



**Designing a didactic approach for GenAI chats in
language teaching: AI literacy and prompt
engineering skills in GFL at university level**

Stefanie John, Oviedo, Spain

ISSN 1470 – 9570

Designing a didactic approach for AI in language teaching: AI literacy and prompt engineering skills in GFL at university level

Stefanie John, Oviedo, Spain

This study addresses the pedagogical gap in university-level German as a Foreign Language (GFL) education regarding the integration of generative Artificial Intelligence (GenAI) chats. Despite the potential of large language model chats for language acquisition, there is a paucity of structured didactic frameworks for critical AI literacy and prompt engineering. The present paper puts forward an online didactic intervention, which is being integrated into an GFL course at the University of Oviedo in Spain. The intervention combines theoretical instruction with practical, task-based workshops. Blind peer evaluation is employed to foster feedback. Preliminary findings indicate high levels of student engagement and satisfaction.

Die vorliegende Untersuchung widmet sich der gegenwärtigen didaktischen Herausforderung der Integration generativer Künstlicher Intelligenz (KI) in den universitären DaF-Unterricht. Obwohl KI-gestützte Chatbots ein erhebliches Potenzial für den Spracherwerb aufweisen, existiert ein Defizit an kohärenten didaktischen Rahmenwerken zur Vermittlung kritischer KI-Kompetenzen und des Prompt Engineerings. Daher wird im Rahmen der vorliegenden Arbeit eine vollständig onlinebasierte didaktische Intervention präsentiert, die in einen DaF-Kurs an der Universität Oviedo in Spanien implementiert wird. Diese kombiniert theoretische Wissensvermittlung mit anwendungsbezogenen, aufgabenorientierten Aktivitäten. Eine anonymisierten Peer-Evaluation wurde durchgeführt. Die ersten Ergebnisse weisen auf ein ausgeprägtes Engagement und eine hohe Zufriedenheit der Studierenden hin.

1. Introduction

The rapid advancement of artificial intelligence (AI) is transforming pedagogical approaches across various disciplines, with foreign language education being no exception (Tekin 2023). Specifically, the integration of generative AI-driven large language model (LLM) chatbots, such as ChatGPT, has emerged as a significant technological development in language learning environments (Klimova 2024). While traditional AI encompasses a broader range of capabilities, including analysis, prediction and decision-making based on existing data, generative AI (GenAI) is a subset of artificial intelligence powered by large-scale pre-training (e.g., LLM) and specifically designed to produce novel content, such as text, images, or audio (College of Education 2024; Ye et al. 2025).

The integration of GenAI chats into foreign language instruction seeks to leverage its capabilities to personalize and enhance the learning experience. Meta-analyses indicate that in the context of English as a Foreign Language instruction, the utilization of AI applications, particularly GenAI chatbots, is on the rise in university foreign language education (AlTwijri & Alghizzi 2024) and to have the potential to enhance language proficiency, motivation, and self-assurance, though it should be noted that research in this domain is still in its nascent stages (Cisłowska & Pena-Acuna 2024; Du & Daniel 2024). In contrast, research in the field of German as a Foreign Language (GFL) remains limited (Tekin 2023), with preliminary experiences concerning university-level learning supported by GenAI chats (Fleischmann 2023). The purpose of this article is to put forward and provide a robust justification for a didactic intervention, based on a sound theoretical foundation, that enables students to effectively and autonomously engage with GenAI chats by fostering critical AI literacy and prompt engineering skills in the teaching of GFL at university level, whilst simultaneously promoting foreign language proficiency. The theoretical framework underpinning this approach will be expounded in the subsequent chapter.

2. Theoretical framework: GenAI chat integration in foreign language education

The integration of AI into educational contexts, particularly within language learning environments, requires pedagogical considerations. The following section will first outline the potential and challenges of integrating GenAI chat programs in foreign language instruction and their implications for language acquisition, followed by an explanation of the didactic considerations regarding the intended integration into GFL instruction.

The integration of GenAI chat programs into foreign language learning can significantly enhance the learning process and linguistic competencies in several ways, for example, by facilitating independent practice and increasing language contact, providing immediate feedback and further language-related information, and enabling the simulation of realistic communication situations in the target language (Law 2024; Lyu et al. 2025). Meta-studies have demonstrated a positive impact of working with GenAI chats in English as a foreign language education on various language skills, particularly in the areas of oral proficiency, reading comprehension, writing and listening skills, vocabulary

learning, and grammar, as well as motivation and interest (Cisłowska & Pena-Acuna 2024; Du & Daniel 2024; Huang et al. 2022). Furthermore, the utilization of GenAI chats in language learning has been demonstrated to personalize the educational experience, customizing it to the individual learner's needs, learning preferences, and interests, while offering flexibility and accessibility (Alshumaimeri & Alshememry 2024; Hartmann 2021; Lyu et al. 2025). Consequently, GenAI chats have been identified as an effective tutoring tool for foreign languages (Lyu et al. 2025; Ye et al. 2025). Research has demonstrated that learners can study independently of time and place and gain low-barrier access to language practice (de Witt et al. 2020). Furthermore, GenAI chats were identified as a medium that offers a secure and more relaxed atmosphere for language practice, thereby reducing nervousness, increasing emotional comfort, self-esteem, self-confidence and potentially mitigating Foreign Language Learning Anxiety (Alsager 2024; Cisłowska & Pena-Acuna 2024). Additionally, GenAI chats have the potential to provide learners with enhanced opportunities to learn autonomously, organize their own learning process, become aware of their errors, and receive feedback, which can promote their metacognitive skills and support them in their autonomous learning (Cisłowska & Pena-Acuna 2024; Hoffmann et al. 2024; Li et al. 2024). These skills are of particular significance in the constructivist paradigm of learning (Hatmanto & Sari 2023) and are indispensable for continued language acquisition within a lifelong learning framework (Alshumaimeri & Alshememry 2024).

The incorporation of GenAI chats into language learning also aligns with the learners' lived reality, as it is already an integral part of it, thereby meeting the pedagogical demand for learner orientation (Fleischmann 2023; Köbis 2023). Concurrently, this prepares learners for the digital future: In the contemporary era of increasing digitalization in the workplace, the capacity to employ AI tools has become imperative (André, Bauer 2021). Engaging with GenAI chats has been demonstrated to enhance digital competence and to impart essential skills for future professional careers (Hartmann 2021). It is also important to note that the ability to critically evaluate and question the responses generated by AI is a crucial aspect of developing critical AI literacy (Makeleni et al. 2023). AI literacy, as defined by Hartmann (2021), denotes the capacity to utilize AI-based technologies with assurance and to critically assess their

capabilities and constraints. The ability to write effective prompts, known as prompt engineering, constitutes a crucial building block (Knoth et al. 2024).

