

# KI-bezogene Kompetenzen von Französisch- und Spanischlehrkräften – eine empirische Untersuchung

Maximilian Irion<sup>1,\*</sup>, Sebastian Breitenbach<sup>1</sup>,  
Johanna Lea Korell<sup>1</sup> & Roland Ißler<sup>1</sup>

<sup>1</sup> *Goethe-Universität Frankfurt am Main*

\* *Kontakt: Goethe-Universität Frankfurt am Main,  
Institut für Romanische Sprachen und Literaturen,  
Norbert-Wollheim-Platz 1, 60629 Frankfurt am Main  
Mail: irion@em.uni-frankfurt.de*

**Zusammenfassung:** Die Integration Künstlicher Intelligenz (KI) in den Fremdsprachenunterricht eröffnet neue didaktische Möglichkeiten, stellt Lehrkräfte jedoch vor Herausforderungen in Bezug auf die kritische Reflexion und Anwendung entsprechender Tools. Die Studie untersucht KI-bezogene Kompetenzen von  $N = 279$  Französisch- und Spanischlehrkräften mittels Wissenstests und Selbsteinschätzungsbögen. Theoretisch basiert die Untersuchung auf zentralen KI-Kompetenzmodellen wie dem AI-PACK-Modell (Lorenz & Romeike, 2023) sowie der „AI Literacy“ (Ng et al., 2021) und „Prompt Literacy“ (Hwang et al., 2023). Methodisch folgt sie einem quantitativ-deskriptiven Ansatz im Sinne einer Bestandsaufnahme der KI-bezogenen Kompetenzen von Lehrkräften in Form ihres KI-bezogenen Wissens sowie ihrer KI-bezogenen Meinungen und selbstberichteten Verwendungsweisen. Die Befunde zeigen, dass Lehrkräfte ihre KI-Kompetenzen als eher niedrig einschätzen, was mit den Ergebnissen der Wissenstests übereinstimmt. Es besteht jedoch großes Interesse an KI, wobei sowohl Chancen als auch Herausforderungen erkannt werden. Konkrete Weiterbildungspotentiale sind vor allem in Hinblick auf die didaktische Integration von KI-Tools in den Fremdsprachenunterricht zu identifizieren, da deren Einsatz im Unterricht bislang nur vereinzelt erfolgt. Die Studie betont die Notwendigkeit gezielter Fortbildungsmaßnahmen, um eine kritische-reflexive KI-Nutzung durch Französisch- und Spanischlehrkräfte zu fördern.

**Schlagwörter:** Künstliche Intelligenz (KI); Lehrerforschung; Kompetenz; Französischunterricht; Spanischunterricht



## 1 Einführung: Ausgangslage und Forschungsbedarf

Durch die fortschreitende Entwicklung Künstlicher Intelligenz (KI), insbesondere im Bereich der Entwicklung der großen Sprachmodelle („large language models“, LLM), ergeben sich zunehmend neue Möglichkeiten für das Fremdsprachenlehren und -lernen, u.a. im Rahmen der Unterrichtsvorbereitung und -gestaltung, der Erstellung von Aufgaben sowie der Individualisierung von Unterricht (Dimitriadou & Lanitis, 2023). Der Einsatz solcher KI-Systeme verlangt jedoch von Lehrenden und Lernenden nicht nur technische Kompetenzen, sondern auch eine kritisch-reflexive Haltung gegenüber der Anwendung, der eigenen Arbeitsweise und der Bewertung der von der KI erzeugten Ergebnisse (Korell et al., 2025), wie sie bspw. im Konzept der „Critical Digital Literacy“ (Hinrichsen & Coombs, 2013) beschrieben wird.

Bisher konzentrieren sich empirische Studien zu schulbezogenen KI-Kompetenzen und zum Umgang mit KI-Anwendungen auf Lernende (u.a. Irion et al., 2025; Korell et al., 2025; Strasser, 2023) oder primär auf Lehrkräfte nicht-fremdsprachlicher Fächer (u.a. Antonenko & Abramowitz, 2023; Lindner & Berges, 2020). Für das Fach Englisch existieren vereinzelte Untersuchungen zu KI-bezogenen Umgangsweisen von Lehrkräften (u.a. Al-khresheh, 2024; Bajaña Ordoñez & Espinal Bailón, 2024; Meyasa et al., 2024). Studien, die explizit Französisch- und/oder Spanischlehrkräfte und ihre KI-bezogenen Verwendungsweisen im Unterricht adressieren, konnten nicht identifiziert werden.

Dieser Forschungslücke nimmt sich der vorliegende Beitrag<sup>1</sup> an und präsentiert die Auswertung schriftlicher Wissenstests und Selbsteinschätzungsbögen von  $N = 279$  Französisch- und Spanischlehrkräften, die im Rahmen des ViFoNet-Projekts<sup>2</sup> eingesetzt wurden. Der Beitrag liefert somit erste Erkenntnisse zu KI-bezogenen Kompetenzen von Lehrkräften romanischer Schulfremdsprachen, konkret zu deklarativem und prozeduralem KI-bezogenem Wissen wie auch zu KI-bezogenen Sicht- und Verwendungsweisen.

## 2 Theoretischer und konzeptueller Rahmen

### 2.1 KI-bezogene Kompetenzen von Fremdsprachenlehrkräften

In den vergangenen Jahren sind neue Modelle digitalisierungsbezogener Kompetenzen für das Lehren und Lernen entwickelt worden (vgl. Korell et al., 2025), darunter z.B. das DigCompEdu- (Redecker & Punie, 2017) sowie das DPaCK-Modell (Digitality-related Pedagogical Content Knowledge, Huwer et al., 2019). Diese Modelle umfassen lehrkraftseitige Fähigkeiten und Voraussetzungen, die zur effektiven Nutzung digitaler Medien im Unterricht notwendig sind, sowie motivationale Bereitschaften, diese Fähigkeiten im Kontext des eigenen Unterrichtens gezielt einzusetzen (Redecker & Punie, 2017; Weinert, 2014). Die dargestellten Fähigkeiten beziehen sich einerseits auf kognitive Aspekte, z.B. auf digitalisierungsspezifisches pädagogisches Wissen (Huwer et al.,

---

<sup>1</sup> Maximilian Irion, Sebastian Breitenbach und Johanna Lea Korell tragen eine Ko-Autorschaft. Sie verantworten gemeinsam die Datenauswertung und -interpretation und die Anfertigung und Überarbeitung des Manuskripts. Maximilian Irion ist zudem korrespondierender Autor. Neben der Verantwortung für die Einreichung und die Koordination der Zusammenarbeit der Autor\*innen konzentriert sich sein Beitrag auf die methodische Anlage und auf die Unterstützung bei der theoretischen Rahmung der Untersuchung. Sebastian Breitenbach ist ebenfalls für die methodische Anlage und darüber hinaus für die Datenaufbereitung zuständig. Johanna Lea Korell ist überdies für die theoretische Anlage, die Teilnehmendengenerierung und die Datenerhebung verantwortlich. Der Beitrag von Roland Ißler konzentriert sich auf das Schlusslektorat und die Akquise der Finanzierung des zugrundeliegenden Forschungsprojekts.

<sup>2</sup> ViFoNet ist ein durch die Europäische Union (NextGenerationEU) finanziertes und durch das Bundesministerium für Bildung und Forschung (BMBF) gefördertes Verbundprojekt von sechs Hochschulen in Deutschland. Die geäußerten Ansichten und Meinungen sind ausschließlich die der Autor\*innen und spiegeln nicht unbedingt die Ansichten der Europäischen Union, Europäischen Kommission oder des BMBF wider. Die genannten Institutionen können für sie nicht verantwortlich gemacht werden.

