compétences langagières variées, allant de la compréhension à la production. En engageant les apprenants dans des scénarios où les réponses peuvent être multiples et originales, le jeu encourage la prise de risques linguistiques dans un environnement sécurisé (Reinhardt, 2019). Cette approche est particulièrement bénéfique pour les apprenants qui ont besoin de renforcer leur confiance en eux et leur fluidité dans l'utilisation de la langue cible.

Ziegler (2016), dans sa méta-analyse sur la communication médiatisée par ordinateur, souligne l'importance des interactions synchrones pour l'acquisition d'une langue seconde. *English Quest* intègre cette dimension en proposant des conversations en temps réel, où les apprenants peuvent pratiquer la langue de manière spontanée et authentique. Ces interactions immédiates ne se contentent pas d'améliorer la fluidité, elles contribuent également à l'acquisition de compétences pragmatiques et sociolinguistiques, qui sont essentielles pour une communication efficace dans des contextes réels.

Les tâches proposées dans le jeu sont conçues pour être flexibles et adaptatives, ce qui permet une différenciation en fonction du niveau de langue des apprenants. Par exemple, bien que le niveau de compétence requis pour entrer dans le jeu soit plutôt élevé (B2/C1 selon le Cadre Européen Commun de Référence pour les Langues, 2018, désormais CECR), les interactions peuvent être ajustées en termes de longueur et de complexité, selon les capacités individuelles des apprenants. Cela permet de stimuler une expression personnelle et émotionnelle, notamment dans les conversations libres. Ces échanges permettent aux apprenants de s'appropriier la langue de manière plus personnelle, renforçant ainsi leur engagement et leur motivation.

Le feedback, un élément crucial pour le suivi des apprentissages, est particulièrement bien intégré dans *English Quest*. Les retours fournis par l'IA sont non seulement détaillés, mais également formulés de manière positive et encourageante. Par exemple, lors de la rédaction d'un message ou de la réalisation d'une autre activité linguistique, l'IA analyse les aspects syntaxiques avec précision et offre

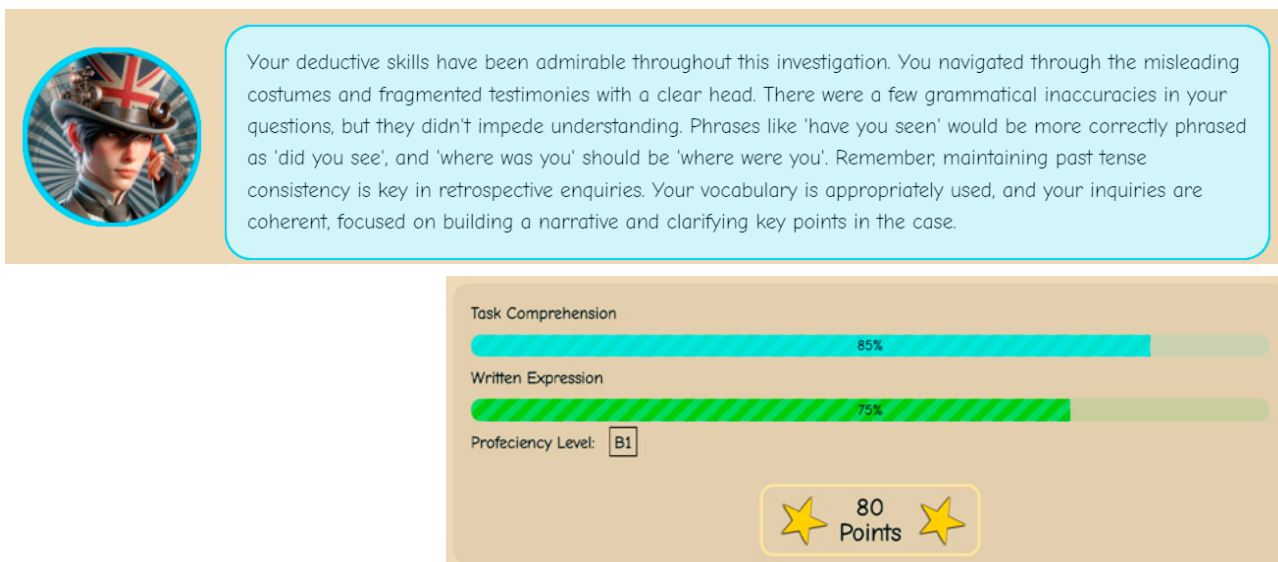


Table 3
Evaluation de la progression des apprentissages

Thank you. So I guess we could say sustainability means that we try to protect the way we use the planet's treasures.

Yes, that's an excellent way to describe sustainability, focusing on the responsible usage of our planet's resources to ensure they're available for future generations. Keep practicing, and you'll find your vocabulary and pronunciation will improve over time. Well done today, and I hope you feel encouraged to continue learning. Have a splendid day!

0:00 / 0:21

Well done! You're making good progress with your pronunciation. The word 'sustainability' is indeed challenging, but your effort is paying off. Remember to take it slow and break down difficult words into smaller parts. Your definition of sustainability is also on the right track. The concept does involve using the planet's resources in a way that doesn't deplete them for future generations. Keep practicing, and you'll continue to improve both your vocabulary and pronunciation.

Table 4
Exemple de feedback personnalisé
fourni par l'intelligence artificielle
dans *English Quest*, suggérant des
améliorations pour renforcer les
compétences linguistiques en
anglais, comme la correction d'une
structure grammaticale ou l'enri-
chissement du vocabulaire utilisé
dans une conversation.

des suggestions pour améliorer l'exactitude linguistique (Table 4). Ce type de feedback soutient le développement des compétences syntaxiques et phonétiques des apprenants, tout en maintenant l'accent sur le contenu et le sens du discours, ce qui est essentiel pour une interaction authentique.

Ainsi, *English Quest* se révèle être un outil didactique prometteur, capable de combiner les avantages de scénarios pédagogiques ouverts avec la rigueur des tâches linguistiques spécifiques. En s'appuyant sur des principes didactiques éprouvés, le jeu réussit à créer un environnement d'apprentissage qui est à la fois engageant, stimulant et créatif.

Analyse du potentiel et des limites

English Quest présente un potentiel pédagogique remarquable, particulièrement grâce à son approche immersive et interactive qui se situe entre les niveaux de Modification et de Redéfinition dans le modèle SAMR (Substitution, Augmentation, Modification, Redéfinition) (Hamilton et al., 2016). À travers l'utilisation de l'intelligence artificielle, le jeu ne se contente pas d'améliorer les méthodes traditionnelles d'enseignement des langues, mais redéfinit complètement la manière dont les apprenants interagissent avec la langue. En créant des scénarios narratifs où les interactions

sont personnalisées en temps réel, *English Quest* peut permettre aux apprenants de vivre des situations de communication authentiques qui étaient auparavant plus difficilement accessibles sans l'usage de la technologie d'intelligence artificielle.

L'une des principales forces de ce jeu réside dans sa capacité à réduire l'anxiété liée aux erreurs linguistiques, un obstacle fréquent dans l'apprentissage des langues (Blyth, 2018). En mettant l'accent sur la communication du sens plutôt que sur l'exactitude formelle, *English Quest* encourage les apprenants à prendre des risques linguistiques dans un environnement sécurisé. Cette approche soutenue par des feedbacks immédiats et constructifs, aide à développer la confiance en soi et la fluidité dans l'utilisation de la langue (Reinhardt, 2019). Selon Neville et al. (2009), l'intégration des jeux numériques dans l'apprentissage des langues peut également faciliter la rétention et le transfert des connaissances, un avantage que ce jeu exploite pleinement en combinant immersion narrative et interaction personnalisée.

Cependant, malgré ses nombreux atouts, *English Quest* n'est pas exempt de limites, principalement dues à son statut de prototype. Par exemple, le jeu exige un niveau de compétence linguistique relativement élevé pour comprendre les textes présentés (B2/C1) (CECR, 2018), ce qui pourrait exclure les apprenants moins avancés. De plus, le niveau de langue n'est pas encore ajustable, ce qui limite la capacité du jeu à s'adapter à une variété plus large de profils d'apprenants. De même, bien que les scénarios soient immersifs et engageants, ils peuvent parfois manquer de profondeur, laissant peu de place à l'exploration libre, ce qui pourrait frustrer certains utilisateurs.

Ces limitations, bien qu'actuelles, ne sont pas insurmontables. Elles offrent des pistes de développement claires pour les futures itérations du jeu. Plusieurs pistes de développement pourraient être envisagées pour enrichir l'expérience de *English Quest*. D'abord, il pourrait être bénéfique de permettre aux apprenants de choisir des objectifs d'apprentissage spécifiques afin d'orienter les activités et les feedbacks, en se concentrant sur des aspects linguistiques précis comme la grammaire, le vocabulaire ou la prononciation. Ensuite, la possibilité de créer leurs propres scénarios pourrait offrir une dimension créative et engageante, tandis que renforcer l'impact des choix faits lors de la création du personnage rendrait l'expérience encore plus immersive.

De plus, étendre la fonctionnalité de dialogue audio pour parler directement aux personnages pourrait permettre de développer les compétences orales de manière interactive. Une autre piste pourrait être de structurer une partie du scénario initialement, puis de laisser l'IA étendre l'histoire de manière flexible, rendant chaque aventure unique. En outre, une progression persistante tout au long du jeu, avec la possibilité de rejouer certaines sections pour débloquer de nouveaux éléments, pourrait offrir un fort potentiel de motivation. Enfin, un personnage accompagnant l'apprenant tout au long de l'aventure, disponible pour répondre aux questions, pourrait ajouter un soutien personnalisé et renforcer l'engagement.

Perspectives et opportunités

En combinant les éléments de *Digital Game-Based Learning* (DGBL), d'*Interactive Fiction* (IF), et de scénarios pédagogiques ouverts, *English Quest* ouvre des perspectives prometteuses pour l'avenir de l'apprentissage des langues, en particulier dans le contexte d'une éducation numérique en constante évolution. Ce prototype, bien qu'encore en développement, montre déjà comment les technologies immersives, associées à l'intelligence artificielle, peuvent transformer les pratiques pédagogiques traditionnelles en créant des environnements d'apprentissage plus interactifs, personnalisés et engageants.

English Quest peut permettre aux apprenants de vivre des situations de communication authentiques qui étaient auparavant plus difficilement accessibles sans l'usage de la technologie d'intelligence artificielle.

Bibliographie complète

Aguilera, M., & Roock, R. (2022). *Digital Game-Based Learning: Foundations, Applications, and Critical Issues*. *Journal of Educational Technology*, 21(2), 123-145.

Blyth, C. (2018). Immersive technologies and language learning. *Foreign Language Annals*, 51(1), 225-232.