In conclusion, the integration of GenAI chat programs presents numerous pedagogical benefits for foreign language teaching. However, it is equally important to consider their inherent limitations and potential drawbacks, such as the occasional generation of grammatical errors and inappropriate responses, which call into question their positive effects on the language learning process, motivation, and anxiety (Cisłowska & Pena-Acuna 2024). Beyond these immediate concerns regarding linguistic output and learner affect, several critical aspects demand attention when considering GenAI integration. These include challenges related to data protection (e.g. the secure handling of unencrypted data), issues of liability stemming from AI-generated *hallucinations* or inappropriate responses, and broader ethical considerations such as bias (e.g., the reproduction of prejudices, clichés, and hateful content due to biased training data). Furthermore, in the context of contemporary educational disparities, the integration of GenAI raises concerns about digital divides, potentially widening competence gaps due to varying financial access to more advanced chat versions. Finally, the prospective impact on the labor market, particularly the potential future elimination of teaching positions, also warrants meticulous examination (Fleischmann 2023; Klimova et al. 2022; Law 2024; Stahl & Eke 2024; Ye et al. 2025).

Acknowledging both their identified potential and weaknesses the integration of GenAI chats into university foreign language education is not merely an option but an imperative. Indeed, higher education institutions are no longer debating *whether* to adapt to the omnipresence of GenAI, but *how* to do so (Hochschule Emden/Leer 2025). This shift is particularly pressing given that students are already making use of GenAI tools – primarily as *ghostwriters* for written assignments (Fleischmann 2023). This practice underscores a key issue: although contemporary students are often referred to as *digital natives*, this label does not necessarily imply that they are capable of effectively self-regulating their learning with GenAI tools, especially in the domain of language acquisition (Hoffmann et al. 2024). As such, explicit instruction and guidance in the educational use of GenAI chats are essential in language teaching, as the mere implementation of GenAI chats does not automatically lead to the desired learning outcomes. These outcomes are highly contingent upon the learner's ability to engage in competent interaction with the AI, which includes the skillful formulation of prompts – i.e. user

inputs that guide the AI's responses (Hartmann & Möller 2025; Zheldibayeva 2025). Importantly, learners also need to understand that GenAI chats are not limited to functioning as *ghostwriters*. As shown, they can serve as multifunctional tools for language learning, supporting the creation of individualized learning paths, offering explanations, generating varied exercises, providing feedback, and suggesting strategies for improvement. Fostering critical and productive engagement with GenAI tools in language learning therefore constitutes a central challenge and responsibility for higher education institutions today (de Witt et al. 2020).

The integration of GenAI chats into GFL instruction, as envisaged in the present intervention, requires a structured didactic approach. As previously mentioned, numerous meta-studies concerning the pedagogical use of GenAI chat tools have been conducted in the context of English as a Foreign Language (Alshumaimeri & Alshememry 2024; Cislowska & Pena-Acuna 2024; Du & Daniel 2024; Huang et al. 2022). In contrast, there is a paucity of comprehensive research in the field of GFL (Ciężka 2024). Köbis (2023) draws attention to a significant pedagogical gap that still needs to be addressed. The author further advances the argument that the integration of AI into GFL instruction constitutes an inherently interdisciplinary endeavor, situated at the intersection of linguistic, pedagogical, and AI-specific domains (see Fig. 1).

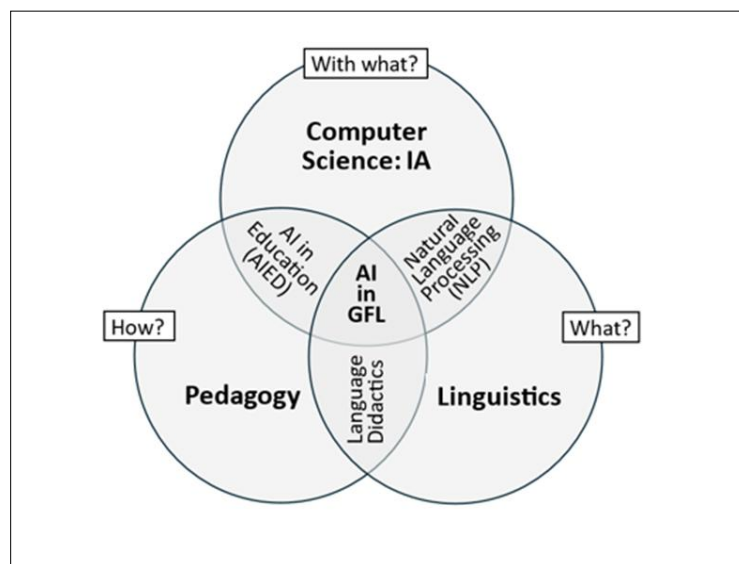


Figure 1. AI in language didactics: An interdisciplinary perspective (Köbis 2023: 34)

This model can also be transferred to the didactic context: the integration of GenAI chats into GFL instruction must consider linguistic, AI-related, and pedagogical dimensions. In practical terms, this entails clearly defining the foreign language content and AI competencies to be developed, as well as the pedagogical methods to be employed. A careful alignment and synthesis of these elements is essential to ensure coherent instructional design and effective learning outcomes. In consideration of the foregoing, and with the objective of contributing to the development and practical implementation of a structured didactic model for the implementation of GenAI chats in university GFL teaching, the intervention presented in this article was designed.

3. Didactic Intervention

This section presents the design of the didactic intervention, building on the theoretical considerations outlined above regarding the integration of GenAI chats into GFL instruction. The following subsections present the pedagogical rationale that informs the intervention in combination with the selected didactic instruments, and the organization of the learning activities.

3.1 Pedagogical Rationale of the Didactic Concept

In applying the presented model of AI in the field of language didactics, the didactic objectives and methods were defined with regard to the three constituent components – namely, AI itself, linguistics, and pedagogy – and integrated into a coherent overall concept in order to achieve the primary pedagogical objective of this intervention, which was to cultivate linguistic proficiency and critical AI literacy among university-level students of GFL in combination with language development. This combination of objectives is particularly suitable, as research has demonstrated that AI-based tasks and projects in language teaching are especially sustainable when they support not only linguistic competence but also a critical and reflective engagement with AI tools (Hartmann & Möller, 2025).