2019), andererseits stehen einstellungs- und verhaltensbezogene Teilkompetenzen im Mittelpunkt. Hierzu zählen etwa die kritische Reflexion des Einsatzes digitaler Medien, aber auch die Stärkung digitaler Kompetenzen von Lernenden, z.B. der Informations- und Medienkompetenz sowie des verantwortungsvollen Umgangs mit digitalen Medien (Redecker & Punie, 2017).

Während sich die o.g. Modelle vorrangig auf die Rezeption und Produktion herkömmlicher digitaler Inhalte beschränken, zielt der vorliegende Beitrag auf KI-Kompetenzen ab, die explizit das Wechselspiel zwischen Mensch und KI-Anwendung fokussieren und Kompetenzen definieren, die Individuen zu einer souveränen und reflektierten Anwendung von KI-Technologien befähigen (Korell et al., 2025). Einige dieser Modelle werden nachfolgend erläutert.

Ein grundlegender Versuch der Einbindung KI-spezifischer Aspekte in digitalisierungsbezogene Kompetenzen von Lehrkräften stellt das sogenannte AI-PACK-Modell dar (Lorenz & Romeike, 2023). Dieses umfasst in Anlehnung an das DPaCK-Modell

- eine allgemeine KI-Kompetenz, die das Erkennen, Verstehen, Reflektieren und Umsetzen von KI-Phänomenen aus einer technologischen, gesellschaftlich-kulturellen und anwendungsorientierten Perspektive umfasst,
- eine KI-bezogene pädagogische Kompetenz (AI-PK), die sich auf das Erkennen und Reflektieren von Potentialen und Grenzen von KI-Systemen hinsichtlich allgemeiner Lehr-Lernprozesse bezieht,
- eine KI-bezogene (fach-)inhaltliche Kompetenz (AI-CK), die sich auf das Erkennen und Reflektieren von Auswirkungen von KI-Systemen auf den eigenen Fachbereich, die eigene Fachwissenschaft bzw. das eigene Unterrichtsfach bezieht, sowie
- eine KI-bezogene pädagogisch-inhaltliche Kompetenz (AI-PCK), die das Erkennen und Reflektieren des Einflusses von KI auf fachspezifische Lehr-Lernprozesse sowie das Gestalten KI-basierter fachspezifischer Lehr-Lernsettings einschließt (vgl. Abb. 1).

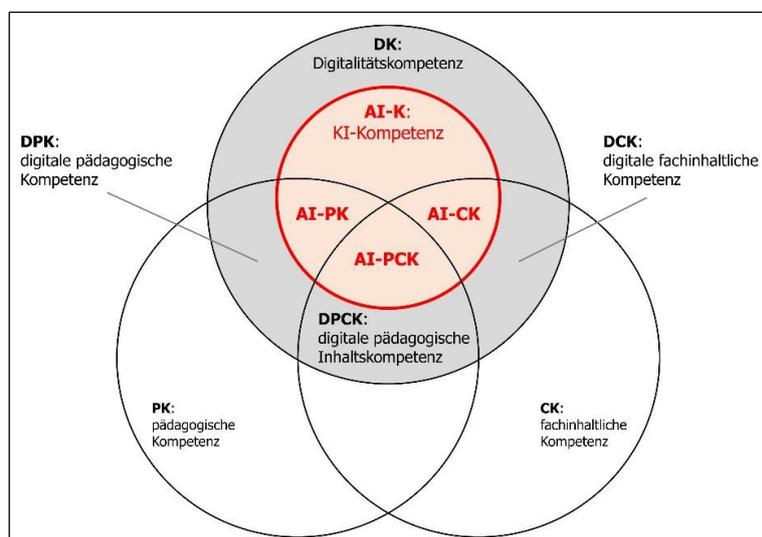


Abbildung 1: AI-PACK-Modell (eigene Darstellung in Anlehnung an Lorenz und Romeike, 2023)

Als Basis für die Entwicklung weiterer KI-spezifischer Kompetenzmodelle kann das Konzept der „Critical Digital Literacy“ (Hinrichsen & Coombs, 2013) betrachtet werden. Dieses basiert auf fünf Komponenten, die für die effektive Verwendung digitaler Medien unabdingbar sind:

- das Dekodieren von Inhalten („decoding“),
- das Untersuchen von Inhalten nach deren Sinngehalt („meaning making“),
- die technische Nutzung („using“),
- das Analysieren von Daten („analysing“) sowie
- die Reflexion über die eigene Rolle in einem medial geprägten Umfeld („persona“).

Dezidiert KI-bezogene Aspekte enthält insbesondere das „AI Literacy Framework“ von Ng et al. (2021). Im Fokus stehen vielfältige Kompetenzbereiche, welche sich u.a. auf das Verstehen technischer Konzepte von KI, das Verstehen und Aneignen von Strategien und Techniken bei der Nutzung von KI und auf die Entwicklung einer kritischen Haltung bei der Nutzung von KI beziehen.

Besonders hervorzuheben ist darüber hinaus das Mehrebenen-Modell der „AI Literacy“ nach Wienrich et al. (2022), das eine Mikro- (Individuum), Meso- (Organisation) und Makroebene (Politik und Gesellschaft) umfasst. Dieses Modell ist in seiner Anlage mit dem DOIT-Modell von Horz und Schulze-Vorberg (2017) vergleichbar, welches generell digitale Werkzeuge fokussiert, weist im Gegensatz dazu allerdings einen expliziten KI-Bezug auf. Die Mikroebene bezieht sich auf die individuelle „AI Literacy“ und damit verbundene Faktoren wie Einstellungen, kognitive, motivationale und emotionale Dispositionen. Auch das sogenannte KI-Verständnis, also z.B. das Wissen über die für die KI-Nutzung erforderlichen Fähigkeiten, ist auf der Mikroebene verortet. Während sich die Mesoebene auf organisationale Rahmenbedingungen bezieht, umfasst die Makroebene vorrangig politische und gesellschaftliche Prinzipien (z.B. soziale Normen). Wechselwirkungen zwischen Mikro- und Meso- bzw. Makroebene manifestieren sich in der KI-Reflexion, die dazu dient, die Grenzen von KI zu erkennen und deren Auswirkungen sowie die Verantwortung des Menschen im Umgang mit KI zu hinterfragen. Das KI-Verhalten stellt schließlich die Veräußerung der „AI Literacy“ dar und zeigt sich bspw. in der individuellen Anwendung von KI-Tools in spezifischen Situationen.

Neben den von Wienrich et al. (2022) genannten individuellen Dispositionen erfordert das KI-Verhalten jedoch auch eine gezielte menschliche Steuerung. In diesem Kontext gewinnt das Konzept der „Prompt Literacy“ (Hwang et al., 2023) zunehmend an Bedeutung. Es beschreibt die Fähigkeit, zielgerichtet mit KI-Systemen zu interagieren, indem Nutzende Eingaben, sog. Prompts, formulieren, die möglichst stichhaltige Antworten hervorbringen. Zusätzlich umfasst sie die Fähigkeit, bereits getätigte Eingaben zu optimieren, um präzisere Ergebnisse zu erzielen.

Die vorgestellten Modelle und Konzepte bieten komplementäre Perspektiven auf KI-bezogene Kompetenzen, die in Abbildung 2 schematisch veranschaulicht werden. Auf Grundlage dieser theoretischen Konzeptualisierung wird nachfolgend das Konstrukt der KI-bezogenen Kompetenzen für den vorliegenden Beitrag erläutert und auf den fremdsprachlichen Kontext übertragen.