Conseil de l'Europe. (2018). *Cadre européen commun de référence pour les langues – Apprendre, enseigner, évaluer. Volume complémentaire avec de nouveaux descripteurs* (G. Breton & C. Tagliante, Trad.). Programme des Politiques linguistiques. Division des Politiques éducatives. Service de l'Éducation.

González-Lloret, M. (2020). *Task-based language teaching in technology-mediated contexts*. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 1-34). John Benjamins Publishing Company.

Hamilton, E., Rosenberg, J., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) Model: A Critical Review and Suggestions for its Use. *TechTrends*, 60(1), 433-439. <https://doi.org/10.1007/s11528-016-0091-y>

Huber, G., Meier, R., & Smith, T. (2024). *Leveraging the Potential of Large Language Models in Education: Opportunities and Challenges*. *Journal of Educational Technology*, 21(2), 123-145.

Montfort, N. (2013). *Riddle Machines: The History and Nature of Interactive Fiction*. In R. Siemens & S. Schreibman (Eds.), *A Companion to Digital Literary Studies* (pp. 267-282). John Wiley & Sons.

Neville, D. O., Shelton, B. E., & McInnis, B. (2009). Cybertext redux: Using digital game-based learning to teach L2 vocabulary, reading, and culture. *Computer Assisted Language Learning*, 22(5), 409-424. <https://doi.org/10.1080/09588220903345168>

Nissen, E. (2019). Pourquoi mettre à contribution le numérique dans l'enseignement des langues à l'école ? In *Conférence de Consensus, Langues vivantes étrangères* (pp. 13-15). http://www.cnesco.fr/wp-content/uploads/2019/04/CCLV_Nissen_MEF-v2.pdf

Reinhardt, J. (2019). *Gameful Second and Foreign Language Teaching and Learning: Theory, Research, and Practice*. Palgrave Macmillan.

Schrader, M. (2023). *Serious Games and Game-Based Learning*. *Educational Game Studies*, 15(1), 47-63.

Ziegler, N. (2016). Synchronous Computer-Mediated Communication and Interaction: A Meta-Analysis. *Studies in Second Language Acquisition*, 38(3), 553-586.

STUDENT INTERACTIONS WITH EDUCATIONAL AI CHATBOTS IN LANGUAGE FOR SPECIFIC PURPOSES: INSIGHTS FROM USAGE AND PERSPECTIVES

Questo studio esamina l'uso di un chatbot educativo IA in un corso di inglese commerciale con 142 studenti in un'università della Germania meridionale. Dall'analisi delle interazioni sono emerse quattro categorie di utilizzo: richiesta di esempi, guida strutturale, feedback e valutazione. Un sondaggio su 73 studenti rivela un'elevata usabilità e fiducia nel chatbot, ritenuto efficace per migliorare le capacità di scrittura. Tuttavia, sebbene il chatbot abbia aiutato dal punto di vista linguistico, sorgono preoccupazioni riguardo al fatto che alcuni studenti si concentrino sulle correzioni superficiali piuttosto che su miglioramenti concettuali più profondi, evidenziando la necessità di strategie che incoraggino un utilizzo più sostanziale dello strumento.

● Alice Gruber
| Technical University
of Applied Sciences
Augsburg



Alice Gruber is a professor of Business English at the Technical University of Applied Sciences Augsburg. She has taught at the secondary and tertiary levels in the UK and Germany. Her research areas include Augmented and Virtual Reality, AI for language learning, Global Englishes and Virtual Exchange.

Background

Artificial Intelligence (AI) has transformed language learning by providing tailored experiences and personalised feedback. AI-powered chatbots simulate real-life conversations, enabling extended practice in speaking, writing, and listening without fear of face loss. However, challenges such as standard language bias in AI (Payne et al., 2024) require critical examination in educational settings. Recent studies have examined the influence of AI-powered chatbots on L2 writing, including individual and collaborative processing of AI-generated feedback on overall writing quality (Yan, 2024), L2 writers' revision strategies (Ziqi et al., 2024), and engagement with corrective feedback provided by ChatGPT (Yan & Zhang, 2024). However, students' engagement with chatbots specifically designed for pedagogical purposes remains underexplored. This exploratory study examined human-AI collaboration in a language-for-specific-purposes setting, focusing on student interactions with a

chatbot to develop business writing skills and their perceptions of its use.

Methodology

Mizou, a GDPR-compliant AI chatbot that allows educators to create tailored chatbots, was used in its paid version during a compulsory in-person B2 level Business English course for 142 second-semester Business Administration students at a German university during the summer term of 2024. The students completed writing tasks, including business emails, with the option to use Mizou. The tasks were completed in class under instructor supervision to ensure independence from external AI tools. Students were explicitly instructed to avoid using AI while writing their texts.

The chatbot was programmed to provide feedback without completing tasks, as configured by the instructor. For this exploratory study, the corpus comprised students' interactions with three

Category	Example
Requesting Examples	'Examples of the beginning of the Body of an Email'; 'Give me an example of how to describe the delivery service package'; 'Can you give me an example for the professional Tone';
Asking for guidance in terms of structure and content	'What should I write for point 2?'; 'How do I express regret'; 'How is the best way to apologize?'; 'what mistakes should I avoid?'; 'can i offer a replacement and discount'; 'is it good to tell the costumer [sic] that she should send us the product back so we can investigate it?'
Asking for feedback (e.g., on their texts and corrections)	'what could still be made better?'; 'Tell me about grammar synthax and language'; 'is that okay?'; 'Bitte gebe mir ein Feedback'; 'Can you check grammar and vocabulary for me?'
Grade and word limit	'What would be the grade'; 'Grade'; 'rate my essay in case of content, grammar fluently and check the word limit'

Table 1

Categories of chatbot use

task-specific Mizou chatbots, accessed 345 times in total. All comments, questions, and commands from the interactions were analysed.

Results

Four distinctive categories of chatbot use were identified. Table 1 shows that some students asked generic questions, while others posed tailored, context-specific ones, indicating deeper engagement and potentially eliciting more precise chatbot responses.

The number of text versions each student uploaded for Mizou's feedback ranged from 1 to 7, though most students uploaded their text only once, often without providing any instruction.

Some students were economical with their language, giving one-word instructions such as "Grade" or "Example" instead of full requests. Others adhered to interpersonal conventions, incorporating greetings such as "Hi there help me starting with my writing", expressing gratitude with phrases like "Thanks for the help" and using an emoji: "Thank you. you too 🍷". Some even reassured the chatbot with comments such as "Ok, calm down" after receiving suggestions. Additionally, some students used language typical of human interaction, such as, "What do you think about this?" and "Okey got it go on". While the students' communication was generally conducted in English, some instances of translanguaging were noted. For example, a student wrote, "How can i say that she become [sic] a retouren Schein for the Post", indicating a struggle with e-commerce terminology. The chatbot successfully understood the students'

request and provided assistance without any difficulties.

Student perceptions and attitudes

At the end of the course, 73 students participated in an anonymous survey administered through Moodle. The survey consisted of 17 questions designed to gather feedback on their experience with Mizou.

Students were asked two general questions about AI to assess their openness and frequency of use. A large majority (85%) indicated that they enjoy using AI for their studies. However, their actual frequency of use varied, with around 22% using it very frequently, 37% frequently, 26% occasionally, 12% rarely, and 3% very rarely.

Most respondents (93%) found Mizou easy to use. Students generally viewed Mizou as a valuable tool to improve their Business English skills, with over half rating it as very useful or useful. Additionally, most students regarded Mizou as effective in providing detailed feedback on their writing, rating it very effective (21%) or effective (55%). Meanwhile, 22% felt neutral about Mizou's feedback, and a small percentage (3%) rated it as ineffective. Notably, no respondents found it very ineffective.

Students identified several specific features of Mizou as particularly useful. Overall feedback, i.e. providing a broader evaluation of their writing, was valued the most (62%). Grammar correction and sentence structure improvement were also appreciated, with 57% of students

Statement	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
1. The use of Mizou encourages me to practise writing more frequently	4%	32%	47%	14%	4%
2. Mizou helped me make good progress.	5%	55%	32%	8%	0%
3. I would recommend using Mizou to other students for writing practice.	18%	58%	19%	4%	1%
4. Using Mizou reduces my fear of making mistakes in my writing.	5%	32%	59%	2%	2%

Table 2
Student Perceptions of the Chatbot's
Usefulness for Writing Practice

recognising their benefits. Additionally, a notable number of students found vocabulary suggestions (44%) and the ability to receive instant feedback (40%) helpful.

When asked about their trust in Mizou, the responses suggest that while many students generally trust the chatbot (27% almost always; 45% often), they also engage critically, demonstrating selective trust (25% sometimes; 3% rarely). This critical engagement is desirable in educational contexts that encourage critical thinking, as it fosters reflective learning.

Regarding motivation to write based on feedback from Mizou, respondents generally did not feel demotivated, yet a strong positive motivation is absent for most. A small portion of students (4%) reported feeling much more motivated, with 27% feeling more motivated. The majority (62%) remained neutral, while only 7% felt less or much less motivated. Most students expressed a positive interest in working with Mizou in the future (19%

strongly agreed and 5% agreed). Meanwhile, 21% were undecided, and a smaller proportion (7%) expressed reservations.

Table 2 table shows a generally positive view of Mizou for writing practice. Most students believe the tool aids their progress and would recommend it, though opinions vary on its impact on reducing fear of mistakes. Many feel motivated to practise more, while a notable portion remains neutral.

Regarding receiving feedback from Mizou vs. human feedback, the responses show that 42% of participants feel neutral about Mizou's feedback compared to human feedback. 30% feel more comfortable, while 7% are much more comfortable. On the other hand, 21% feel less comfortable, and no participants reported feeling much less comfortable. Overall, the results are mixed, with a slight preference for comfort or neutrality towards Mizou's feedback.

Regarding feedback preferences, over a third of respondents somewhat prefer receiving feedback from a human instructor, with a smaller group strongly preferring human feedback. In contrast, 22% somewhat prefer feedback from Mizou, while only a few strongly favour it. About a quarter of respondents expressed no preference between human feedback and Mizou. Overall, there is a general preference for human feedback, although many respondents also appreciate Mizou's feedback or remain neutral.

Students were invited to comment on their chatbot experience. Many praised its effectiveness in improving writing

Some students focused on quick fixes, avoiding deeper dialogue that could support their writing skills development.

skills, providing structured and immediate feedback (e.g., “helped me to get feedback at any time”), being easy to use, and providing grammar assistance. One student appreciated that “Mizou is not judging like a human could do”, while another expressed a preference for human feedback, stating, “it’s okay but I prefer a human being – it’s just personal”. However, some students mentioned technical issues, such as slow performance and non-functioning links, and noted the repetitive nature of the feedback, with one commenting, “It was okay, but often the replies are quite the same – I missed some deepness”.