Regarding the linguistic element, topics from the course curriculum were expanded and elaborated to meet institutional requirements and to ensure the relevance of the learning content for students. In relation to AI, the intervention sought to cultivate AI literacy by integrating fundamental concepts, potential applications, and limitations of AI with the practical utilization of GenAI chats for GFL learning across a range of activities. This

approach was complemented by a critical examination of AI-generated content, structured in a gradual and cumulative manner. The development of prompt engineering skills was identified as a fundamental aspect of this transformation, with these skills forming the basis for productive and qualitatively demanding learning with GenAI chats (Knoth et al., 2024). Additionally, an understanding of the variety of possible learning activities with such tools was emphasized.

From the didactic-pedagogical perspective, innovative teaching methods were employed, as Aboura (2024) has recommended based on a series of research papers on AI-supported education. To facilitate this, online instruments were integrated into the course's Moodle platform, thus enabling students to work independently on the topic. Consequently, the integration of generative AI into conventional university-level language instruction can be achieved without necessitating any alterations to its fundamental structure. This pedagogical concept, which is based on a proven model for promoting learner autonomy in university-level GFL online contexts (John, 2025), can be easily transferred to other learning environments.

The following didactic instruments were selected for the intervention:

Moodle learning environment: The intervention was centrally hosted on the university's Moodle platform. Moodle facilitates autonomous learning through customizable learning pathways and collaborative activities (Godwin-Jones 2011), while also fostering social interaction and student engagement (Kunze & Frey 2021).

Moodle SCORM package: Introductory content on AI fundamentals was provided through a SCORM package. SCORM (Sharable Content Object Reference Model) packages are standardized, trackable learning units, such as texts, quizzes, and videos, which can be integrated into Moodle as ZIP files. These tools facilitate the monitoring of learner progress and engagement (https://docs.moodle.org/500/en/SCORM_activity). The SCORM packages were created using *eXeLearning*, an open-source authoring tool for designing structured, multimedia-rich modules without the necessity of programming knowledge. Given its capacity for SCORM export, *eXeLearning* is especially well-suited for the creation of reusable and interactive digital content (<https://exelearning.net/en/>).

Moodle Workshop: The Moodle *Workshop* constitutes a peer-assessment activity, in which students submit their work and evaluate that of their peers using predefined rubrics provided by the instructor. The process entails the phases of submission, assessment and grading. Students are graded on both their own work and the quality of their peer assessments (https://docs.moodle.org/500/en/Workshop_activity). Peer assessment has been identified as a valuable didactic instrument for promoting self-regulated learning by enhancing awareness of criteria, developing evaluative skills, fostering social responsibility, and increasing motivation (Winter, 2007; Hiltmann et al., 2019). The use of rubrics as transparent instruments for formative evaluation has been demonstrated to foster the development of evaluative skills and critical thinking, while concurrently ensuring transparency and learner autonomy by clarifying performance expectations (Kreuzer, 2018; Lissmann, 2007). Moreover, rubrics have been found to contribute to the structuring and

objectification of assessment processes (Hiltmann et al., 2019). Furthermore, rubric-guided feedback has been evidenced to foster critical reflection and metacognitive awareness (Schmoll & Braun, 2019).

Based on the fundamental decisions previously described, which related to both linguistic content and AI-related content, as well as the didactic instruments which were selected, the intervention was designed and implemented.

3.2 Structure of the Intervention

The didactic intervention, a semester-long project, was implemented during the spring term (January to May) of 2025 within the context of an undergraduate beginner-level GFL course (*Lengua Alemana II* of the bachelor's degree program *Lenguas Modernas y sus Literaturas*) at the University of Oviedo, Spain. The bachelor's degree in *Modern Languages and their Literatures* is a four-year program (240 ECTS credits) that provides students with a solid foundation in language, literature, culture, and linguistics, while allowing specialization in two languages. Students who are learning German enter the course with no previous knowledge of the language and are expected to reach level B2 of the Common European Framework of Reference for Languages (CEFR) by the time they graduate. The course in which the intervention was conducted was in the second year of study, designed for students at the A2.2 CEFR level, and convened for three instructional hours per week. Attendance was not compulsory. A total of 16 students were enrolled in the intervention, with 14 successfully completing it.

The didactic intervention was implemented entirely on Moodle in the form of homework tasks, operating independently from but in close alignment with the regular academic instruction. In accordance with the adopted continuous assessment approach, successful participation in the intervention was incorporated into the comprehensive assessment portfolio. The intervention was implemented across two main phases: The first phase entailed theoretical instruction and content delivery, followed by a second phase of practical application, in which students engaged in independent work with GenAI chats.

At the beginning of the intervention, students were provided with a Moodle-based document outlining the structure and objectives of the project, in conjunction with instructions for establishing user accounts for Gemini and ChatGPT. It was also communicated to the participants that alternative platforms, such as the recently released DeepSeek, could be utilized.

The subsequent initial theoretical phase, which lasted two weeks, focused on providing a basic understanding of AI and its relevance to language learning. This was achieved by providing a SCORM package located in Moodle. The package included an introduction to core AI concepts and a comprehensive exploration of prompt engineering as a method for effectively guiding AI systems, supplemented with practical examples of AI applications in GFL learning. Consequently, students participated in an in-depth study of prompt engineering and prompt refinement to optimize the effective use of AI systems. Moreover, the package incorporated a critical examination of the limitations and potential biases inherent in AI systems, with the aim of cultivating critical AI literacy (see Fig. 2).



Figure 2. First phase of the intervention: theoretical input delivered via Moodle.

The integration of exercises to assess prior knowledge and knowledge gain, such as matching tasks and multiple-choice questions, enabled students to evaluate their understanding and monitor their learning progress (see Fig. 3).