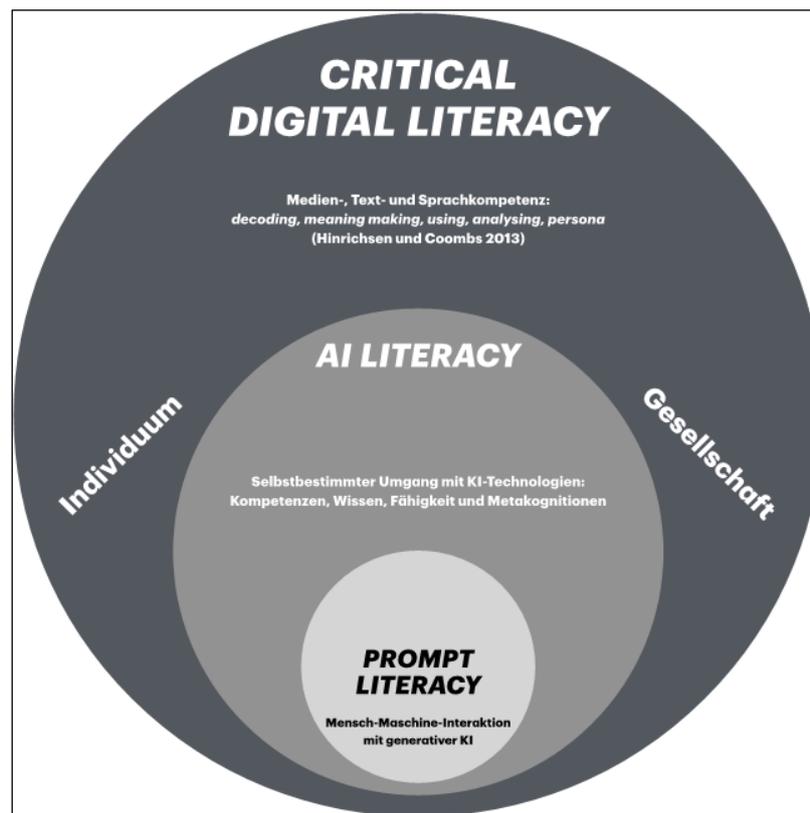


Abbildung 2: Zusammenhänge zwischen „Critical Digital Literacy“, „AI Literacy“ und „Prompt Literacy“ nach Korell et al. (2025, S. 141)

## 2.2 Konstruktdefinition

Unter Rückgriff auf die in Kapitel 2.1 skizzierten Modelle ergibt sich für die Konstruktdefinition eine dreidimensionale Struktur, die eine kognitive, meinungs- und handlungsbezogene Dimension umfasst (vgl. Korell et al., 2025). Eine solche Struktur steht im Einklang mit zentralen Annahmen der Lehrkräfteforschung, die ein multidimensionales und über Wissen hinausgehendes Verständnis von professioneller Kompetenz zugrunde legen (Baumert & Kunter, 2011) und ein Kontinuum zwischen Kompetenz und Performanz im Sinne des konkreten Unterrichtshandelns postulieren (Blömeke et al., 2015). Die Fokussierung auf KI-relevante Aspekte durch Konzepte wie „AI Literacy“ und „Prompt Literacy“ ermöglicht ebenso eine umfassende und ganzheitliche Betrachtung des Konstrukts der KI-bezogenen Kompetenzen von Lehrkräften. Tabelle 1 ist die Konzeptualisierung der einzelnen Dimensionen des Untersuchungsgegenstands unter Rückgriff auf die gelisteten (Teil-)Aspekte der verschiedenen Modelle zu entnehmen.

Tabelle 1: Dimensionen und Facetten von KI-bezogenen Kompetenzen von Fremdsprachenlehrkräften

Dimension des Untersuchungsgegenstands	Modelle (alphabetisch nach Autor*innen)
<p><b>KI-bezogenes deklaratives und prozedurales Wissen</b></p> <ul style="list-style-type: none"> <li>• Kenntnis zentraler KI-bezogener Fachbegriffe und Funktionsweisen (deklarativ)</li> <li>• Kenntnis verschiedener KI-Tools (deklarativ)</li> <li>• Kenntnis über sachgemäße Verwendung von KI-Tools, auch im Fremdsprachenunterricht (prozedural)</li> <li>• Reflexion über Auswirkungen des Einsatzes von KI für allgemeine und fremdsprachenspezifische Lehr-Lernprozesse (prozedural)</li> </ul>	<ul style="list-style-type: none"> <li>• Baumert &amp; Kunter (2011): Professionswissen</li> <li>• Blömeke et al. (2015): „cognition“ („disposition“)</li> <li>• Hinrichsen &amp; Coombs (2013): „analysing“</li> <li>• Hwang et al. (2023): „Prompt Literacy“</li> <li>• Lorenz &amp; Romeike (2023): AI-PACK</li> <li>• Ng et al. (2021): Wissen und Verstehen von KI, Entwicklung einer kritischen Haltung bei der Nutzung von KI</li> <li>• Wienrich et al. (2022): KI-Verständnis (Wissen, welche Fähigkeiten die Nutzung von KI benötigt)</li> </ul>
<p><b>KI-bezogene Meinungen</b></p> <ul style="list-style-type: none"> <li>• Interesse an KI</li> <li>• Einschätzung zur Sinnhaftigkeit des Einsatzes von KI-Tools im Fremdsprachenunterricht</li> <li>• Meinung zur Ersetzung des Fremdsprachenunterrichts durch KI-Tools</li> <li>• Meinung zu Gefahren und Problemen bei der Nutzung von KI-Tools zum Fremdsprachenlernen</li> <li>• Selbstwirksamkeitserwartung bezüglich des Einsatzes von KI-Tools im Fremdsprachenunterricht</li> </ul>	<ul style="list-style-type: none"> <li>• Baumert &amp; Kunter (2011): Überzeugungen, Werthaltungen, Ziele, motivationale Orientierungen</li> <li>• Blömeke et al. (2015): „affect-motivation“ („disposition“)</li> <li>• Hinrichsen &amp; Coombs (2013): „persona“</li> <li>• Wienrich et al. (2022): „AI Mental State“ und KI-Reflexion (Grenzen der KI, Verantwortung des Menschen im Umgang mit KI)</li> </ul>
<p><b>KI-bezogene Verwendungsweisen</b></p> <ul style="list-style-type: none"> <li>• Nutzungshäufigkeit von KI-Tools im Fremdsprachenunterricht</li> <li>• Nutzungszwecke von KI-Tools für den Fremdsprachenunterricht</li> </ul>	<ul style="list-style-type: none"> <li>• Blömeke et al. (2015): „observable behavior“ („performance“)</li> <li>• Hinrichsen &amp; Coombs (2013): „using“ und „analysing“</li> <li>• Ng et al. (2021): Verstehen und Aneignen von Strategien und Techniken bei der Nutzung von KI</li> <li>• Wienrich et al. (2022): KI-Verhalten (Sichtbarkeit der „AI Literacy“ durch die Mensch-Maschine-Interaktion)</li> </ul>

Tabelle 1 verdeutlicht, dass KI-bezogene Kompetenzen von Fremdsprachenlehrkräften eine Vielzahl an Dimensionen und Facetten umfassen. Vor diesem Hintergrund ist besonders kritisch zu sehen, dass bislang kaum Erhebungsinstrumente zur validen Erfassung dieser Dimensionen vorliegen. Eine Ausnahme stellt der AI-Literacy-Test von

Hornberger et al. (2023) dar, der sich allerdings auf deklaratives und prozedurales KI-bezogenes Wissen und damit auf die kognitive Dimension von KI-bezogenen Kompetenzen beschränkt, diese jedoch nicht in ihrer Ganzheitlichkeit adressiert. Instrumente zur gezielten Erhebung von pädagogisch relevanten kognitiven Facetten und zur hier aufgezeigten holistischen Betrachtung des Untersuchungsgegenstandes fehlen ebenfalls noch (vgl. Kap. 3.3).