Chatbot interactions often required students to process lengthy responses, serving as covert reading tasks that may support language development

Discussion and conclusion

Students used the chatbot to explore different writing aspects, including seeking examples, guidance on structure and content, grammar checks, better phrasing and syntax improvement. Engagement levels varied; some students interacted minimally with the chatbot, while others engaged intensively, demonstrating a proactive approach by requesting further feedback after making revisions. These findings highlight key differences in how students use AI for writing, suggesting varied level of writing autonomy and learning strategies.

The findings indicate that some students value iterative learning, using the chatbot as a tool to refine their skills. In contrast, some students focused on quick fixes, avoiding deeper dialogue that could support their writing skills development. Future research should investigate the factors that influence learners’

willingness to engage deeply with the chatbot, including motivation and perceptions of AI tools, particularly regarding academic integrity.

Chatbot interactions often required students to process lengthy responses, serving as covert reading tasks that may support language development. Future studies should examine how the quality and types of student questions evolve with increased engagement with educational chatbots, with particular attention to variations across proficiency levels. This includes investigating the impact of chatbots on students’ reading skills, and exploring students’ pragmatic behaviour in human – machine interactions. It is essential to explore how educational chatbots can be optimised for different educational contexts while also identifying the conditions and learner behaviours that support effective use.

References

Payne, A. L., Austin, T., & Clemons, A. M. (2024). Beyond the front yard: The dehumanizing message of accent-altering technology. *Applied Linguistics*, 45(3), 553–560. <https://doi.org/10.1093/applin/amae002>

Yan, D. (2024). Comparing individual vs. collaborative processing of ChatGPT-generated feedback: Effects on L2 writing task improvement and learning. *Language Learning & Technology*, 28(1), 1–19. <https://hdl.handle.net/10125/73597>

Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. *Humanities and Social Sciences Communications*, 11(1), 1–14.

Ziqi, C., Xinhua, Z., Qi, L., & Wei, W. (2024). L2 students’ barriers in engaging with form and content-focused AI-generated feedback in revising their compositions. *Computer Assisted Language Learning*, 1–21. <https://doi.org/10.1080/09588221.2024.2422478>

PROJECT TOOLS@SCHOOLS:IMMERSIVE – TASKS FOR A CRITICAL USE OF GENERATIVE AI IN IMMERSIVE FOREIGN LANGUAGE SEQUENCES IN SUBJECT TEACHING

Mit den rasanten Fortschritten in der digitalen Übersetzung und Künstlichen Intelligenz steigt der Bedarf an Materialien, die den reflektierten Einsatz von KI im Sprachenlernen fördern. Im Rahmen des Tools@Schools:immersiv-Projekts entwickeln wir daher Aufgaben, die digitale und fremdsprachliche Fähigkeiten stärken und gleichzeitig generative KI in immersive Fremdsprachen-Sequenzen (*ilots immersifs*) einbinden. Bisher haben wir zwei Aufgaben erstellt – flexibel an verschiedene Sprachen anpassbar für die Fächer Hauswirtschaft und Geschichte – und setzen KI-Tools wie ChatGPT ein, um eine kooperative Interaktion mit KI zu ermöglichen. Dieser Ansatz fördert nicht nur das Sprachenlernen und das Verständnis von Inhalten, sondern auch kritisches Denken und eigenständiges Lernen.

● Catherine Ferris
| St.Gallen University
of Teacher Education
Sandrine Wild
| St.Gallen University
of Teacher Education
Stefano Negrinelli
| St.Gallen University
of Teacher Education

1. Introduction

The proliferation of text generation tools, such as ChatGPT, indicates a potential expansion of their influence on education, including primary and secondary levels (LCH, 2024; SBFI, 2019). While the possibilities and impact of AI tools for text generation in higher education have received quite a lot of attention (cf. Meyer & Wessels, 2023; Schmohl et al., 2020) specifically with regards to academic writing, AI tools are increasingly used for all sorts of tasks for language learning, for instance listening and speaking development, writing gains, vocabulary increase, and grammar learning (Dizon, 2020; Huang et al., 2023; Lee, 2019; Woo & Choi, 2021; Zhang & Huang, 2024). As repeatedly stated in the research, the question for the foreign language classroom is not whether teachers and learners welcome or consider the use of AI tools in language learning, but rather that they do this with a certain degree of attention and reflection. Language learning is not confined to traditional

foreign language classrooms, it can also take place in immersive foreign language environments. Digital tools and platforms can facilitate the implementation of immersive learning. This shift aims to integrate foreign language use into a wide range of real-world scenarios, mirroring natural language immersion. Moreover, non-language subjects can simulate cultural immersion, which complements foreign language use.

This paper examines the products and preliminary results from the project “Tools@Schools:immersiv”, in which we are creating tasks to enhance digital and foreign language skills. There are many ways to integrate a foreign language into subject teaching. For example, by incorporating an original Italian source in history lessons or a recipe for a French specialty in French for home economics classes. These texts in the original foreign language are authentic and thus interesting (Efverlund & Anyadi, 2022), but they can be linguistically challenging for the students. A simple translation using a

machine translation tool does not, or only partially, serve language learning and the objectives of subject teaching in a foreign language. Even in this context, the use of online translation tools and automated AI tools requires guidance to ensure that language learning occurs alongside content learning. Our goal is to incorporate generative AI into immersive foreign language learning sequences (*îlots immersifs*) within subject teaching, thereby addressing contemporary didactic methods in foreign language education at the lower secondary level. An *îlot immersif* or 'immersive island' refers to a shorter or longer teaching sequence in a foreign language that integrates the teaching and learning of foreign languages in subject teaching (Freytag Lauer, 2019; 2018). In this sense, immersive islands can be understood as a special format of bilingual education (Gajo et al., 2018).

The tasks are designed to incorporate generative AI as a resource that users engage with and use cooperatively, rather than as a replacement for their own thinking. Through activities such as interactive translation practice, writing assistance and post-editing, and preparation of oral presentations or debates, students can enhance their language skills and subject knowledge in collaboration with AI. This approach fosters critical thinking and promotes autonomous learning.

2. Methodology

In line with the tasks from the first project Tools@Schools (Perrin et al., 2022; 2021), we developed tasks which consist of four parts or sub-tasks:

- a) familiarization with the topic,
- b) consultation of the AI tool,
- c) language production using the tool, and
- d) final reflection.

The task structure is based on the LUKAS model, a process model for creating competence-oriented tasks, developed at the University of Teacher Education Lucerne (Wespi et al., 2015). This model describes a natural learning process where learners initially familiarize themselves with a topic or issue, then deepen their understanding, and eventually engage with it more thoroughly by transferring their gained knowledge to a new context. The goal is for the task structure to support these sequential learning steps,

promoting the learning process from both a usage- and competence-based perspective. The topics of the tasks were chosen based on their thematic occurrence in the current curriculum ('Lehrplan 21' in the Eastern part of Switzerland) for lower secondary level, their connection to an authentic language and cultural immersion, and their suitability for immersive contexts. Completing the tasks takes one standard lesson of 45 minutes, depending on the general pace in the classroom and whether the teacher has already spent some time previously discovering AI tools with their students.

To improve manageability and allow a certain degree of flexibility, the tasks were modularized. This allows teachers to prioritize certain sub-tasks, such as the familiarization phase, before progressing to a more intensive work phase with the tool. The tasks include guidance for teachers on how to adjust them based on their position within the curriculum. Further, they provide sample scenarios in which the task can be implemented.

At this stage in the development and revision process, the tasks are tested in classrooms. The tasks can be handled by foreign language teachers or teachers of non-language subjects, as the goal is to promote implicit language learning in collaboration with AI tools. This approach eliminates the need for both students and teachers to possess advanced foreign language skills, which would otherwise be essential without the support of AI. In the following, we describe a sample task before reporting on the testing of this task in an 8th grade (lower secondary level) in the Canton of St.Gallen, Switzerland.

We develop tasks to enhance digital and foreign language skills while supporting generative AI in immersive foreign language sequences. This approach enhances both language learning and content understanding, while also promoting critical thinking and autonomous learning.



Dr. Catherine Ferris is a lecturer in English linguistics for future teachers of lower secondary education and a research associate at the Institute for Language and Literary Education at St.Gallen University of Teacher Education.



M.A. Sandrine Wild is a research associate at the Institute of Social Studies Education and a lecturer in home economics and in foreign language didactics at St.Gallen University of Teacher Education.



Dr. Stefano Negrinelli is a research associate at the Institute for Language and Literary Education and a lecturer in foreign language didactics for lower secondary education at St.Gallen University of Teacher Education.

3. Sample task

One task presents a so-called *îlots immersif* in the foreign language French in home economics class. The topic for the task is the economic cycle. In the first sub-task, students explore the notion of economic cycle by reflecting on the sources of their pocket money and mapping out how they spend it. They then imagine a specific scenario in which they spend 50 Swiss Francs on a game at the Swiss supermarket Migros and note down where the money goes next, thereby activating the students' prior concepts and knowledge.

Following this, they consult an AI-tool for the first time as depicted in Figure 1. They ask ChatGPT where the money goes, thereby familiarizing themselves with different ways of prompting, such as adopting the role of an accountant at Migros. Students note down the answers provided by the chatbot and compare them with their own, as well as those of their peers. By engaging with their peers, they compare and critically reflect on different prompting outcomes. This first phase can be done in the language of schooling (German in this case).

The second phase introduces the foreign language. In subjects such as home economics, many topics have language or culture-specific character, and the topics themselves relate to everyday notions which we may encounter when immersing ourselves in the language

and culture. For these tasks, emphasis is placed not only on linguistic correctness but also on subject-specific language. Therefore, we chose a topic that is relevant to the students' daily lives: money. They watch a YouTube tutorial in French that explains the economic cycle. In an engaging video with French subtitles, the influencer Hanna Cash shows how the economic cycle works, why things cost money, and why we must work to pay for them.

After watching the video, the students complete a table with key concepts and keywords mentioned in French, adding their German translations (they may use machine translation to assist with this). Next, they compare their word lists with the descriptions they used in phase 1 (cf. Figure 1) to complement and hence explain the concept more holistically.

In the production phase, they sketch an entire economic cycle, using a scenario of their choice (pocket money or other) and using the French words they have collected throughout the task. They can do this in the form of a mini poster, on a flip chart or on the blackboard. The aim is for every student (alternatively in pairs) to present their depiction to the class in the foreign language. They are shown how to practice their short input with the help of the listening / pronunciation function of the AI tool. In groups of 2-3 students, they discuss how they used the chatbot and other digital tools, what difficulties they encountered, and what worked out well.