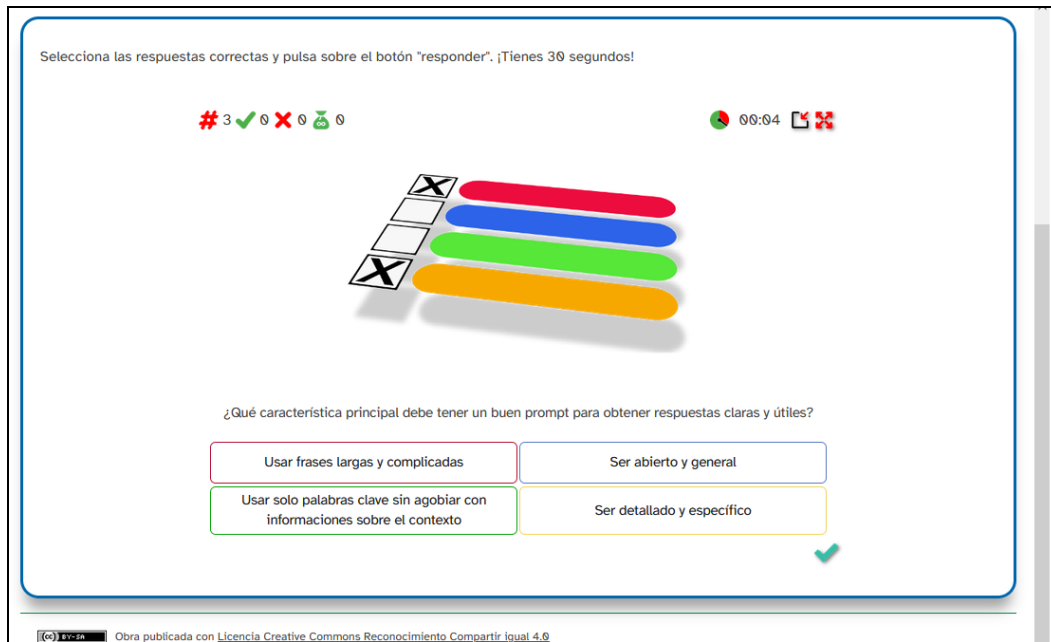


Figure 3. First phase of the intervention: multiple-choice questions.

The subsequent practical phase was structured as a recurring cycle of consecutive stages, namely independent practical work with GenAI chats, process reflection, and peer assessment. This activity provided students with opportunities to apply the knowledge acquired in the theoretical phase and was implemented through the *Workshop* assessment activity. Specifically, students progressed through the following stages:

1. *Completion of homework assignments:* These tasks consisted of written work carried out with one or more self-selected GenAI chats, with the objective of consolidating and expanding linguistic course content. The initial focus was on working with and refining predefined prompts, with a gradual increase in autonomy that culminated in students independently developing prompts tailored to specific learning objectives. These prompts were formulated in Spanish, as the students' proficiency in German was not yet sufficient for carrying out such tasks in the target language. The prompts that underwent testing, refinement, or adaptation by the students are outlined below (translation from Spanish):

- Task 1: I am learning German and want to practice comparisons with “so... wie” and the comparative + “als”. Provide five sentences with two blanks each. In each sentence, I must decide whether “so... wie” or “als” fits best (only one option per blank). Then, give me the solutions, the Spanish translation, and an explanation in Spanish of why “so... wie” or “als” is used in each case.

- Task 2a (two-part activity): I am a beginner-level German student and want to practice conjugation in the *Präteritum*. Please give me simple sentences in the *Perfekt*. I will rewrite each sentence in the *Präteritum*. If I make a mistake, respond with “Leider nicht ganz richtig” and allow me to try again. If I make a second mistake, give me a hint (without providing the correct answer) and allow me a third attempt. If I make a third mistake, provide the correct answer and then give me a new sentence in the *Perfekt*. If the answer is correct, respond with “Richtig!” and provide a new sentence in the *Perfekt*. We will continue in this manner until I say “Vielen Dank”. At the end, tell me how many sentences I completed, how many I got right on the first try, and then quiz me on the *Perfekt* verbs where I made mistakes. I will have to write each verb in the correct *Präteritum* form.
- Task 2b: You are my German teacher. Give me a topic for writing a text in German using the *Präteritum*. Then correct my mistakes, explain them, and create a short activity to help me practice.

After this first phase of working with predefined prompts, students were subsequently tasked with formulating their own prompts. Through this exercise, their proficiency in prompt construction and a range of practice activities was enhanced.

The following instructions were provided (translation from Spanish):

- Task 3a (two-part activity): Formulate a prompt that instructs the AI to play a guessing game with you using only “yes” or “no” questions. First you will think of a profession and the AI must guess it. When formulating the prompt, decide whether you want to ask the AI to show you the English translations of the questions directly or only on request.
- Task 3b: Formulate a prompt that instructs the AI to write a text about the world of work in German, adapted to your language level, and to ask you comprehension questions about the text and comment on your answers.
- Task 4a (two-part activity): Formulate a prompt that instructs the AI to create a review exercise on one of these two topics: verbs with *Dativ* + *Akkusativ* (+ *Personalpronomen*) / subordinate phrases with “wenn”.
- Task 4b: Formulate a prompt that instructs the AI to converse with you about films (written or with a conversational voice AI system), adapted to your language level. For a spoken conversation (with voice) you can use for example: ChatGPT (with voice plugins or the mobile app): the app in particular works very well without any additional plugins; Deep AI has a ‘voice chat’ option: <https://deepai.org/voice>
- Task 5: Choose a communicative situation that you would like to practice with the AI chat in preparation for the oral exam. Formulate a prompt that instructs the AI to converse with you on the chosen topic (written or – preferably! – through a conversational AI voice system), adapted to your language level. For a spoken conversation (with voice) you can use for example: ChatGPT (with voice plugins or the mobile app): the app in particular works very well without any additional plugins; Deep AI has a ‘voice chat’ option: <https://deepai.org/voice>. Tip: You can also upload a list with the vocabulary you want to practice!

In these exercises, the students were encouraged to refine their prompts and, if necessary, subsequently develop additional prompts for exercises aimed at address-

ing their errors and weaker points. In this way, both their linguistic proficiency and their confidence and aptitude in adeptly utilizing AI as a personalized learning instrument in various types of exercises were strengthened.

2. *Guided reflection*: Following the utilization of GenAI chat, students further conducted a written reflection as a homework assignment, guided by pre-defined reflection cues, addressing the following domains: the process of prompt engineering, their individual learning progress, the outcomes accomplished, and their observations and conclusions concerning future learning strategies.
3. *Submission*: The reflections and copy of the chat history were then submitted in a Word Document to the Moodle *Workshop*.
4. *Self-evaluation and blind peer evaluation*: Within the *Workshop* activity, students were assigned their own work products (chat history and reflection) as well as those of two other students, anonymized for self- and peer-assessment, which was carried out as homework. The students awarded grades based on a rubric provided in the Moodle activity, which had previously been made available to them for transparency. In this rubric, the intensity and outcomes of the work with the learning contents, the prompt refinement process, and the quality of the reflection were evaluated. This process not only cultivated critical engagement with the prompting process but also provided valuable insights into the diversity of approaches to AI-assisted language learning by allowing students to recognize other students' approaches.

The final grade for the intervention was calculated from the averages of peer- and self-assessment scores, complemented by the automatically generated measure of peer-assessment quality in Moodle. This approach ensured that evaluation reflected not only individual performance but also students' capacity for critical and constructive evaluation of their peers' work. The resulting grade was then integrated into the overall course assessment portfolio alongside other graded activities.