### 2.3 Der Einsatz von KI im Fremdsprachenunterricht: Empirische Befunde

Die systematische Literaturrecherche von Korell und Albrecht (2025) zeigt, dass das KI-gestützte Französisch- und Spanischlehren und -lernen einen noch jungen Forschungsbereich darstellt und diesbezügliche empirische Untersuchungen erst seit 2023 ansteigen. Unter den identifizierten Studien finden sich allerdings keine, die sich dezidiert mit KI-bezogenen Wissensbeständen sowie Sicht- und Verwendungsweisen von Französisch- und Spanischlehrkräften beschäftigen.

Mit der im Rahmen des vorliegenden Beitrags nach der PRISMA-Methode durchgeführten Literaturrecherche können die vorausgehenden Ergebnisse bestätigt werden: Bisher konnten ausschließlich empirische Studien identifiziert werden, die sich mit Perspektiven von Englischlehrkräften auf die Integration von KI-Anwendungen in den Unterricht befassen. In den meisten Studien werden qualitative Erhebungsmethoden (v.a. Interviews) verwendet, vereinzelt finden auch quantitative Methoden (v.a. Fragebögen, statistische Modelle) Anwendung. Im Folgenden werden die wesentlichen Aspekte dieser Literaturrecherche exemplarisch zusammengefasst, denn auch wenn sie nicht dezidiert auf Französisch- und Spanischlehrkräfte abzielen, können sie Hinweise für die vorliegende Studie bieten.

Die Studien untersuchen die Einbindung von KI, insbesondere ChatGPT, in den Englischunterricht und analysieren, wie Lehrkräfte entsprechende Technologien wahrnehmen und anwenden. Im Großteil der gesichteten Studien werden ethische Bedenken und Skepsis hinsichtlich der Abhängigkeit von KI-Technologien vonseiten der Lehrkräfte geschildert (Al-khresheh, 2024; Bajaan Ordoñez & Espinal Bailón, 2024). Dazu gehören (a) Datenschutzprobleme und fehlende technisch-strukturelle Voraussetzungen (Meyasa et al., 2024), (b) mangelnde Zuverlässigkeit (Mohammad Ali, 2023; Nazaretsky et al., 2022) und (c) Vorurteile (Meyasa et al., 2024). Auch Bedenken bezüglich der Kreativität von Lernenden und des Übergebrauchs von Technologie werden häufig genannt (Al-khresheh, 2024). Insgesamt zeichnet sich ab, dass Lehrkräfte über ein eher rudimentär ausgeprägtes KI-bezogenes deklaratives Wissen verfügen und vor allem ein Wunsch an prozeduralem Wissen besteht (Nazaretsky et al., 2022).

Demgegenüber werden KI-Anwendungen mit Blick auf eine Verbesserung des personalisierten Lernens und individuellen Feedbacks positiv betrachtet. Eine Studie untersucht bspw. die Verbesserung der Sprechfähigkeiten (v.a. Aussprache und Flüssigkeit) unter Einsatz von KI-Anwendungen (Bajaan Ordoñez & Espinal Bailón, 2024), andere Untersuchungen fokussieren den Einsatz von KI für personalisiertes Lernen und die Bewertung von Schreibprodukten (Meyasa et al., 2024; Mohammad Ali, 2023). Eine weitere Untersuchung zeigt hingegen die Einschränkungen auf, die mit dem Einsatz von KI im Bereich des fremdsprachlichen Hörens und Sprechens einhergehen können (Al-khresheh, 2024). Auch wenn diese Untersuchungen insgesamt ein recht homogenes Bild der Sichtweisen von Englischlehrkräften zeichnen, erweisen sich für die vorliegende Studie v.a. quantitative Ansätze als gewinnbringend, um Erkenntnisse über die KI-Kompetenzen einer möglichst großen Bandbreite von Französisch- und Spanischlehrkräften zu gewinnen und überindividuelle Tendenzen zu identifizieren (Al-khresheh, 2024).

## 2.4 Die vorliegende Studie

KI-bezogene Kompetenzen von Fremdsprachenlehrkräften umfassen eine Vielzahl an Facetten, die sich auf kognitive, meinungs- und handlungsbezogene Dimensionen beziehen. Inwieweit Französisch- und Spanischlehrkräfte über eine solche komplexe Kompetenzstruktur verfügen, ist bislang unerforscht, was sich u.a. durch den Mangel an verfügbaren Erhebungsinstrumenten erklären lassen könnte. Untersuchungen zu KI-bezogenen Kompetenzen von Englischlehrkräften deuten allerdings auf Lücken im KI-bezogenen deklarativen Wissen sowie auf Bedenken gegenüber KI-Technologien hin.

Um zu überprüfen, ob diese Beobachtungen auf Lehrkräfte romanischer Schulfremdsprachen übertragbar sind, zielt die vorliegende Studie darauf ab, anhand bewährter und neu entwickelter Messinstrumente die KI-bezogenen Kompetenzen von Französisch- und Spanischlehrkräften explorativ-quantitativ zu untersuchen, um eine Ausgangsbasis für spezifischere Folgestudien zu legen. Konkret stehen dabei die folgenden Forschungsfragen im Fokus:

- 1) Über welches KI-bezogene deklarative und prozedurale Wissen verfügen Französisch- und Spanischlehrkräfte und wie schätzen sie ihre Kenntnisse ein? (kognitive Dimension)
- 2) Welche KI-bezogenen Meinungen liegen bei Französisch- und Spanischlehrkräften vor? (meinungsbezogene Dimension)
- 3) Von welchen KI-bezogenen Verwendungsweisen inner- und außerhalb des Fremdsprachenunterrichts berichten Französisch- und Spanischlehrkräfte? (handlungsbezogene Dimension)

## 3 Methodologie und empirisches Design

### 3.1 Stichprobe

Die Stichprobe der vorliegenden Studie besteht aus  $N = 279$  Französisch- ( $n = 146$ ; 1. Kohorte  $n = 20$ , 2. Kohorte  $n = 126$ ) und Spanischlehrkräften ( $n = 133$ ), die an Online-Fortbildungen mit dem Titel „Der Einsatz Künstlicher Intelligenz beim Lehren und Lernen der französischen/spanischen Sprache“ teilnahmen.<sup>3</sup> Diese Fortbildungen wurden im Rahmen des BMBF-geförderten Verbundprojekts ViFoNet von August 2024 bis Februar 2025 an der Goethe-Universität Frankfurt durchgeführt. Zusätzlich bewarb die Goethe-Lehrkräfteakademie (GLA) die Fortbildungen. Da sich interessierte Lehrkräfte auf freiwilliger Basis für die entsprechende Fortbildung ihres Fachs selbst registrieren konnten, liegt eine Ad-Hoc-Stichprobe (Grum & Legutke, 2022) vor, die auf dem Prinzip der Selbstaktivierung beruht (Reinders, 2016).