In the second task, designed for history class, students explore primary sources on the topic of migration in their original language, Italian. Students, who have very little to no prior knowledge of Italian, attempt to understand the texts using intercomprehension (Tafel et al., 2009) and, later, with the assistance of AI. Initially, they examine the texts to identify familiar words and expressions in Italian. Then, they use a chatbot in German to delve deeper, uncovering more nuanced meanings. Imagining themselves as early 20th-century emigrants traveling from the Italian-speaking part of Switzerland, the Ticino, to the USA, they immerse in the historical context. Finally, they reflect on how the chatbot aided their comprehension and enriched their understanding of the texts.

a) Nehmen wir an, dass du bei der Migros ein Spiel für 50 Franken gekauft hast. Was denkst du: Wohin fließen nun die 50 Franken, die du der Migros eben bezahlt hast?

Frage: Wohin fließen die 50 Franken, die ich in der Migros ausgegeben habe?

b) Gehe nun zu ChatGPT und stelle dem Chatbot («Chat-Roboter») Fragen, um herauszufinden, wo das Geld tatsächlich hinfließt. Achte beim Befehl ('prompting') darauf, dass du dem Chatbot sagst, welche Rolle es übernehmen soll. Also dass du seine Expertise willst, z.B. «Du arbeitest bei der Migros in der Buchhaltung...» Welche Ideen hat ChatGPT genannt, an die du nicht gedacht hast?

Figure 1

Excerpt from the task "Economic cycle" ('Le circuit économique' – der Wirtschaftskreislauf)

4. Discussion

The classroom testing of the “pocket money” task was conducted in June 2024 with an 8th grade secondary class of 18 students. Following this, the project team interviewed the teacher to gain some insights on her as well as her students’ perception of the task. Although the task was tested in a language-learning context rather than in home economics, the teacher’s extensive instructional experience provided valuable perspectives on student comprehension and task feasibility. With her background in French, German, English, and visual arts, she offered observations on the task’s adaptability, assessing its suitability for teachers with differing levels of French proficiency. The teacher generally had a positive impression of the task. In the previous weeks, they had been covering AI awareness in the classroom. For instance, all students had previously tried out ChatGPT and done some prompting. The teacher provided them with insights, such as how the role of perspective in prompting can lead to more accurate results.

The task contains several learning goals stated at the beginning. Interestingly, the teacher prepared her students for the task by engaging in a meta-reflexive discussion on the learning aims. In this way, she clarifies student expectations and promotes students’ active collaboration in skill development. This procedure contributes strongly to the overarching goal of promoting students’ general reflective skills. As stated, the tasks themselves and the final step in particular aim to strengthen students’ critical thinking and reflection on the use of the tools, their benefits and pitfalls. Incorporating these competences into a general reflective setting is beneficial.

The teacher assessed the task difficulty as relatively low, which may be attributed to her students’ familiarity with chatbots like ChatGPT or to the general competence level in the classroom, despite her view that the class is evenly split between stronger and weaker performers. Therefore, it is important to acknowledge that this conclusion is not definitive and would gain reliability through comparisons across different classrooms. Moreover, a task perceived as easy might be even less challenging for strong students,

while failing to adequately reflect the challenges faced by less proficient students in the class.

According to the teacher, the students were highly engaged with the chatbot during sub-tasks where they asked the chatbot to elaborate on processes (e.g., the economic cycle). Working in pairs, they discussed and compared the vocabulary provided by the chatbot with their own. The teacher noted that the students showed a strong interest in this vocabulary work and could have benefited from spending more time on this aspect. This observation ties in nicely with previous research on the use of chatbots to broaden one’s vocabulary. Zhang & Huang’s study compared two groups of foreign language learners: one using a large language model-based chatbot and one without. Over eight weeks, the chatbot group showed significantly better vocabulary knowledge, especially in long-term retention, suggesting large language models as effective language-learning tools. (Zhang & Huang, 2024).

The pocket money task focuses on vocabulary. With a process-related topic such as the economic cycle, it would be interesting to focus on grammatical aspects as well, for instance describing the past or hypothesizing about the future in a foreign language. According to the teacher, we need to have more confidence in students to do more language work with the help of the chatbot, including exploring linguistic nuances such as specific grammatical instances. This is particularly interesting as the emphasis of immersion lies not on coerced language acquisition but rather on engaging with the foreign language in an alternative context, without the same pressure to enhance linguistic proficiency as in traditional language classes. But with the chatbots offering increasingly versatile possibilities, it is essential to consider integrating specific, targeted linguistic objectives that go beyond vocabulary development. This feedback might suggest that teachers encourage students with varying language proficiencies to engage with the AI tool to different extents for language learning, setting specific content and language aims.

The small-scale testing of this one task shows the potential of an application at the lower secondary level. Also, in this

case, the application worked well even though the teacher is not a home economics teacher. This raises the question of how the tasks can be implemented if the subject teacher is not a language teacher or if the teacher is responsible for both the subject and language instruction. There are three dimensions which deserve more attention for the development and revision of further tasks: alignment of subject-specific competencies and language learning, differentiation of students' performance, and implementation of classroom policies. The described materials for immersive contexts are developed by teachers and teacher trainers teaching in both the foreign language and the subject matter. There is need to involve additional educators in the co-creation of the tasks to ensure they address both linguistic demands and subject content comprehensively. Furthermore, the final sub-task, which encourages reflection, should be designed to actively engage students in thinking about their language use and content understanding, helping them to become more aware of how language skills support their learning across other subjects.

Currently, the tasks contain commentary for teachers on how to adapt or modify the tasks according to students' differing proficiency levels. We have yet to gain more insights across classrooms as to the difficulty of the tasks. As highlighted by the teacher who kindly volunteered to test the task, many teachers are left on their own when it comes to systematically implementing machine translation and AI tools in their classrooms. While some teachers are highly engaged in exploring new digital developments and eager to try them out in the classroom, others feel overwhelmed by the fast-moving developments and oftentimes do not know where to begin. Wider, possibly level-specific, classroom policies should be considered. Additionally, teachers should have access to further education courses, along with teaching materials and guidance on implementation.

Through activities such as interactive translation practice, writing assistance and post-editing, and preparation of oral presentations or debates, students can enhance their language skills and subject knowledge in collaboration with AI.

5. Bibliography

Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, 24 (1), 16–26. <https://doi.org/10125/44705>

Efverlund, Y. & Anyadi, M. (2022). Briefly put, authenticity makes it more interesting! : A study of the role of authentic written communication (Dissertation). Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-50589>

Freytag, A. (2018). Les îlots immersifs au primaire : une pratique pédagogique transcurriculaire. In : F. Caspani, V. Manna, V. Todisco & M. Trezzini (eds.), *Mehrsprachiger Unterricht im Fokus. Atti del Convegno sui modelli d'insegnamento veicolare delle lingue straniere*. Chur: Pädagogische Hochschule Graubünden, pp. 184–195.

Freytag, A. (2019). Les documents dans les îlots immersifs : authenticité et authentification. *Revue TDFLE*, 74. https://doi.org/10.34745/numerev_1369

Gajo, L., Steffen, G., Freytag, A., & Vuksanovic, I. (2018). Quelle frontière entre enseignement bilingue et enseignement de la langue orienté vers le contenu ? Insertion curriculaire et moyens d'enseignement. *Babylonia: revue pour l'enseignement et l'apprentissage des langues*, 2, 26–31.

Huang, X., Zou, D., Cheng, G., Chen, X. & Xie, H. (2023). Trends, Research Issues and Applications of Artificial Intelligence in Language Education. *Educational Technology & Society*, 26 (1), 112–131.

LCH. (2024). *Künstliche Intelligenz in der Schule: Chancen nutzen, Herausforderungen meistern*. Positionspapier des Dachverbands Lehrerinnen und Lehrer Schweiz.

Lee, S.-M. (2019). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33 (3), 157–175. <https://doi.org/10.1080/09588221.2018.1553186>

Meyer, E. & Wessels, D. (2023). Natural Language Processing im akademischen Schreibprozess - mehr Motivation durch Inspiration? Positionspapier basierend auf einer Fallstudie an der Fachhochschule Kiel. In: Schmohl, T., Watanabe, A. & Schelling, K. (eds.), *Künstliche Intelligenz in der Hochschulbildung. Chancen und Grenzen des KI-gestützten Lernens und Lehrens*. Bielefeld: transcript 2023, pp. 227–251.

Perrin, R., Diederich, C., Wild, S. & Grimm, L. (2022). Überzeugungen von Lernenden der Sekundarstufe I zur Nützlichkeit der Übersetzungstools DeepL und LEO für die Wortschatzarbeit im Fremdsprachenunterricht. *Babylonia: Journal of Language Education*, 1, 38–41. <https://doi.org/10.55393/babylonia.v1i.156>

Perrin, R., Wild, S., Diederich, C. & Grimm, L. (2021). Digitale Übersetzungstools im Fremdsprachenunterricht der Sekundarstufe I. Nutzung von DeepL und LEO aus der Sicht von Lernenden am Beispiel einer Schreibaufgabe. In: Juska-Bacher, B., Brugger, L., Fischer, A. & M. Röthlisberger M. (eds.), *Defining, assessing and fostering vocabulary skills*, Bulletin VALS-ASLA, pp. 113–129.

SBFI. (2019). *Herausforderungen der künstlichen Intelligenz*. Bericht der interdepartementalen Arbeitsgruppe «Künstliche Intelligenz» an den Bundesrat. Staatssekretariat für Bildung, Forschung und Innovation.

Schmohl, T., Watanabe, A., Fröhlich, N. & Herzberg, D. (2020). How can artificial intelligence improve the academic writing of students? In: Pixel (eds.), *Conference proceedings. International conference "The future of education", 10th edition*. Florence, Italy, 18–19 June 2020. Bologna: Filodiritto Editore, pp. 168–171.

Tafel, K., Durić, R. Lemmen, R., Olshevska, R. & Przyborska-Stolz, A. (2009): *Slavische Interkomprehension. Eine Einführung*. Tübingen: Gunter Narr Verlag.

Wespi, C., Luthiger, H. & Wilhelm, M. (2015). Mit Aufgabensets Kompetenzaufbau und Kompetenzförderung ermöglichen. *Haushalt in Bildung & Forschung*, 4, 31–46.

Woo, J. & Choi, H. (2021). Utilizing Machine Translation Systems to Generate Word Lists for Learning Vocabulary in English. *Journal of Digital Contents Society*, 22 (1), 71–75. <https://doi.org/10.9728/dcs.2021.22.171>

Zhang, Z. & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, 10. <https://doi.org/10.1016/j.heliyon.2024.e25370>

PER UNA DIDATTICA DELLE LINGUE STRANIERE CONNETTIVISTA

Dieser Beitrag entstand aus den Überlegungen zum Lehrbuch "Es macht klick: Künstliche Intelligenz bei schriftlichen Arbeiten clever nutzen" (Alloatti & Montemarano 2024), das 2025 in italienischer Sprache bei Paravia erscheinen ist und sich insbesondere auf die transversalen Kompetenzen konzentriert, die für den Einsatz von KI-Tools im Unterricht aller Fächer erforderlich sind. In diesem Beitrag fokussiert sich die Perspektive hingegen nur auf den schulischen Fremdsprachenunterricht und wechselt den Fokus von der Unterrichtspraxis zur Theorie, die die Praxis leitet.