The integrated nature of these tasks, in conjunction with the concomitant phases of peer evaluation, established a cyclical learning process that encouraged active experimentation, critical reflection, and the application of constructive feedback for the continuous

improvement of both language and AI skills.¹ The workflow of this practical phase is visually represented in Fig. 4.

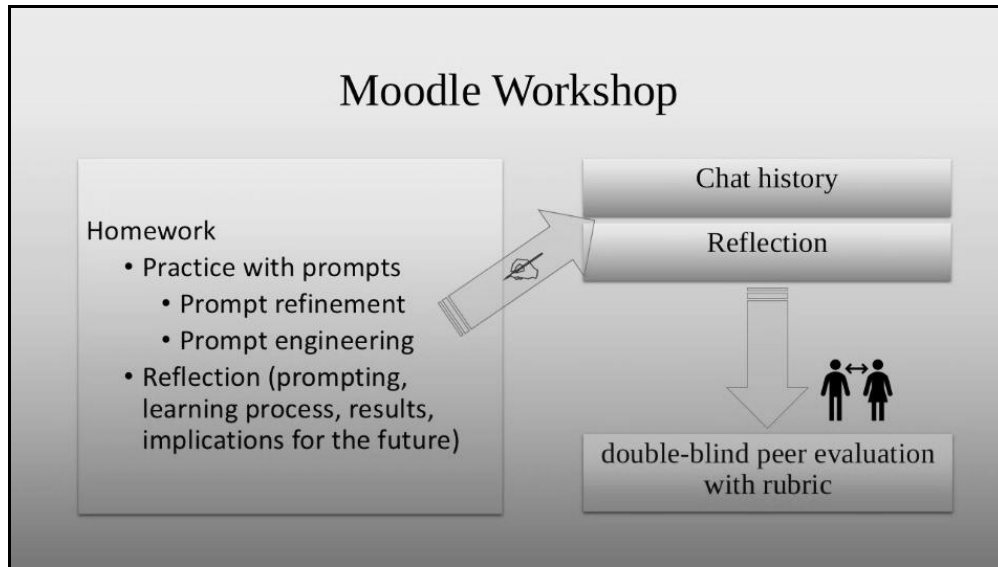


Figure 4. Second phase of the intervention: Workflow.

To encourage autonomous learning and to ensure sustained engagement with the GenAI chats, a period of one to two weeks was allocated for the independent completion of each of the five learning tasks. After submission, a one-week phase of self- and peer-assessment followed. Table 1 summarizes the specific linguistic content and AI-related activities of each task (see the above detailed formulations), along with the corresponding timeline.

Timeline	Linguistic Content	AI-related Content
3-9/2 Evaluation: 10-22/2	Grammar exercise: comparisons with <i>so... wie</i> and the comparative + <i>als</i>	Testing the provided prompt Prompt refinement: refining the prompt (e.g., with AI assistance, adding characteristics, comparing results)
24/2-8/3 Evaluation: 8-18/3	Grammar exercise: Präteritum Text writing: Präteritum	Testing the provided prompt Prompt refinement: refining the prompt (e.g., with AI assistance, adding characteristics, comparing results) Testing the provided prompt Prompt refinement: refining the prompt (e.g., with AI assistance, adding characteristics, comparing results)

¹ A project conducted by Hartmann & Möller (2025) on GenAI chats in university-level language instruction followed similar phases (pre-instruction, independent work with GenAI chats, written reflection, oral presentation of results, feedback session, and online questionnaires).

18-26/3 Evaluation: 27/3-6/4	Vocabulary: professions Reading comprehension: world of work	Prompt engineering: creating a prompt for a guessing game about professions (yes/no) and vocabulary reinforcement Prompt engineering: creating a prompt for the AI to generate a text (theme: world of work) and formulate comprehension questions
6-14/4 Evaluation: 15-23/4	Grammar exercise: verbs with dative + accusative (+personal pronouns) and/or subordinate clauses with <i>wenn</i> Oral (written) expression and comprehension: movies	Prompt engineering: creating a prompt for reviewing grammatical content Prompt engineering: creating a prompt to converse with AI about movies (in writing or using a voice-based conversational AI system)
21-27/4 Evaluation: 28/4-5/5	Oral (written) expression and comprehension: preparation for the oral exam	Prompt engineering: creating a prompt to converse with AI about course topics (in writing or using a voice-based conversational AI system)

Table 1. Second phase of the intervention: Timeline, linguistic and AI-related content.

As demonstrated by the data presented in Table 1, the intervention spanned a duration of three months, concluding at the end of the lecture period.

5. Methods of data collection and data analysis

The evaluation of the intervention's success and design was carried out through the administration of questionnaire surveys via the Microsoft Forms platform. Surveys were conducted in two phases: the first prior to the intervention (t1), and the second after the intervention (t2). The initial questionnaire (t1) was completed by all 21 enrolled students, including those who did not regularly attend lessons or the intervention. In contrast, the final questionnaire (t2) was completed by 13 of the 14 students who fully completed the intervention. The analysis was conducted exclusively on the 13 complete t1–t2 data sets, primarily relying on frequency analyses. Furthermore, the written reflections that accompanied students' work with GenAI were systematically analyzed, ensuring a comprehensive perspective on both quantitative and qualitative dimensions of the intervention.

6. Findings

All participants had prior experience of using GenAI chats prior to the intervention, with ChatGPT being the only one employed by the entire group. In addition, one participant reported using Character AI, while another opted for Consensus. Prior to the

intervention, AI was used across several domains, including leisure and entertainment (4 participants), work and productivity (7 participants), and, most prominently, academic studies and learning (12 participants). However, only eight participants had previously used GenAI chats for language learning, and all of them had relied exclusively on ChatGPT. During the intervention, students expanded their repertoire by incorporating other platforms such as Gemini, DeepSeek, and Character AI for language-learning purposes (see Fig. 5).