Insgesamt 13.6 Prozent der teilnehmenden Lehrkräfte ordneten sich dem männlichen und 85.3 Prozent dem weiblichen Geschlecht zu. Drei Personen machten keine Angabe. Die Möglichkeit der diversen Zuordnung wurde nicht genutzt. Die Lehrkräfte waren zum Zeitpunkt der Erhebung im Durchschnitt 42.6 Jahre alt ( $SD = 8.96$ ,  $Min = 25$ ,  $Max = 67$ ) und unterrichteten zum größten Teil an Gymnasien (92.8 %), danach folgten Gesamtschulen (5.0 %) und weitere Schulen (1.7 %). Im Durchschnitt verfügten die Lehrkräfte über eine Berufserfahrung von 14.2 Jahren ( $SD = 8.04$ ,  $Min = 1$ ,  $Max = 44$ ). Neben Französisch bzw. Spanisch unterrichtete die Mehrheit der Lehrkräfte Englisch ( $n = 59$ ) oder Deutsch ( $n = 57$ ) als weiteres Fach. MINT-Zweifächer wie Informatik ( $n = 7$ ) und Mathematik ( $n = 11$ ) sind nur geringfügig in der Stichprobe vertreten.

---

<sup>3</sup> Insgesamt waren in diesen Fortbildungen bundesweit 457 Lehrkräfte eingeschrieben. Die Datengrundlage des vorliegenden Beitrags bezieht sich dabei nur auf diejenigen Lehrkräfte, die die Fragebögen der zugrundeliegenden Studie bearbeitet haben ( $N = 279$ ).

### 3.2 Design

Um die KI-bezogenen Kompetenzen der Französisch- und Spanischlehrkräfte der vorliegenden Stichprobe zu beschreiben, wurde ein explorativ-quantitativer, zunächst rein deskriptiver Ansatz gewählt (Schramm, 2022). Die analysierten Daten wurden zu Beginn der Fortbildungen erhoben und bieten einen Einblick in KI-bezogene Kompetenzen von Lehrkräften romanischer Schulfremdsprachen. Die Wirksamkeitsuntersuchung der Fortbildungen anhand inferenzstatistischer Analysen sowie die statistische Evaluation der eingesetzten Messinstrumente (s. Kap. 3.3) erfolgen in gesonderten Studien.

### 3.3 Messinstrumente

Die in Kapitel 2.2 identifizierte kognitive Dimension der KI-bezogenen Kompetenzen wird über drei verschiedene Skalen operationalisiert: (a) Selbsteingeschätzte KI-Kompetenz (Eigenentwicklung), (b) AI-Literacy-Test nach Hornberger et al. (2023) und (c) KI-Wissenstest für Lehrkräfte (Eigenentwicklung). Die Inklusion der beiden Tests (b) und (c) basierte auf der Überlegung, die Selbsteinschätzung des KI-bezogenen Wissens abzusichern. Die Eigenentwicklung des zweiten Tests (c) verfolgte zudem das Ziel, konkret KI-bezogenes Wissen zu erfassen, das für Fremdsprachenlehrkräfte und den -unterricht von Relevanz ist, da diese Wissensbereiche nach Ansicht der Autor\*innen im AI-Literacy-Test von Hornberger et al. (2023) nicht hinreichend berücksichtigt werden. Als Grundlage für die Entwicklung dienten die in Abschnitt 2.1 dargestellten KI-Kompetenzmodelle. Darauf aufbauend wurden ebenso Aspekte berücksichtigt, die sich spezifisch auf deutsche Rahmenbedingungen des Einsatzes von KI im Unterricht beziehen. Die handlungs- und meinungsbezogenen Dimensionen wurden über Einzelitems operationalisiert. Alle Messinstrumente sind vollständig im Online-Supplement abgebildet.

### 3.4 Datenerhebung, -aufbereitung und -auswertung

Sämtliche Daten wurden mittels Online-Fragebögen erhoben. Schwierigkeiten bei der Datenaggregation in der ersten Kohorte der Französischfortbildung konnten in den nachfolgenden Erhebungen behoben werden. Da viele Teilnehmende der Fortbildung aus Zeit- und Belastungsgründen nur einen Teil der Fragebögen ausfüllten, liegen für diese Personen keine vollständigen Datensätze vor. Der Rücklauf vollständiger Fälle mit 58.2 Prozent in der Französisch- und 64.6 Prozent in der Spanischfortbildung bewegt sich im für Lehrkräftefortbildungen erwartbaren Rahmen.

Die Datensätze wurden in R eingelesen und via Personencode aggregiert. Alle Items der selbsteingeschätzten KI-Kompetenz sind unidirektional formuliert und basieren auf einer 4-stufigen Likert-Skala von 0 = *gering ausgeprägte Kompetenz* bis 3 = *stark ausgeprägte Kompetenz*. Im Single-Choice-basierten AI-Literacy-Test nach Hornberger et al. (2023) wurden richtige Antworten mit 1 und falsche Antworten mit 0 kodiert; anschließend erfolgte die Bildung eines Summenwerts. Im Multiple-Choice-KI-Wissenstest wurde hingegen für jede korrekte Antwort innerhalb eines Items ein Punkt vergeben, während falsche Antworten mit einem Minuspunkt bewertet wurden.

Die Vervollständigung vereinzelt fehlender Werte zur selbsteingeschätzten KI-Kompetenz erfolgte über eine Modus-Imputation. Bei mit „weiß nicht“ beantworteten Items wurde ein Wert von 0 = *keine Kompetenz* imputiert. Diese Entscheidung lässt sich damit begründen, dass man von keiner gesicherten Kompetenz ausgehen kann, wenn ein entsprechendes Item mit „weiß nicht“ beantwortet wird. Die Imputationen betrafen nur die Daten der  $n = 20$  Französischlehrkräfte der 1. Kohorte. Ab zehn fehlenden Werten fand keine Berechnung der selbsteingeschätzten KI-Kompetenz statt und der Fall wurde als nicht vollständig behandelt. Für losgelöste Einzelitems zu Einstellungen erfolgte keine Imputation.

## 4 Ergebnisse der empirischen Untersuchung und Diskussion

### 4.1 KI-bezogenes deklaratives und prozedurales Wissen

Die selbsteingeschätzte KI-Kompetenz lag in der vorliegenden Stichprobe im Durchschnitt bei  $M = 1.1$  ( $SD = 0.5$ ; vgl. Abb. 3). Die Lehrkräfte schätzten ihre Fähigkeit im Umgang mit KI somit insgesamt eher niedrig ein. Dies bestätigte sich auch in der deskriptiven Analyse der Summenwerte des AI-Literacy-Tests, die mit  $M = 14.6$  ( $SD = 4.3$ , potentieller Wertebereich: Min = 0, Max = 31) bis auf einige Ausreißer (vgl. Abb. 3) im Allgemeinen niedrig ausfielen. Der KI-Wissenstest für Lehrkräfte erbrachte hingegen etwas höhere Werte und streute, gemessen am potentiellen Wertebereich, weniger ( $M = 3.7$ ,  $SD = 5.8$ , potentieller Wertebereich: Min = -37, Max = 32). Die Ergebnisse stehen in Einklang mit anderen Studien, die ebenfalls auf eine relativ gering ausgeprägte KI-Kompetenz von Lehrkräften nicht-fremdsprachlicher Fächer (Lindner & Berges, 2020) hindeuten.