● Sara Alloatti
| Università di Zurigo
Filomena
Montemarano
| Scuola Pedagogica
PH FHNW



Sara Alloatti, docente di didattica all'università di Zurigo, insegnante di italiano al liceo di Uetikon am See.



Filomena Montemarano, docente di didattica alla Scuola Pedagogica PH FHNW.

1. L'IA che svolge i compiti al posto dell'apprendente

La popolazione scolastica – sia apprendenti che insegnanti – ha reagito in molti modi diversi alla diffusione dell'intelligenza artificiale generativa, in particolare dei traduttori automatici (p. es. DeepL) e dei generatori di testo (p. es. ChatGPT). C'è chi, fin dalla prima ora, li ha usati assiduamente, imparando a combinarli in maniera efficace (o all'opposto fidandosi ciecamente di essi) e chi, invece, li ha banditi dal proprio repertorio – per paura, per principio o per sfiducia. Il passare dei mesi ha messo alla luce significative disparità sia tra apprendenti che tra insegnanti riguardo: a) alle **conoscenze** di diversi strumenti di intelligenza artificiale, b) alle **competenze** nel loro utilizzo, c) agli **atteggiamenti** nei confronti delle loro potenzialità e dei rischi associati e, di conseguenza, d) anche in relazione alle **reazioni emotive** individuali.

È questo divario che abbiamo osservato giorno per giorno crescere nelle scuole

e nelle università in cui insegniamo: al contempo, nel tentativo di cogliere la portata dell'innovazione sulla nostra materia, abbiamo cominciato a chiederci quali vantaggi e svantaggi comportasse l'uso mirato di traduttori e generatori ai fini dell'apprendimento della lingua straniera. Così facendo, ci siamo rese conto che la domanda riguardava un ambito ben più ampio: quello della didattica della scrittura nel contesto scolastico, indipendentemente dalla lingua e dalla materia insegnata.

Ma torniamo alle lingue straniere: i generatori di testo ci pongono indubbiamente di fronte a interrogativi cruciali sulla **funzione**, sulla **validità** e sull'**autenticità** dei formati testuali scritti che tradizionalmente si producono a scuola, dal semplice esercizio a completamento fino alle più complesse presentazioni in classe. Per esempio: ha ancora senso far scrivere a casa un testo per far esercitare i tempi del passato se sappiamo che una buona parte della classe delegherà il compito a un'IA, saltando un'occasione per

acquisire nuove forme verbali e non permettendo più all'insegnante di osservare progressi individuali? Ha ancora senso il tema in cui esprimere un pensiero personale se poi, delegato all'IA, l'output non rappresenta più autenticamente le posizioni dell'individuo? Si tratta dunque di ripensare le attività in modo che in esse sia riconoscibile la prestazione individuale dell'apprendente. Questo implica ripensare obiettivi, approcci e metodi di insegnamento e di valutazione – ben sapendo quali compiti è in grado di svolgere con successo la macchina, quali sono i suoi limiti e quali le potenzialità didattiche delle diverse attività di produzione testuale gestite dall'apprendente in co-costruzione con l'IA.

2. Ripensare gli obiettivi

Partiamo dai macro-obiettivi che guidano l'insegnante nel design dei propri corsi:

1. da un lato le materie scolastiche “lingue straniere” mirano ad **abilitare l'apprendente a fronteggiare con successo situazioni comunicative extrascolastiche**: è questo il principio che ha portato alla redazione del QCER e allo sviluppo della didattica orientata all'azione e, più marginalmente, della didattica interculturale;
2. dall'altro, la didattica delle lingue straniere prevede che, ai fini dello sviluppo delle suddette competenze comunicative, si miri all'**acquisizione di numerose microstrutture linguistiche**, perlopiù grammaticali e lessicali. All'apprendente si propongono ricorrenti occasioni per “allenarle” e automatizzarle – dagli esercizi *drill*, alle attività di vocabolario e alla recitazione di dialoghi preconfezionati.

Entrambi questi ambiti di azione didattica sono tuttavia spesso “oscurati” da un altro obiettivo meno evidente delle lezioni di lingua, dettato dal “sistema scuola”: pensiamo alle non poche attività didattiche sviluppate in un dato formato primariamente per soddisfare la necessità dell'insegnante di assegnare dei voti, e ai non pochi sforzi intrapresi dall'apprendente per acquisire selettivamente le competenze parziali necessarie per una valutazione positiva. È a questo micro-livello istituzionale che si gioca attualmente la partita dell'accettazione o

“L'uso competente di traduttori e generatori di testo non è affatto scontato né in apprendenti né in insegnanti: richiede abilità complesse. Si tratta dunque di ripensare le attività in modo che in esse sia riconoscibile la prestazione individuale dell'apprendente.”

meno, da parte della scuola, dei generatori di testo. Occorre tuttavia mettere da parte per un momento la discussione sull'(ab)uso dell'IA nei contesti valutativi e considerare l'effetto della tecnologia anche ai fini del raggiungimento dei due macro-obiettivi sopra descritti: l'uso mirato dell'IA può sostenere il macro-obiettivo 1 (abilitare l'apprendente a fronteggiare con successo situazioni comunicative extrascolastiche) e/o l'obiettivo 2 (favorire l'acquisizione delle microstrutture linguistiche necessarie a farlo)?

2.1 L'IA che porta a distinguere tra «competenza linguistica autonoma» e «competenza linguistica assistita»

Per quanto riguarda il macro-obiettivo 1, ovvero quello di abilitare l'apprendente a saper fronteggiare con successo situazioni comunicative reali, si fa sempre più frequente la voce di chi constata che anni di apprendimento linguistico e ingenti investimenti non sempre portano ai risultati desiderati (SWR 2023, Elmiger 2021) e ricorreranno quindi a software di traduzione automatica o generazione di testo. Sta vacillando allora la validità di una “didattica orientata all'azione”? Sì, secondo Berthele e Udry (2023): [es ist] “nicht auszuschliessen, dass Fremdsprachenunterricht im Rahmen solcher Debatten ganz grundsätzlich in Frage gestellt werden wird, denn das zentrale Ziel des kommunikativen Handelns mit Anderssprachigen kann auch via Übersetzungsmaschinen erfolgen”. Berthele e Udry arrivano a questa conclusione dopo aver analizzato quali (e quanti!) obiettivi didattici degli attuali piani di studio delle scuole dell'obbligo in Svizzera tedesca siano, già nel 2023, raggiungibili senza competenze linguistiche individuali, grazie a generatori di testo – e quanti lo saranno presto. Interrogarsi sulla pertinenza

degli attuali obiettivi comunicativi nei curricula delle lingue straniere porta a chiedersi quali competenze l'individuo debba ancora sviluppare per comunicare con successo – da solo e con l'aiuto di risorse esterne.

A nostro avviso è qui che va introdotto un distinguo: quello tra **competenza linguistica autonoma** – il saper raggiungere obiettivi comunicativi senza aiuti esterni, né di insegnanti, di amici, di traduttori o di generatori di testo – e quella di **competenza linguistica assistita** ovvero quella capacità di “fare lingua” facendo ricorso alle tante risorse che il mio ambiente mi offre e che mi possono anche permettere di esprimermi a un livello superiore rispetto alla comunicazione autonoma.

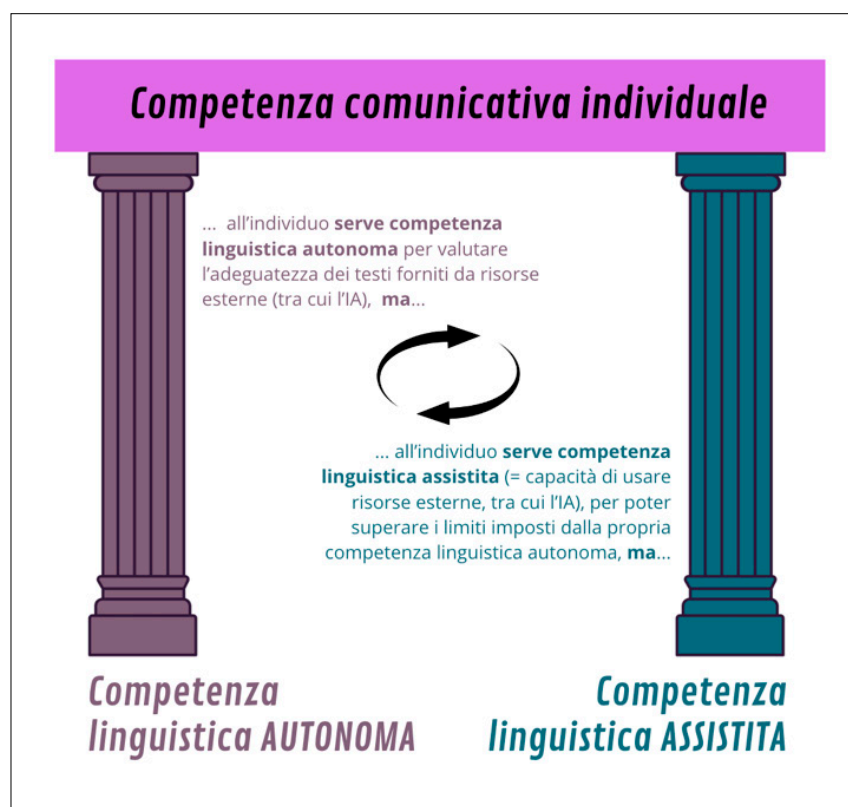
Sviluppare competenza linguistica assistita dovrebbe allora essere un ulteriore obiettivo della scuola? Secondo noi sì, in particolare considerando che

- questi strumenti non sono destinati a sparire, ma a restare, e a migliorare;
- l'apprendente in contesto scolastico oggi ne fa, a volte, un uso poco riflettuto;
- chi comunica in lingua straniera in contesti reali, oggi spesso fa uso di traduttori automatici e generatori di testo con lo scopo di superare i limiti imposti dalle proprie competenze linguistiche autonome;
- gli strumenti di IA hanno il potenziale per raggiungere meglio i propri obiettivi comunicativi se usati in modo competente e al contempo rischiano, se usati senza le competenze necessarie, di minare la comunicazione stessa;
- l'uso competente di traduttori e generatori di testo non è affatto scontato né in apprendenti né in insegnanti: richiede abilità complesse;
- la capacità di sfruttare al meglio le risorse disponibili è una premessa per il *life long learning*.