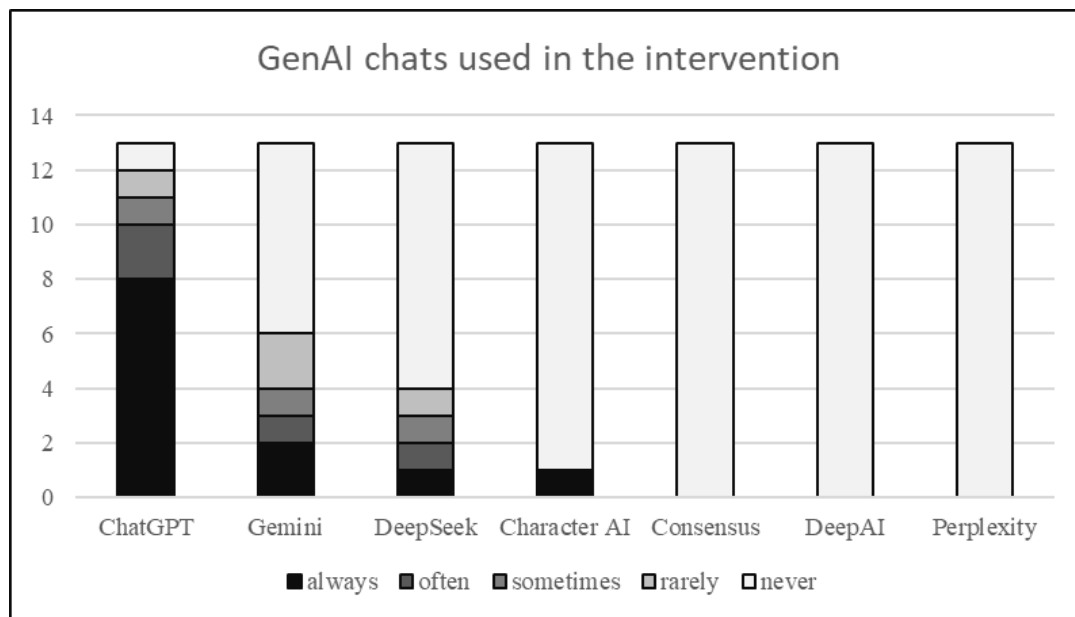


Figure 5. GenAI chat platforms used by students during the intervention (questionnaire t2, $n = 13$).

Following the intervention, most participants expressed a strong desire to continue using GenAI chats for the purpose of language learning. Eight students stated that they were certain they would continue doing so, while three reported being fairly certain. A mere two participants expressed either certainty or relative certainty that they would not persist in utilizing GenAI chats for this purpose (see Fig. 6).

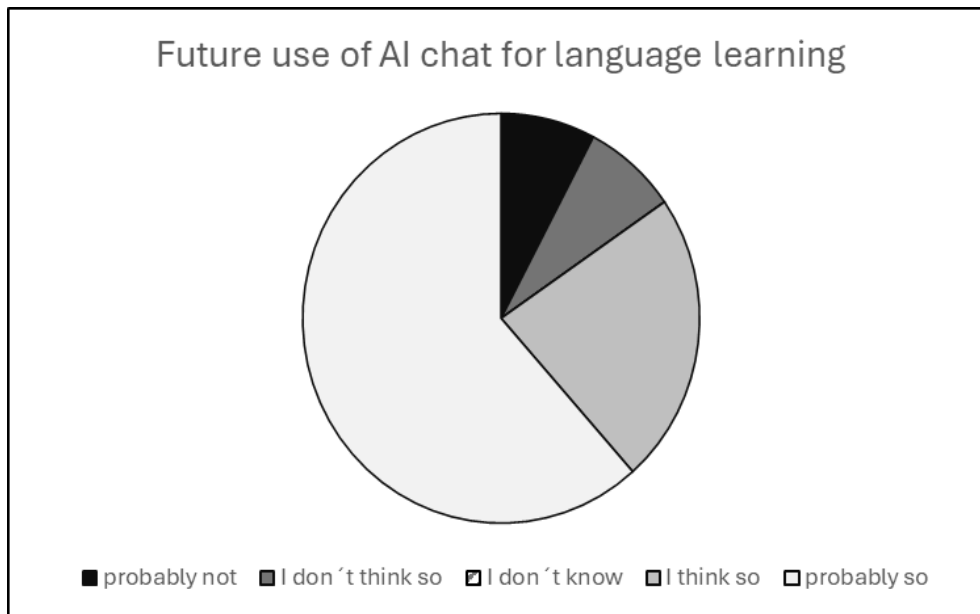


Figure 6. Future use of GenAI chat for language learning, as stated by students (questionnaire t2, $n = 13$).

Beyond their intentions, participants also reported significant improvements in their competencies related to prompt formulation, a core objective of the didactic intervention. In the initial evaluation, the participants self-assessed their ability in this domain at an average of 2.38 on a 5-point scale (1 = *not at all well*, 5 = *very well*), reflecting a general perception of limited competence. By the conclusion of the intervention, however, all students reported having acquired at least some knowledge, with none indicating that they had learned little or nothing. The highest perceived gains concerned knowledge of AI's potential for language learning ($M = 4.08$ on a scale from 1 = *no learning* to 5 = *very intense learning*), followed by prompt engineering ($M = 3.92$), prompt refinement ($M = 3.77$), and awareness of AI limitations ($M = 3.77$). Students also assigned high importance to these competencies for effective language learning, with prompt engineering ($M = 4.62$ on a scale from 1 = *not at all important* to 5 = *very important*) and prompt refinement ($M = 4.62$) rated as the most critical, followed by awareness of AI limitations ($M = 4.31$) and knowledge of AI learning possibilities ($M = 4.08$).

The quantitative findings were corroborated by the written reflections of the students. The participants demonstrated a nuanced understanding of the prompting process, their own learning trajectories, and the linguistic and AI-related outcomes achieved through their interactions with GenAI chats. Furthermore, the reflections indicate that students had reached a stage where they were able to draw actionable conclusions from their

experiences, suggesting an aptitude to apply their knowledge in future learning contexts. This assertion is supported by the following two examples. One student emphasized the perceived value of AI-generated feedback, while also contrasting and evaluating different Generative AI chats in a differentiated manner:

Above all, I would highlight the feedback I have received from DeepSeek and ChatGPT by being quite didactic and subtle with the clues they gave me when I got a verb wrong, helping me to try to figure it out myself without needing to look at PONS ... However, I have sometimes encountered serious inaccuracies in ChatGPT when it came to putting the correct form of the Präteritum.

Another student provided a balanced evaluation of the strengths and limitations of GenAI chats as a learning tool:

I consider that AI does an excellent job as a teacher for the elaboration of texts as you can ask it questions or have explained anything you want. However, I would still recommend looking at the dictionary or resorting to the help of a private tutor in some cases as artificial intelligence is still not perfect and sometimes it is the human factor that really makes the difference.

In conclusion, a synthesis of the quantitative and qualitative findings indicates that the intervention augmented students' technical competencies in prompt engineering and refinement, while also cultivating a critical awareness of the advantages and limitations of GenAI chats in language learning.