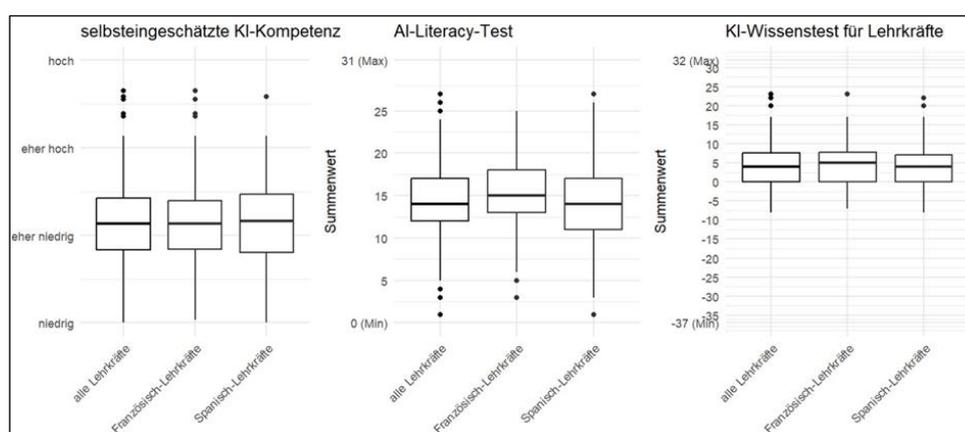


Abbildung 3: Deskriptive Statistiken der KI-bezogenen Kompetenzen von Französisch- und Spanischlehrkräften

Die an Gymnasien praktizierenden Lehrkräfte unterschieden sich hinsichtlich der selbsteingeschätzten KI-Kompetenz ( $M = 1.1$ ,  $SD = 0.5$ ) und der Summenwerte im KI-Wissenstest für Lehrkräfte ( $M = 3.6$ ,  $SD = 5.8$ ) nicht von den Lehrkräften an Gesamtschulen (selbsteingeschätzte KI-Kompetenz:  $M = 1.1$ ,  $SD = 0.5$ ;  $z = -0.63$ ,  $p = .526$ ; KI-Wissenstest für Lehrkräfte:  $M = 4.0$ ,  $SD = 5.1$ ;  $z = -0.22$ ,  $p = .826$ ). Im AI-Literacy-Test erzielten die Gymnasiallehrkräfte ( $M = 14.7$ ,  $SD = 4.3$ ) deskriptiv marginal höhere Werte als die Gesamtschullehrkräfte ( $M = 12.5$ ,  $SD = 4.5$ ;  $z = -1.34$ ,  $p = .180$ ). Lehrkräfte anderer Schulformen waren in der Stichprobe kaum vorhanden, weshalb eine genauere Differenzierung an dieser Stelle für nicht sinnvoll erachtet wird. Mittels Kendall-Rangkorrelationsstest kann jedoch gezeigt werden, dass das Alter signifikant negativ mit der selbsteingeschätzten KI-Kompetenz zusammenhängt ( $z = -2.08$ ,  $p = .037$ ). Ältere Lehrkräfte schätzen sich demnach als weniger KI-kompetent im Vergleich zu jüngeren Lehrkräften ein. Die Zusammenhänge zwischen Alter und KI-Wissenstest ( $z = -1.58$ ,  $p = .114$ ) bzw. AI-Literacy-Test ( $z = 1.79$ ,  $p = .074$ ) fallen hingegen nicht signifikant aus. Die Befunde legen somit nahe, dass sich jüngere Lehrkräfte offensichtlich stärker als ältere Lehrkräfte mit KI-Anwendungen vertraut fühlen, was z.B. auf ein umfassenderes privates Nutzungsverhalten oder eine stärkere Beschäftigung mit KI-spezifischen Themen im Freizeitbereich zurückzuführen sein könnte. Gleichzeitig manifestiert sich die als höher eingeschätzte KI-Kompetenz jedoch nicht in tatsächlich umfangreichem KI-bezogenem Wissen.

Nachfolgend werden die deskriptiven Statistiken in Bezug auf Einzelitems präsentiert und diskutiert. Im Hinblick auf das deklarative Wissen stimmte die Mehrzahl der Teilnehmenden der Aussage „zu wissen, was ein large language model (LLM) ist“, *gar nicht* zu (64.0 %). Eine Analyse der Einzelitems des Tests ermöglicht zusätzliche Erkenntnisse darüber, in welchen Bereichen KI-bezogenes Wissen vorhanden war und in welchen Wissenslücken bestanden. Insbesondere können so auch übliche Fehlkonzepte erkannt werden (Antonenko & Abramowitz, 2023). Nur 11.5 Prozent der Teilnehmenden gaben die korrekte Antwort, dass ein lokales LLM in einigen Fällen nicht in der Lage ist, die von den Nutzer\*innen gewählte Sprache zu verstehen. Es lag offenbar das Fehlkonzept vor, dass alle LLMs jede Sprache verstehen. Einige LLMs sind für spezifische Sprachen entwickelt; andere sind zwar universell, aber der Datensatz, mit dem diese Modelle trainiert wurden, enthält eventuell zu wenig Textkorpus der jeweiligen Sprache. In Bezug auf die in Abschnitt 2.1 thematisierte „Prompt Literacy“ erkannten nur 13.3 Prozent der Teilnehmenden, dass die Strategie, dem LLM dreimal dieselbe Frage zu stellen und die KI anschließend zu bitten, diese Antworten als ein Experte für das Themenfeld zu vergleichen, sinnvoll ist und zu besseren Ergebnissen führt. Zudem kannten die meisten Lehrkräfte keine (31.8 %) oder nur wenige (44.4 %) Strategien oder Techniken zur Nutzung von LLMs.

Bei den am häufigsten richtig beantworteten Items wurde deutlich, dass die Lehrkräfte gut über Datenschutz informiert waren. Die Mehrzahl der Teilnehmenden gab korrekt an, dass Lernende nicht dazu aufgefordert werden dürfen, sich mit privaten (95.3 %) oder schulischen (75.3 %) E-Mailadressen auf ChatGPT anzumelden, um die KI anschließend im Unterricht zu nutzen. Erstaunlicherweise wurde die Rechtskonformität von Offline-Lösungen via lokalen LLMs auf Schullaptops (34.2 % richtige Antworten) oder Schulservern (29.7 % richtige Antworten) oft nicht erkannt. Dies deutet auf eine Tendenz zur Gleichsetzung von KI und ChatGPT bzw. auf eine Identifikation von LLMs als reine (fremdgehostete) Browserapplikationen hin. Evidenz für diese Vermutung liefert auch die Tatsache, dass einer der bedeutendsten Konkurrenten von ChatGPT bzw. Open AI, Anthropic Claude Sonnet, nur von 20.0 Prozent der Lehrkräfte als LLM erkannt wurde. Zusätzlich wussten die Lehrkräfte nur selten, dass LLMs intern über Matrixoperationen gesteuert werden (29.3 %), und waren sich ihrer algorithmischen Funktionsweisen *eher nicht* bewusst. Es sollte nicht das Ziel sein, Fremdsprachenlehrkräfte mit KI-bezogenem Detailwissen zu überfordern, aber es ist wichtig, sie für die Tatsache zu sensibilisieren, dass LLMs nur aufgrund stochastischer Methoden funktionieren. Zumindest ein grundlegendes Verständnis des Aufbaus sollte vorhanden sein, um bspw. die Outputs entsprechend kritisch hinterfragen zu können (Nazaretsky et al., 2022).

Im Bereich des prozeduralen Wissens gaben 75.3 Prozent der Lehrkräfte an, *gar nicht* oder *eher nicht* zu wissen, wie sie Lernende zu einem kritischen Umgang mit LLMs anregen können. Bei maschinellen Übersetzungstools waren es hingegen nur 56.4 Prozent, die von sich behaupteten, dies *gar nicht* oder *eher nicht* zu wissen. Dies deutet darauf hin, dass der Umgang mit maschinellen Übersetzungstools bereits stärker in der Schulpraxis verankert scheint, auch wenn noch große Potentiale brachliegen. Knapp 70 Prozent der Befragten führten aus, *gar nicht* oder *eher nicht* zu wissen, wie ihre Lernenden LLMs für fremdsprachliche Zwecke nutzen, während es bei maschinellen Übersetzungstools nur etwa 50 Prozent waren. Auffällig ist, dass die meisten Lehrkräfte angaben, *gar nicht* (33.0 %) oder *eher nicht* (55.2 %) über konkrete Unterrichtsaktivitäten informiert zu sein, durch welche Lernende den verantwortungsbewussten Umgang mit LLMs und maschinellen Übersetzungstools zur Kommunikation und Zusammenarbeit erlernen können.