Come sviluppare allora la competenza linguistica assistita? Si tratta di una competenza complessa che comprende

- a. strategie di autoregolazione: “Quando sono più efficace usando risorse esterne e quando no?”;
- b. competenza nell'uso di funzioni specifiche degli strumenti: “So quale software e quale sua funzionalità potrebbe produrre ciò che mi serve?”;
- c. spirito critico e competenza linguistica autonoma per verificare la qualità dell'output: “sono in grado di valutare la qualità dei testi prodotti dall'IA?”;
- d. competenza nelle procedure di co-costruzione di testi con l'IA: “so selezionare le parti di output di cui ho bisogno riunendole in un testo coeso che rappresenta i miei pensieri e nel quale, pertanto, si sente la mia voce?”.

Diverse “ricette” didattiche, che mirano a raggiungere gli atteggiamenti, le competenze e le sensibilità sopra descritti, vengono attualmente sperimentate da insegnanti che intuiscono il bisogno di sviluppare la “competenza linguistica assistita” (Alloatti & Martini 2021), ma nei manuali di lingue straniere perlopiù (ancora) mancano. Speriamo tuttavia che per



chi scriverà corsi di lingua attualmente e nel prossimo futuro la domanda-guida seguente sia considerata prioritaria:

Domanda-guida 1

Come allenare l'apprendente a valutare, di caso in caso, quale sia l'uso (o il non uso) dell'IA più proficuo ai propri obiettivi comunicativi e, successivamente, a farne uso in modo veramente competente?

2.2 L'IA che mina o sostiene i processi di apprendimento di una lingua

Per quanto riguarda il secondo macro-obiettivo delle lezioni di lingua, ovvero quello di fornire occasioni frequenti ed efficaci per osservare, esercitare e produrre strutture linguistiche nuove, la prospettiva sull'uso dell'IA è differente. Nelle nostre classi e in quelle dei colleghi e delle colleghe con cui cooperiamo, l'uso di traduttori e di generatori di testo è frequente: l'IA è usata per esempio per svolgere compiti a casa, per scrivere tesi, per preparare presentazioni, ma anche per rispondere oralmente a domande che l'insegnante pone in classe. Osserviamo che la cooperazione o la delegazione di determinati compiti all'IA può avere due effetti opposti: da un lato può minare il processo d'apprendimento, dall'altro può sostenerlo.

L'uso di strumenti di IA indebolisce il processo di apprendimento, e quindi lo sviluppo di competenza linguistica autonoma, quando implica uno *skill skipping* (un "saltare lo sviluppo di una competenza", vedi Nuxoll 2024): lo *skill skipping* può essere dovuto alla riduzione del tempo di esposizione alla lingua straniera e dalla riduzione della qualità e dell'intensità dei processi cognitivi annessi – tra cui il *noticing* (la presa di coscienza o l'analisi di una data struttura linguistica) e dell'*involvement load* (grado di coinvolgimento individuale nello svolgimento dei *task*, cfr. Berthele 2021). La problematica non si limita all'ambito delle lingue straniere: anche in altre materie si è dimostrato che, in determinate condizioni, l'uso di strumenti di IA porta a saltare fasi essenziali per l'acquisizione delle competenze di base, inducendo una forma di pensiero pigro e impedendo all'apprendente di sviluppare abilità fondamentali (Bastani et al. 2024).

Al contrario, l'uso di strumenti di IA può favorire uno *skill enhancement*, sia guidato dall'insegnante e focalizzato su determinate strutture o funzioni comunicative sia invece personalizzato, "caotico" rispetto alla progressione prevista (spesso guidato da un effetto *serendipity* che porta l'apprendente a trasformare *input* linguistico prodotto dall'IA in *intake* laddove una struttura corrisponde ai propri bisogni o interessi). Ecco un esempio per entrambi i processi di *skill enhancement*, con due scenari sperimentati nelle nostre classi:

- acquisito il presente indicativo, chiediamo a ogni apprendente di preparare, con l'aiuto dell'IA, una dozzina di domande per un'intervista a compagni e compagne su eventi passati (p. es. "Hai già vinto una gara?"). Il setting prevede che le domande vadano non solo lette, ma anche ascoltate con l'auricolare e poi poste nel modo più fluente possibile: in tal modo abbiamo favorito un buon *involvement load* e una triplice esposizione all'*input* linguistico, a cui si aggiunge il coinvolgimento emotivo dovuto al poter parlare di sé stessi/e e al soddisfare una propria curiosità. Abbiamo osservato che, trattando il passato prossimo esplicitamente la settimana dopo, l'uso dell'IA sembra aver avuto un valore propedeutico e facilitare l'acquisizione delle nuove forme verbali.
- nel corso di un progetto di visione individuale di video su Netflix supportato dall'*add-on Language Reactor*, l'apprendente sceglie, sull'arco di diversi mesi, dei video da visionare e da documentare in un portfolio (con riassunti, selezione di parole chiave, ecc.). Il portfolio serve anche per consigliare a compagni/e la visione di determinati contenuti. Al termine della fase di visione è previsto un esame orale su alcuni video scelti dall'apprendente. Questo momento di valutazione sommativa ci ha sorpreso: gli/le apprendenti erano in grado di raccontare e analizzare i propri video con una fluenza notevole; molti/e di loro dimostravano di aver acquisito il vocabolario specifico necessario.

Questi, e altri scenari didattici che fanno ricorso all'IA in quanto risorsa a supporto dell'apprendimento linguistico e che sono disponibili tramite il codice QR 1 a margine, mostrano quanto, al momento,

si stia sperimentando nella prassi. Al momento, è soprattutto grazie al confronto con colleghi e colleghe che noi insegnanti di lingua straniera possiamo ampliare il nostro repertorio didattico, riprendendo e adattando scenari che si sono, in contesti analoghi ai nostri, rilevati efficaci. La domanda-guida che può sostenere la costruzione del nostro curriculum in classe o la progettazione di corsi di lingua può riassumersi così:

Domanda-guida 2

Come integrare gli strumenti di IA nelle attività didattiche da proporre all'apprendente affinché gli strumenti stessi favoriscano (e non rallentino) i processi di apprendimento della lingua – e con essi lo sviluppo di competenza linguistica autonoma?

2.3 L'IA che richiede competenze trasversali

Se dalle considerazioni fatte finora si deducono le domande-guida 1 e 2 che possono guidarci, in quanto insegnanti di lingua straniera, nella pianificazione didattica specifica alla nostra materia, occorre anche considerare quanto l'uso dei generatori di testo costituisca per l'apprendente una sfida che richiede lo sviluppo di competenze e di spirito critico che prescindono quindi dalla singola materia. La domanda-guida 3 ci aiuta a focalizzarci sullo sviluppo di tali competenze:

Domanda-guida 3

Come portare l'apprendente a sperimentare sulla propria pelle il potenziale e i rischi dei generatori di testo, indipendentemente dalle specifiche

esigenze delle materie lingue straniere, in modo da sviluppare nei loro confronti un'attitudine costruttivamente critica?

La domanda-guida 3 si focalizza su competenze trasversali, a monte delle lingue straniere: ogni disciplina può contribuire a svilupparle, e per ogni disciplina sono importanti. Il manuale che ha ispirato questo articolo, *Es macht klick* (e la sua versione in italiano "Io scrivo con l'IA"), si pone proprio questo obiettivo e porta ad allenare le competenze trasversali in 26 unità didattiche: la descrizione della struttura del volume e delle singole unità si trova online (cfr. codice QR2). Tutte mirano a far luce su diversi processi di co-costruzione di testi prodotti con l'aiuto dell'IA. Ma cosa ne è della capacità di fare lingua da soli/e, senza nessun aiuto? Con il dilagare di strumenti di IA se ne perderà il valore?

3. L'autenticità al centro

Senza entrare nei dettagli delle tante discussioni che mettono in questione la legittimazione stessa delle lezioni di lingua, va sottolineato quanto la competenza linguistica autonoma sia e resti

- necessaria a valutare ogni output di strumenti di IA in modo autonomo: senza competenza autonoma, infatti, non è possibile riconoscere dove l'IA sbaglia;
- necessaria a esprimere la propria identità nei confronti nei confronti di un'alterità con le sue convenzioni linguistiche e le specificità culturali;
- ragione di orgoglio e di fonte di gratificazione, allorché siamo autosufficienti nel comunicare con l'altra persona;
- un ingrediente essenziale per avviare comunità di lingue e culture differenti, favorendo l'intercomprensione, il rispetto e la tolleranza.

In altre parole: la competenza linguistica autonoma è indispensabile per risultare autentici nei confronti dei destinatari dei nostri messaggi e per percepire l'interlocutore o l'interlocutrice come autentici. Per capire il concetto di "autenticità" nell'ambito comunicativo è utile rifarsi a studi di psicologia, per esempio quelli di Deci e Ryan nell'ambito della teoria dell'autodeterminazione: l'autenticità riguarda il bisogno di sentirsi agenti delle

“La competenza linguistica autonoma è indispensabile per risultare autentici nei confronti dei destinatari dei nostri messaggi e per percepire l'interlocutore o l'interlocutrice come autentici.”

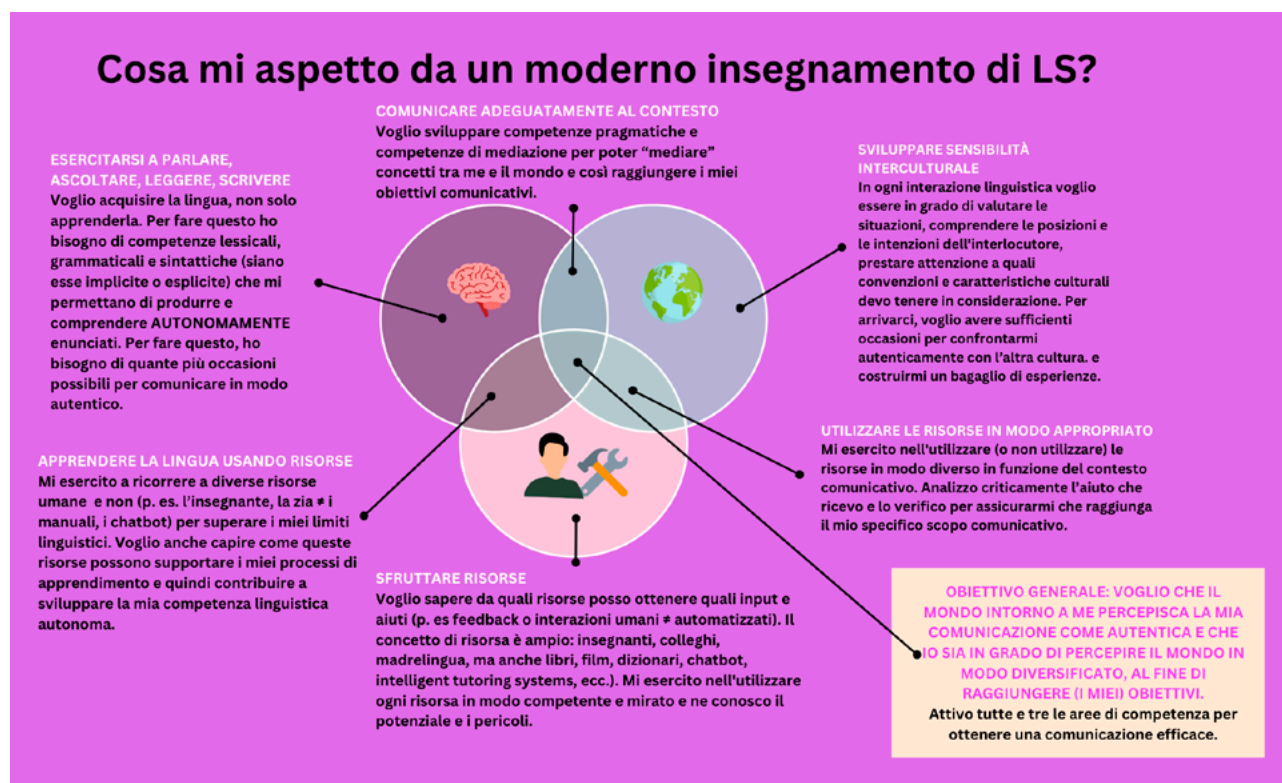
proprie affermazioni, e non di essere controllati da altre entità. Anche il bisogno di sentirsi competenti implica autenticità: se il messaggio è mio e rappresenta la mia identità, e so che sono stato/a in grado di formularlo così che sia percepito come tale anche da chi mi ascolta o legge, mi sento competente; oppure, se percepisco l'interlocutore o l'interlocutrice come non autentico/a, non riuscirò a instaurare una vera connessione o un legame, e la comunicazione sarà inefficace, superficiale o addirittura raggiungerà l'effetto contrario. Il rischio di non risultare autentici/che usando l'IA si contrasta solo se si dispone di sufficiente competenza autonoma.