7. Discussion

The findings demonstrate that prior to the intervention, students perceived their ability to use GenAI chats for language learning as limited and engaged with these tools only to a restricted extent. Consequently, their capacity to effectively self-regulate their language learning with GenAI chats was insufficiently developed, a pattern also identified among current students by Hoffmann et al. (2024). This underscores the importance of explicit instruction in prompt engineering and critical evaluation to enable productive use of GenAI in educational contexts.

The intervention was found to be effective in enhancing students' competencies in prompt formulation, as well as their critical awareness of the affordances and limitations of GenAI chats. These outcomes corroborate claims in recent scholarship that effective language learning with GenAI chats requires the development of prompt engineering skills (Hartmann & Möller, 2025; Zheldibayeva, 2025) and critical engagement with AI-generated content (de Witt et al., 2020). Moreover, this finding is consistent with

contemporary scientific understanding of the importance of prompt engineering in such contexts (Hartmann & Möller, 2025; Knoth et al., 2024; Zheldibayeva, 2025).

Furthermore, an analysis of students' reflective accounts revealed a nuanced metacognitive engagement, whereby they identified pedagogical benefits of AI feedback and critically assessed its shortcomings. This finding aligns with the observations of Cislowska & Peña-Acuña (2024), Hoffmann et al. (2024), and Li et al. (2024), who have previously reported that GenAI-supported tasks can facilitate metacognitive skills and promote autonomous learning. The active questioning of AI responses also contributed to the development of critical AI literacy, as described by Makeleni et al. (2023). The combination of language learning with critical engagement proved especially effective, as evidenced by Hartmann & Möller (2025).

The didactic design applied in this study was derived from Köbis's (2023) interdisciplinary model, which provided a robust framework for planning and analyzing the intervention. The integration of linguistic, pedagogical, and AI-related dimensions constituted the foundation for the intervention's success in fostering meaningful GenAI chat use in foreign language education.

Nevertheless, the study's limited sample size restricts the external validity of the results. Furthermore, the heterogeneity in participants' prior experience with AI technologies may have influenced individual learning trajectories. These limitations mandate a cautious generalization of the findings.

8. Conclusion and Outlook

The present study demonstrates that explicit instruction in prompt engineering and critical evaluation can substantially enhance students' ability to use GenAI chats for autonomous and reflective language learning. It thus demonstrates a viable pathway for integrating GenAI chats into higher education language learning in a manner that fosters autonomy and reflective practice.

Future research should build on these findings by testing the approach with larger and more diverse cohorts, as well as exploring longitudinal effects on learners' self-regulation and sustained critical engagement. Further refinement of didactic models derived from Köbis's (2023) framework may also contribute to ensuring that AI

integration addresses both AI-relevant outcomes and the linguistic and the pedagogical dimensions of AI use in language education.

Bibliography

- Aboura, Bouchra (2024) Neue Perspektiven in der universitären Sprachbildung: Integration von Künstlicher Intelligenz ins DaF-Curriculum der Universität Oran 2. *Langues & Cultures* 5 (1), 230–249.
- Alsager, Harron N. (2024) To Look from Another Window in Education: Artificial Intelligence Assisted Language Learning and Its Reflections on Academic Demotivation, Foreign Language Learning Anxiety and Autonomy. *Computer-Assisted Language Learning Electronic Journal* 25 (4), 124–147.
- Alshumaimeri, Yousif A.; Alshememry, Abdulrahman K. (2024) The Extent of AI Applications in EFL Learning and Teaching. *IEEE Transactions on Learning Technologies* 17, 653–663.
- AlTwijri, Lujain; Alghizzi, Talal Musaed (2024) Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon* 10, e31053.
- André, Elisabeth; Bauer, Wilhelm (Hrsg.) (2021) *Kompetenzentwicklung für Künstliche Intelligenz – Veränderungen, Bedarfe und Handlungsoptionen*. Whitepaper aus der Plattform Lernende Systeme. https://www.plattform-lernende-systeme.de/files/Downloads/Publikationen/AG2_WP_Kompetenzentwicklung_KI.pdf, zuletzt geprüft am 03.07.2025.
- Cieźka, Agnieszka (2024) Generative KI-Tools: Die Zukunft des kreativen Lernens. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 29 (1), 375–405.
- Cisłowska, A. I.; Pena-Acuna, B. (2024) Integration of chatbots in additional language education: A systematic review. *European Journal of Educational Research* 13, 1607–1625. doi: <https://doi.org/10.12973/eu-jer.13.4.1607>.
- College of Education (2024) *Traditional AI vs. Generative AI: What's the Difference?* University of Illinois Urbana-Champaign. <https://education.illinois.edu/about/news-events/news/article/2024/11/11/what-is-generative-ai-vs-ai>.
- Witt, Claudia de; Rampelt, Florian; Pinkwart, Niels (Hrsg.) (2020) *Künstliche Intelligenz in der Hochschulbildung*. Whitepaper KI-Campus. https://ki-campus.org/sites/default/files/2020-10/Whitepaper_KI_in_der_Hochschulbildung.pdf
- Du, Jinming; Daniel, Ben Kei (2024) Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence* 6, 100230. [Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice - ScienceDirect](https://doi.org/10.1016/j.caei.2024.100230).
- Fleischmann, Andreas (2023) ChatGPT in der Hochschullehre. Wie künstliche Intelligenz uns unterstützen und herausfordern wird. *Neues Handbuch Hochschullehre* 110. <https://www.nhhl-bibliothek.de/media/263cce49cd3391acb6fb6b8ca603b68b541f1986/064f1a3e3b0b581f2835cb6852d369ad50dde3a4.pdf>.