Es wird häufig darauf verwiesen, dass KI-Anwendungen bei der Adaptivität, Binnendifferenzierung und konstruktiven Beratung als zentrale Merkmale guten Unterrichts helfen können (Dimitriadou & Lanitis, 2023). Allerdings gab eine verschwindend geringe Zahl von Lehrkräften an, LLMs *eher* (8.2 %) oder *voll und ganz* (1.4 %) zu nutzen,

um gezielt Beratung oder Unterstützung anzubieten. Die meisten Lehrkräfte berichteten, dies *gar nicht* (40.1 %) oder *eher nicht* (50.2 %) zu können. Besorgniserregend ist, dass sich die meisten Lehrkräfte *gar nicht* (51.1 %) oder *eher nicht* (45.7 %) in der Lage sahen, Maßnahmen ergreifen zu können, um das physische, psychische oder soziale Wohlergehen von Lernenden bei der Nutzung von LLMs zu gewährleisten. Kritisch ist dies v.a. hinsichtlich der parasozialen Beziehung von Nutzer\*innen zu KI-Anwendungen – häufig aktuell medial diskutiert bezüglich der KI-Anwendung *character.ai*, in der die Nutzer\*innen mit einem von ihnen erschaffenen künstlichen Charakter diskutieren und schreiben können (Youn & Jin, 2021). Außerdem bieten vollständig unzensierte LLM-Modelle Zugang zu schadhaften Inhalten, während zensierte Modelle wiederum von KI-Anbietern dafür genutzt werden können, bestimmte Weltbilder zu transportieren.

## 4.2 KI-bezogene Meinungen

Bezüglich der in Kapitel 2.2 adressierten Meinungsdimension KI-bezogener Kompetenzen lässt sich feststellen, dass die meisten Lehrkräfte aus der Stichprobe entweder *voll und ganz* (52.7 %) oder *eher* (39.8 %) angaben, an KI interessiert zu sein. Dies ist zu interpretieren als eine grundsätzliche Aufgeschlossenheit gegenüber dem Themenfeld, auch wenn diese bei Teilnehmenden einer freiwilligen Fortbildung vorausgesetzt werden kann. Die Teilnehmenden zeigten zudem mehrheitlich eine kritisch-reflexive Grundhaltung: 90 % der Lehrkräfte führten aus, sich nicht automatisch auf die Existenz oder die Korrektheit von Quellen, die ChatGPT angibt, verlassen zu können.

Der überwiegende Anteil der Befragten schätzte den Einsatz von LLMs im Fremdsprachenunterricht als *eher* sinnvoll (52.6 %) oder als *voll und ganz* sinnvoll (24.5 %) ein. Nur wenige Teilnehmende fanden den Einsatz *gar nicht* sinnvoll (2.9 %). Diese grundsätzliche Verteilung änderte sich kaum, wenn anstelle von LLMs nach maschinellen Übersetzungstools gefragt wurde. Dass die Lehrkräfte den Unterrichtseinsatz von LLMs dennoch nicht uneingeschränkt befürworteten, könnte sich dadurch begründen lassen, dass eine große Skepsis hinsichtlich der technischen Realisierung des Einsatzes im Unterricht besteht (Nazaretsky et al., 2022). Dabei ist jedoch zu beachten, dass maschinelle Übersetzungstools meist browserbasiert sind und im Gegensatz zu LLMs schnell und unkompliziert in den Unterricht integriert werden können. Trotzdem wurden sie ähnlich bewertet. Es ist zu vermuten, dass dies daran liegt, dass Lehrkräfte maschinellen Übersetzungstools primär über die missbräuchliche Verwendung durch maschinell übersetzte Hausaufgaben begegnen und die Tools daher nicht in ihrem didaktischen Potential erkennen. Fragte man hingegen nach dem Einsatz für Lehrkräfte außerhalb des Unterrichts, so war sowohl bei LLMs (38.9 %) als auch bei maschinellen Übersetzungstools (36.0 %) eine Zunahme von Fortbildungsteilnehmenden zu beobachten, die den Einsatz dieser Tools *voll und ganz* sinnvoll fanden. Die Potentiale der Tools wurden für die eigene Nutzung also sehr wohl erkannt.

Die Lehrkräfte befürchteten mehrheitlich *gar nicht*, dass aktuell verfügbare KI-Tools den Fremdsprachenunterricht ersetzen können (75.7 %). Allerdings lässt sich aus diesen Zahlen gleichermaßen eine latente Verunsicherung herauslesen, da 18.8 Prozent der Teilnehmenden dieser Aussage nur *eher nicht* zustimmten und 5.4 Prozent *voll und ganz* oder *eher* der Überzeugung waren, dass diese Tools den Fremdsprachenunterricht ersetzen können. Diese Zahlen änderten sich strukturell nur wenig, wenn man fragte, ob LLMs in naher Zukunft den Fremdsprachenunterricht ersetzen können (61.3 % stimmten der Aussage *gar nicht* zu). Obwohl also grundsätzlich von der bleibenden Relevanz des eigenen Berufs ausgegangen wurde, stimmten die meisten Lehrkräfte der Aussage, dass sich der Fremdsprachenunterricht in naher Zukunft durch KI fundamental verändern wird, dennoch *eher* (50.9 %) oder *voll und ganz* zu (34.1 %).

Die Mehrzahl der Fortbildungsteilnehmenden gab an, Gefahren und Probleme beim Einsatz von LLMs (86.7 %) und maschinellen Übersetzungstools (91.5 %) *voll und ganz* bzw. *eher* zu erkennen. Hieraus ist keine Technikfeindlichkeit abzuleiten, sondern ein

gesundes Maß an Kritik und Skepsis. An dieser Stelle zeigen sich Parallelen zu einer Interviewstudie von Le (2023), in welcher die befragten Französischlehrkräfte in der Lage waren, über Herausforderungen und Risiken digitaler Medien im Fremdsprachenunterricht zu reflektieren, und gleichzeitig die Vermutung einer zunehmenden Relevanz entsprechender Medien für den Unterricht äußerten. Die Ergebnisse der vorliegenden Studie legen nahe, dass diese Befunde durchaus auch auf die KI-spezifischen Meinungen von Französisch- und Spanischlehrkräften übertragen werden können.

Dennoch gilt es zu berücksichtigen, dass 91.0 Prozent der Lehrkräfte sich *gar nicht* oder *eher nicht* auf den Einsatz von LLMs im Fremdsprachenunterricht vorbereitet fühlten. Dieser Befund ist insbesondere im Vergleich mit Lehrkräften anderer Fächer auffällig: Laut einer aktuellen Studie der Robert Bosch Stiftung (2025), in der  $N = 1540$  Lehrkräfte unterschiedlicher Fächer an allgemein- und berufsbildenden Schulen in Deutschland befragt wurden, gaben zumindest 38 Prozent der Befragten an, sich im Umgang mit KI-Tools für die berufliche Tätigkeit *eher sicher* oder *sehr sicher* zu fühlen. Der Vergleich der Studienergebnisse deutet somit darauf hin, dass insbesondere bei Französisch- und Spanischlehrkräften ein besonderer Weiterbildungsbedarf im Bereich des KI-gestützten Unterrichtens besteht, wenngleich die jeweils verwendeten Items nicht im Wortlaut identisch sind. Mögliche Ursachen hierfür werden - in Verbindung mit den nachfolgenden Befunden zu den KI-bezogenen Verwendungsweisen der Lehrkräfte - in Abschnitt 4.3 diskutiert.