Ecco allora che una buona competenza linguistica autonoma (cosa so dire, scrivere o capire da solo/a) è premessa per sviluppare anche una buona competenza linguistica assistita (cosa so dire, scrivere o capire grazie ad aiuti esterni). Entrambe non sono tuttavia sufficienti a comunicare in modo efficace se non è

presente quella sensibilità interculturale che permette di valutare in quale contesto quale tipo di comunicazione garantirà una comunicazione efficace.

Partendo da queste considerazioni ci siamo chieste cosa dovremmo aspettarci, oggi, da un insegnamento delle lingue straniere in contesto scolastico che considera l'entità dell'innovazione "intelligenza artificiale". L'IA amplia il catalogo di risorse a cui possiamo ricorrere, ma pone un nuovo accento sulla qualità della comunicazione: quella che sempre più sentiamo "finta" (pensiamo agli innumerevoli siti web autogenerati dall'IA che rifuggiamo perché sentiamo "falsi") e quella che invece vogliamo continuare a sentire "vera", "autentica".

Il seguente schema riassume queste aspettative e introduce alcune nuove relazioni che approfondiamo di seguito:



“Una buona competenza linguistica autonoma (cosa so dire, scrivere o capire da solo/a è premessa per sviluppare anche una buona competenza assistita (cosa so dire, scrivere o capire grazie ad aiuti esterni).”

Come si può notare, nello schema l'insegnante di lingua straniera è considerato una "risorsa"; un'altra risorsa può essere una chatbot, una zia, un libro... Il termine "risorsa" non definisce affatto la qualità del loro apporto, ma è stato scelto per rappresentare ogni aiuto esterno come un *node* (cioè una fonte di informazione) all'interno del paradigma del connettivismo. Il connettivismo già nel 2005 anticipava quale sarebbe stato l'apporto del *machine learning* ai processi di apprendimento umani: i suoi principi si rivelano adattissimi per sostenere quel "ripensare gli obiettivi dell'insegnamento (linguistico e non)" di cui abbiamo parlato all'inizio dell'articolo. Il connettivismo definisce infatti come obiettivo prioritario dell'apprendimento lo sviluppo di un pensiero umano elastico e critico, capace sia di continuare ad imparare connettendo fonti (*nodes*) di informazione e expertise che rappresentano diverse posizioni o opinioni, sia capace di prendere delle decisioni oculate basandosi su un sapere attuale e accurato, ma in costante evoluzione.

Ecco i principi del connettivismo (Siemens 2005):

1. Learning and knowledge rests in diversity of opinions.
2. Learning is a process of connecting specialized nodes or information sources.
3. Learning may reside in non-human appliances.
4. Capacity to know more is more critical than what is currently known.
5. Nurturing and maintaining connections is needed to facilitate continual learning.
6. Ability to see connections between fields, ideas, and concepts is a core skill.

7. Currency (accurate, up-to-date knowledge) is the intent of all connectivist learning activities.
8. Decision-making is itself a learning process. Choosing what to learn and the meaning of incoming information is seen through the lens of a shifting reality. While there is a right answer now, it may be wrong tomorrow due to alterations in the information climate affecting the decision.

L'IA sta trasformando il mondo, e il connettivismo sembra essere il paradigma che meglio ne coglie l'impatto. L'IA rende accessibili a un ampio pubblico testi non sempre accurati, ma capaci di influenzare le decisioni e le azioni delle persone. Oggi, più che mai, la scuola ha il compito di formare individui responsabili, in grado di prendere decisioni consapevoli e fondate, come sintetizzato nell'espressione *vertiefte Gesellschaftsreife* ("la profonda maturità che permette di partecipare consapevolmente alla società") dei piani di studio liceali svizzeri.

Di conseguenza, insegnare a verificare la qualità e la provenienza delle informazioni seguendo i metodi validati dalle rispettive discipline scolastiche è essenziale, anche di fronte alla crescente tentazione di delegare questo compito all'IA, che impara a compiere operazioni sempre più complesse. Alcune di esse le possiamo delegare senza remore, altre invece richiedono un controllo accuratissimo. È qui, a nostro avviso, che le lingue straniere potranno fornire un apporto essenziale: grazie ad esse si impara a interpretare messaggi "stranieri" che richiedono conoscenze sulle diverse nature e culture umane, e la sensibilità per incontrare l'alterità di nuovi *nodes*. Se così sarà, l'insegnamento delle lingue straniere non potrà che dare un'importanza crescente alla componente culturale, intendendo il proprio compito come qualcosa che va ben al di là del mero imparare vocabolario, esercitarsi a produrre frasi senza errori di grammatica e a saper effettuare un reclamo per un guasto in albergo. L'obiettivo ultimo sarà invece l'incontro autentico tra esseri umani "diversi" e una comunicazione accurata in grado di superare barriere linguistiche e culturali intrinseche alla natura umana.

Appendice 1

ALCUNE PISTE PER USARE L'IA NELLE LEZIONI DI LINGUA STRANIERA

- l'uso di software (come p. es. Univerbal) per sviluppare le competenze dialogiche orali parlando con un chatbot. L'IA prende qui un ruolo di tutor personalizzato che sostiene l'apprendente, per esempio, svolgendo conversazioni e fornendo correzioni immediate;
- banalmente, l'esercitazione della propria pronuncia facendo ricorso a software con funzionalità di text-to-speech o a software di riconoscimento verbale (come p. es. integrato negli Assignments di Teams di Microsoft);
- la classica "presentazione di un tema in classe", in cui la presentazione stessa (produzione orale monologica con alto grado di pianificazione + diapositive con elementi visuali e verbali) sarà seguita da un colloquio orale con l'insegnante (produzione orale dialogica con basso grado di pianificazione);
- l'uso di strumenti di IA in grado di fornire feedback formativo: si può andare da tool come EditGPT focalizzati sulla correttezza linguistica e sulla fluenza fino a strumenti che sono in grado di valutare componimenti, come FelloFish;
- strumenti di IA in grado di assistere l'insegnante in processi di valutazione sommativa (facendo attenzione alla protezione dei dati e alle limitazioni imposte dall'AI Act europeo);
- la programmazione, da parte dell'insegnante, di "chatbot confezionati per la propria classe" (ad es. my GPTs su ChatGPT). Questi chatbot offrono un'assistenza personalizzata all'apprendente. Ne sono un esempio i chatbot che alcuni nostri colleghi stanno già sperimentando nelle loro classi: caricando su un proprio GPT un determinato testo e corredandolo di un prompt ad hoc, questo permette agli studenti e alle studentesse di porre domande sul testo o farsene porre, il tutto secondo il livello di difficoltà linguistica prestabilita dall'insegnante;
- lo sviluppo di strategie di *life long learning* delle lingue straniere, promuovendo nell'apprendente l'abitudine di fare ricorso all'IA, per esempio per

accedere all'enorme quantità di materiale autentico in lingua straniera online (testuale, audio o video). Un esempio è un *add-on* come Readlang, con cui ogni parola su un sito in lingua straniera può essere istantaneamente tradotta; oppure un software come ChatGPT for Youtube, in grado di riassumere un video prima che venga visto e quindi facilitarne la comprensione durante la visione;

- l'analisi di testi letterari sostenuta da strumenti di IA, volta a scoprire quali livelli di significato non sono accessibili alla macchina, ma all'essere umano;
- la creazione di video esplicativi con focus sulla co-costruzione dei contenuti con l'IA, con un risultato che, linguisticamente, va al di là della momentanea competenza linguistica autonoma dell'apprendente. In un tale scenario si può sviluppare il pensiero critico e la capacità di comprensione dei documenti sfruttando l'aiuto di Perplexity.ai o ChatPDF, la ricerca di ulteriori fonti affidabili (p. es. con Elicit) e usare strumenti come Napkin.ai per sintetizzare e organizzare le informazioni su slides.
- la creazione di canzoni (p. es. con Suno) per illustrare un momento importante della propria vita, i sentimenti di un personaggio letterario, le difficoltà di un'epoca, ecc.

Appendice 2

BREVE DESCRIZIONE DEL MANUALE «ES MACHT KLIKK» E DELLA SUA VERSIONE IN ITALIANO "IO SCRIVO CON L'IA"

Il manuale "Es macht klick. Künstliche Intelligenz bei schriftlichen Arbeiten clever nutzen" (Alloatti & Montemarano 2024) fornisce a insegnanti di tutte le discipline del secondario II 26 sequenze didattiche pronte all'uso in classe, modulari e non interdipendenti. Nel caso ideale, il corpo insegnante di una medesima classe si ripartisce gli scenari: se per esempio in ogni disciplina si trattano da 1 a 3 sequenze, una classe avrà modo di svolgerle tutte e 26.