- Godwin-Jones, Robert (2011) Autonomous Language Learning. *Language Learning & Technology* 15 (3), 4–11.
- Hartmann, Daniela (2021) Künstliche Intelligenz im DaF-Unterricht? Disruptive Technologien als Herausforderung und Chance. *Informationen Deutsch als Fremdsprache* 48 (6), 683–696.
- Hartmann, Daniela; Möller, Max (2025) Future Skills und KI-Kompetenz im Sprachunterricht. Ein Projekt zum Einsatz generativer KI im DaF-Unterricht. In: Baum, Luisa; Günay, Gülsüm (Hrsg.): *Künstliche Intelligenz in DaF/DaZ*. Berlin: Frank & Timme, 135–151.
- Hatmanto, Endro Dwi; Sari, Mariska Intan (2023) Aligning Theory and Practice: Leveraging Chat GPT for Effective English Language Teaching and Learning. *E3S Web of Conferences* 440 (05001), 1–18.
- Hiltmann, Stefanie; Hutmacher, Fabian; Hawelka, Birgit (2019) Selbstlernphasen Studierender unterstützen. In: Jahn, Dirk; Kenner, Alessandra; Kergel, David; Heidkamp-Kergel, Birte (Hrsg.) *Kritische Hochschullehre: Impulse für eine innovative Lehr- und Lernkultur*. Diversität und Bildung im digitalen Zeitalter. Wiesbaden: Springer, 305–322.
- Hochschule Emden/Leer (2025) *Umgang mit generativen KI-Tools an der Hochschule Emden/Leer. Ein Positionspapier*. Version 12. https://www.hs-emden-leer.de/fileadmin/user_upload/cd/Positionspapier_KI_Tools_V12.pdf.
- Hoffmann, Isabel; Hofmann, Katrin; Gür-Şeker, Derya (2024) *KI für AI-nsteiger - Künstliche Intelligenz im DaF/DaZ-Kontext. Ein Einführungspapier*. Version 1.0, Januar 2024. https://pub.h-brs.de/frontdoor/deliver/index/docId/7942/file/KI-Paper_Hoffmann_et_al_Version_1_0.pdf.
- Huang, Weijiao; Hew, Khe Foon; Fryer, Luke K. (2022) Chatbots for language learning – Are they really useful? A systematic review of chatbot-supported language learning. *Computer Assisted Learning* 38 (1), 237–257.
- John, Stefanie (2025) *Effekte einer Online-Intervention zur Förderung von Lernenden-autonomie und Textkompetenz im universitären DaF-Unterricht unter Einbeziehung der Studierendenperspektive*. Dissertation. Oviedo: Universidad de Oviedo.
- Klimova, Blanka; Pikhart, Marcel; Kacetyl, Jaroslav (2022) Ethical issues of the use of AI-driven mobile apps for education. *Front. Public Health* 10, 1118116.
- Klimova, Blanka; Pikhart, Marcel; Al-Obaydi, Liqaa Habeb (2024) Exploring the potential of ChatGPT for foreign language education at the university level. *Frontiers in Psychology* 15, 1269319.
- Knoth, Nils; Tolzin, Antonia; Janson, Andreas; Leimeister, Jan Marco (2024) AI literacy and its implications for prompt engineering strategies. *Computers and Education: Artificial Intelligence* 6, 100225.
- Köbis, Laura (2023) Künstliche Intelligenz in der DaF Lehre: Theoretischer Überblick und praktischer Einsatz. *German as a Foreign Language* 2023 (2), 31–53.
- Kreuzer, Pia (2018) *Kriterienraster. Handreichung der Prüfungswerkstatt*. LOB Lehren – Organisieren – Beraten. Zentrum für Qualitätssicherung und -entwicklung, Johannes-Gutenberg-Universität Mainz. https://www.zq.uni-mainz.de/files/2018/08/6_Kriterienraster-erstellen.pdf.
- Kunze, Louisa; Frey, Dieter (2021) Digitale Lehre an der Hochschule: Warum Blended Learning so gut funktioniert. In: Dieter Frey; Monika Uemminghaus (Hrsg.)

- Innovative Lehre an der Hochschule: Konzepte, Praxisbeispiele und Lernerfahrungen aus COVID-19.* Berlin, Heidelberg: Springer, 69–86.
- Law, Locky (2024) Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open* 6, 100174.
- Li, Zixi; Wang, Chaoran; Bonk, Curtis (2024) Exploring the Utility of ChatGPT for Self-directed Online Language Learning. *Online Learning Journal* 28 (3), 157–180.
- Lissmann, Urban (2007) Beurteilungsraster und Portfoliobeurteilung. In: Gläser-Zikuda, Michaela; Hascher, Tina (Hrsg.) *Lernprozesse dokumentieren, reflektieren und beurteilen: Lerntagebuch und Portfolio in Bildungsforschung und Bildungspraxis.* Bad Heilbrunn: Julius Klinkhardt, 87–108.
- Lyu, Boning; Lai, Chun; Guo, Jianing (2025) Effectiveness of Chatbots in Improving Language Learning: A Meta-Analysis of Comparative Studies. *International Journal of Applied Linguistics* 35, 834–851.
- Makeleni, Sive; Mutongoza, Bonginkosi Hardy; Linake, Manthekeleng Agnes (2023) Language Education and Artificial Intelligence: An Exploration of Challenges Confronting Academics in Global South Universities. *Journal of Culture and Values in Education* 6 (2), 158–171.
- Schmoll, Lars; Braun, Dirk (Hrsg.) (2019) *Kompetenzorientiert unterrichten – kompetenzorientiert ausbilden: Ein Kompetenzraster für die schulische Aus- und Fortbildung.* Hohengehren: Schneider.
- Stahl, Bernd Carsten; Eke, Damian (2024) The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management* 74, 102700.
- Tekin, Özlem (2023) ChatGPT im Unterricht Deutsch als Fremdsprache. In: *Alman Dili ve Kültürü Araştırmaları Dergisi* 5 (2), 135–163.
- Winter, Felix (2016) Formative Leistungsbeurteilung – Wie geht das? Mit den Lernenden über Qualität und Ergebnisse des Lernens sprechen. *Gemeinsam lernen* 2, 8–14.
- Ye, Jingheng; Wang, Shen; Zou, Deqing; Yan, Yibo; Wang, Kun; Zheng, Hai-Tao et al. (2025) *Position: LLMs Can be Good Tutors in Foreign Language Education.* <https://arxiv.org/pdf/2502.05467>.
- Zheldibayeva, Raigul (2025) GenAI as a Learning Buddy for Non-English Majors: Effects on Listening and Writing Performance. *Educational Process International Journal* 14, e2025051.

Biographical information

Dr Stefanie John is a German as a Foreign Language lecturer at the University of Oviedo in Spain. She completed her studies at Ruhr University Bochum in Germany, subsequently obtaining her PhD from the University of Oviedo. Her research focuses on second language acquisition and foreign language didactics, particularly the development of learner autonomy. At present, she is engaged in a project that explores the integration of artificial intelligence in GFL lessons as a component of an innovation

project. She also explores the potential of incorporating the United Nations' Sustainable Development Goals into language curricula to link foreign language instruction with broader issues of global citizenship and education for sustainability.

E-Mail: johnstefanie@uniovi.es

Schlüsselwörter

Künstliche Intelligenz; Sprachenlernen; Universität; Deutsch als Fremdsprache; Autonomie; Kritische KI-Kompetenz

Key words

Artificial Intelligence; Language Learning; University; German as a Foreign Language; Autonomy; Critical AI Literacy