Il volume è suddiviso in tre sezioni: esperimenti, lezioni e giochi di pensiero. Gli esperimenti sono sequenze brevi (30'-45') che portano l'apprendente a sviluppare *skills* nell'uso di diversi generatori di testo. Le lezioni mirano invece ad allenare degli atteggiamenti critici o disposizioni (*attitudes*) nei confronti dei generatori di testo. I giochi di pensiero, infine portano a riflettere e a valutare l'impatto dei generatori di testo sulla nostra società, confrontandolo con quello di altre innovazioni del nostro tempo. Tutte le sequenze tematizzano diverse fasi dei processi di scrittura supportati dall'IA e portano l'apprendente a scoprirne limiti e potenziali, sviluppando un'attitudine critica e una ricerca dell'autenticità e della precisione, piuttosto che della genericità e la fluenza.

Più in dettaglio, negli esperimenti si esercita l'uso di diversi traduttori e generatori di testo per:

- capire un testo o un argomento (esperimento 1);
- ampliare la propria prospettiva o circoscrivere un tema prima di redigere testi descrittivi e argomentativi (esperimento 2);
- trovare fonti e bibliografia critica (esperimento 3)
- farsi correggere errori e farsi proporre miglioramenti stilistici nei propri testi (esperimento 4);
- redigere formati testuali standardizzati, rielaborandoli per raggiungere specifici scopi comunicativi (esperimento 5);

- sviluppare nuove idee facendosi ispirare da spunti più o meno validi ai propri scopi (esperimento 6);
- ampliare le proprie competenze linguistiche sviluppando strategie per confrontarsi attivamente con le correzioni e i feedback dell'IA (esperimento 7);
- farsi mettere alla prova dall'IA chiedendole di porci domande di comprensione su un testo letto o proponendoci esercizi mirati (esperimento 8);
- migliorare le proprie strategie nel valutare criticamente l'output dei traduttori automatici (esperimento 9);
- interagire per iscritto o oralmente con l'IA simulando ed esercitando così situazioni comunicative reali (esperimento 10);

Nelle lezioni si mira a usare traduttori e generatori di testo per:

- abituarsi a vagliare la veridicità di ogni output fin nel dettaglio, coscienti del *halo effect* che ci porta invece a non farlo (lezione 1);
- acquisire competenza nella formulazione dei prompt (lezione 2);
- apprezzare l'autenticità di un testo e imparare a produrre testi che risultano autentici anche in co-costruzione con l'IA (lezione 3);
- conoscere i limiti della "logica" dell'IA e allenarsi a un'attitudine critica nei confronti di affermazioni solo apparentemente accurate (lezione 4);
- fare ricorso a diversi strumenti specializzati nello svolgimento di compiti parziali, assemblandone gli output in un testo coeso (lezione 5);
- imparare a documentarne in modo trasparente (per sé stessi e per l'insegnante) l'uso di strumenti di IA (lezione 6);
- capire quanto portano a sviluppare *lazy thinking* e con questo *skill skipping*, scegliendo consapevolmente quando evitarli (lezione 7);
- focalizzarsi sulla scelta del tono corretto in cui scrivere un testo al fine di raggiungere i propri scopi comunicativi (lezione 8);
- sviluppare un'attenzione particolare alle tipiche high risk zones dei traduttori (forme flesse, espressioni idiomatiche, bias) e un'attitudine critica sul proprio uso dei traduttori per imparare lingue (lezione 9);

- scoprire in quali rischi si incorre usando: consolidamento di bias, assenza di protezione dei dati, cognitive overload, incapacità a riconoscere fake news... (lezione 10);

Nei giochi di pensiero (che potrebbero anche essere tradotti in ogni lingua straniera, per usarle in quanto attività comunicative per il livello B2-C2), lo sguardo si apre a innovazioni passate e future:

- simulando un viaggio nel tempo, si aiuta chi ancora mai ha visto un computer a immaginare cosa implica l'avvento della scrittura non lineare al computer, la scrittura su supporti digitali e la scrittura assistita dall'IA (gioco 1);
- analizzando gli scenari utopici e distopici formulati da chi ha visto nascere e espandersi internet, confrontandoli con quelli che attualmente si formulano sull'IA (gioco 2);
- ripercorrendo come sono cambiati i processi di lettura nelle ultime tre generazioni (nonni, genitori e attuali adolescenti) e ipotizzando come cambieranno per i figli di chi oggi va a scuola (gioco 3);
- immaginando un megavirus che blocca i sistemi informatici: quali competenze ci accorgeremo di non aver più, rimpiangendole? (gioco 4);
- inventando un proprio assistente IA personale che ha lo scopo di sostenere i processi di apprendimento e acquisizione linguistica (gioco 5);
- considerando quanto stanno cambiando le forme di valutazione delle competenze (linguistiche) oggi, a causa dell'IA – e come cambieranno ulteriormente in futuro? (gioco 6).

Bibliografia

Alloatti, S., Martini, L. (2021).

Maschinelle Übersetzungstools im Fremdsprachenunterricht: Einführung mit Unterrichtsbeispielen für das Fach Italienisch. www.mut2021.ch (sito consultato l'11.10.24)

Alloatti, S., Montemarano, F. (2024). *Es macht*

klick. Künstliche Intelligenz bei schriftlichen Arbeiten clever nutzen. Bern: hep Verlag

Alloatti, S., Montemarano, F. (2025). Io scrivo

con l'IA: per un uso critico dell'Intelligenza Artificiale, Sanoma Italia, Edizioni scolastiche Bruno Mondadori e Paravia: Milano.

Bastani, H.; Bastani, O., Sungu, A., Ge, H.;

Kabakci, Ö. & Mariman, R. (2024). Generative AI Can Harm Learning (July 15, 2024). The Wharton School Research Paper, <https://ssrn.com/abstract=4895486> (consultato l'11.10.24)

Berthele, R., Udry, I. (2023). Digitale

Übersetzungsprogramme und Online-Wörterbücher im Fremdsprachenunterricht: Erkenntnisse aus einer Umfrage bei Lehrpersonen und Lernenden der obligatorischen, post-obligatorischen und tertiären Bildung, in: *Linguistik Online* 120:2

Elmiger, D. (2021). *L'enseignement des langues*

étrangères en Suisse. Enjeux et tensions actuelles. Neuchâtel: Alphil

Moerth, M. (2023). "KI – An Tools abgeben

oder selber können", Webinar al Berliner Zentrum für Hochschullehre, 5.7.2023. <https://www.youtube.com/watch?v=69faqHMTyIM>.

Nuxoll, F. (2024). KI darf das Lernen nicht

ersetzen!, in: *Die Zeit* n°45, 24.10.2024

Siemens, G. (2005). *Connectivism: A Learning*

Theory for the Digital Age. International Journal of Instructional Technology and Distance Learning, 2 https://jotamac.typepad.com/jotamacs_weblog/files/connectivism.pdf (sito consultato l'11.10.24)

SWR (2023). Kretschmann fordert stärkere

Förderung des Französischen an Schulen. SWR Aktuell, Südwestrundfunk (SWR) mit Sendung des 4.7.23, 16:00 Uhr, SWR1 Baden-Württemberg, <https://www.swr.de/swraktuell/baden-wuerttemberg/kretschmann-franzoesisch-schule-bildung-fremdsprache-technik-100.html> (sito consultato l'11.10.24)

IMPRESSUM

EDITORE

Association Babylonia Suisse

WWW.BABYLONIA.ONLINE

GRAFICA

Acadabra communication visuelle
www.acadabra.ch

CONCETTO GRAFICO

Filippo Gander | distillerie grafiche
www.distillerigrafiche.ch
filippo.gander@gmail.com

AUTORI E AUTRICI DI QUESTO NUMERO

Sara Alloatti | sara.alloatti@ife.uzh.ch
Corinne Baradel | Ca' Foscari University of Venice | corinne.baradel@gmail.com
Alice Delorme Benites | Zürcher Hochschule für Angewandte Wissenschaften | delr@zhaw.ch
Raphael Berthele | Université de Fribourg | raphael.berthele@unifr.ch
Erica Borloz | HEP Valais | Erica.Borloz@hepvs.ch
Nicola Brocca | Universität Innsbruck | nicola.brocca@uibk.ac.at
Damien Cavaleri | HEP Vaud | damien.cavaleri@hepl.ch
Anna-Maria de Cesare | TU Dresden | anna-maria.de_cesare_greenwald@tu-dresden.de
Sara Cotelli | Université de Neuchâtel | sara.cotelli@unine.ch
Catherine Ferris-Diederich | PH St. Gallen | catherine.ferris@phsg.ch
Alice Gruber | Technical University of Applied Sciences Augsburg | alice.gruber@hs-augsburg.de
Antje Kolde | HEP Vaud | antje-marianne.kolde@hepl.ch
Mara Haller Berner Fachhochschule | mara.haller@bfh.ch
Micha Hersch | HEP Vaud | micha.hersch@hepl.ch
Rebeca Iniesta-Jimenez | Universität Innsbruck | rebeca.iniesta-jimenez@uibk.ac.at
Ferit Kilickaya | Mehmet Akif Ersoy University | Ferit.Kilickaya<ferit.kilickaya@gmail.com
Carolnie Lehr | Zürcher Hochschule für Angewandte Wissenschaften | lehc@zhaw.ch
Filomena Montemarano | filomena.montemarano@fhnw.ch
Stefano Negrinelli | PH St. Gallen | stefano.negrinelli@phsg.ch
Hasti Noghrechi | Université de Neuchâtel | hasti.noghrechi@unine.ch
Meike Raaflaub | PH Bern | meike.raaflaub@phbern.ch
Anne Raveling | Universität Wien | anne.raveling@univie.ac.at
Brigitte Reber | PH Bern | brigitte.reber@phbern.ch
Holli Schaubert, University of Geneva | Holli.Schauber@unige.ch
Elizabeth Steele | Berner Fachhochschule | elizabeth.steele@bfh.ch
Elana Summers | Zürcher Hochschule für Angewandte Wissenschaften | summ@zhaw.ch
Guillaume Tschupp | HEP Valais | guillaume.tschupp@hepvs.ch
Sandrine Wild|PH St. Gallen| sandrine.wild@phsg.ch

REDAZIONE

Matteo Casoni | OLSI
matteo.casoni@ti.ch
Sabine Christopher | OLSI
sabine.christopher@ti.ch
Anna Ghimenton | Université de Grenoble
anna.ghimenton@univ-grenoble-alpes.fr
Edina Krompák | PH Luzern
edina.krompak@phlu.ch
Amelia Lambelet | HEP Vaud
amelia.lambelet@hepl.ch
Karine Lichtenauer | Université de Genève
karine.lichtenauer@unige.ch
Laura Loder-Büchel | PH Zürich
laura.loder@phzh.ch
Flavio Manetsch | PH Fribourg
flavio.manetsch@edufri.ch
Elisabeth Peyer | KFM
elisabeth.peyer@unifr.ch
Verónica Sánchez Abchi | IRDP
veronica.sanchez@irdp.ch
Ingo Thonhauser | HEP Vaud
ingo.thonhauser@hepl.ch