### 4.3 KI-bezogene Verwendungsweisen

Abschließend erfolgt eine analytische Betrachtung der handlungsbezogenen Dimension KI-bezogener Kompetenzen. 44.8 Prozent der Befragten gaben an, ChatGPT bereits im Unterricht zu nutzen. Nur 34.8 Prozent der Lehrkräfte berichteten hingegen, DeepL im eigenen Unterricht zu verwenden. Fragte man nach der Nutzungshäufigkeit von LLMs wie ChatGPT für den Fremdsprachenunterricht, so führten die wenigsten Fortbildungsteilnehmenden aus, diese Tools *täglich* einzusetzen (2.9 %). Etwas weniger als ein Drittel gab an, LLMs *gar nicht* zu verwenden (30.9 %). Der Rest verteilte sich recht ausgeglichen auf *einmal im Monat* (23.2 %), *alle zwei Wochen* (18.8 %) und *mehrmals pro Woche* (22.1 %). Interessanterweise änderten sich die Nutzungshäufigkeiten kaum, wenn man nach maschinellen Übersetzungstools fragte, obwohl diese Technologie deutlich älter und einfacher zugänglich ist – DeepL oder google translate erfordern zum Beispiel keinen Account. Dies lässt sich möglicherweise damit erklären, dass das Item hauptsächlich als eigene Nutzung interpretiert wurde – nicht als Nutzung im Unterricht mit den Lernenden. Andere KI-Tools wie Dall·E, stable diffusion oder Suno finden nach Angabe von 65.3 Prozent der Befragten keine Verwendung im Fremdsprachenunterricht.

Generell fällt bei der Nutzung von LLMs auf, dass Lehrkräfte entsprechende Tools nur selten direkt ins Unterrichtsgeschehen einbanden, sondern diese primär zur Vor- und Nachbereitung des Unterrichts nutzten (vgl. Abb. 4). Der intentionale Einsatz durch Lernende im Fremdsprachenunterricht erweist sich somit eher als Ausnahme. Dies könnte u.a. darauf zurückzuführen sein, dass sich die große Mehrheit der befragten Lehrkräfte nicht ausreichend auf den Einsatz von LLMs im Fremdsprachenunterricht vorbereitet fühlte (s. Kap. 4.2) und deshalb eher von einer Implementation absah. Darüber hinaus steht die Entwicklung von Unterrichtskonzepten und Aufgaben, die eine kritisch-reflexive Verwendung von KI-Tools durch Lernende im Französisch- und Spanischunterricht erfordern, noch am Anfang. Mögliche Konzepte könnten sich beispielsweise auf die Arbeit mit sogenannten KI-Co-Texten beziehen, also auf die Analyse von Textprodukten, die in der Zusammenarbeit von Lernenden und KI-Anwendungen entstanden sind (Hartmann, 2021). Dabei ist allerdings zu beachten, dass der Stärkung der digitalisierungsbezogenen Kompetenzen von Lernenden durch Lehrkräfte hohe Bedeutung zukommt (s. Kap. 2.1) und diese zugleich eine Grundlage für die lernförderliche

Einbindung von KI-Anwendungen ins Unterrichtsgeschehen darstellen. In diesem Zusammenhang könnte auch die potenzielle Sorge von Lehrkräften, Lernende durch KI-gestützte Unterrichtskonzepte zu überfordern, eine hemmende Wirkung auf die Implementation entsprechender Tools ausüben.

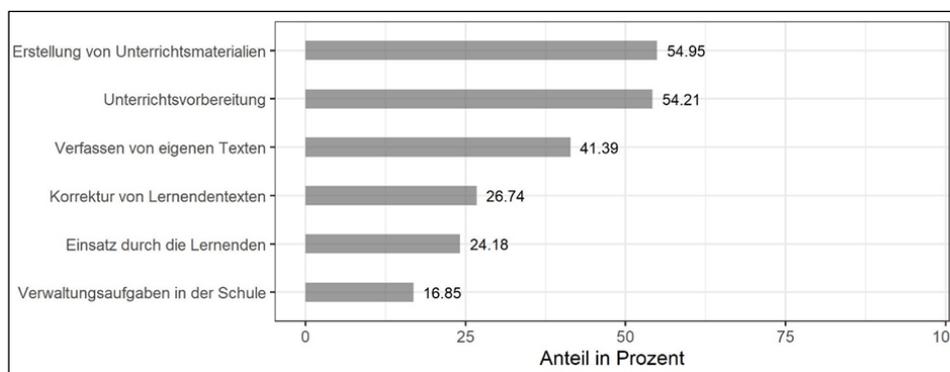


Abbildung 4: Nutzungszwecke von LLMs von Französisch- und Spanischlehrkräften

Im Gegensatz zur meinungsbezogenen Dimension stehen die Befunde zu KI-bezogenen Verwendungsweisen in Einklang mit den Untersuchungsergebnissen der Robert Bosch Stiftung (2025): Auch in dieser Studie gaben 31 Prozent der Lehrkräfte an, KI-Tools *nie* zu verwenden, wohingegen nur 2 Prozent der Befragten diese nach eigenen Angaben *täglich* nutzten. Auch die Nutzungszwecke fallen interessanterweise nahezu identisch aus: Die Verwendung von KI-Tools begrenzt sich zumeist auf das Entwickeln von Aufgaben für den Unterricht (58 %) oder auf die Unterrichtsplanung (56 %). Daran wird erkennbar, dass Französisch- und Spanischlehrkräfte mit ihrer eher zurückhaltenden KI-Nutzung keine Ausnahme darstellen, sondern in dieser Hinsicht vergleichbar mit Lehrkräften anderer Fächer sind.

## 5 Limitationen

Abschließend ist auf einige methodische und inhaltliche Limitationen der Studie hinzuweisen. Aufgrund der geringen Stichprobengröße sind die Ergebnisse nur bedingt verallgemeinerbar. Die meinungs- und handlungsbezogenen Dimensionen von KI-Kompetenzen wurden zudem nur durch Selbstauskünfte der Teilnehmenden erhoben, sodass nicht ausgeschlossen werden kann, dass sozial erwünschte Antworten gegeben wurden. Zudem ist davon auszugehen, dass gewisse Verzerrungen vorliegen, da vermutlich vor allem jene Lehrkräfte an einer Fortbildung zum KI-gestützten Unterrichten in den Fächern Französisch und Spanisch teilnehmen, die bereits im Vorfeld Interesse an diesem Themenfeld zeigen.

Durch die gegenwärtigen Entwicklungen im Feld der KI sind zudem einige Items des validierten AI-Literacy-Tests nach Hornberger et al. (2023) nicht mehr ganz aktuell (z.B. hinsichtlich des Umgangs von LLMs mit Ironie). In Hinblick auf diese Items sind Anpassungen bzw. Aktualisierungen nötig. Zu demselben Ergebnis kommt Lintner (2024) bei der Untersuchung des Instruments im Rahmen einer Meta-Analyse verschiedener AI-Literacy-Skalen. Allerdings ist zu bedenken, dass im dynamischen Feld der KI vermutlich nur bedingt Messinstrumente mit überdauernder Gültigkeit entwickelt werden können.

## 6 Fazit und Implikationen

Zusammenfassend lässt sich festhalten, dass Französisch- und Spanischlehrkräfte ein großes Interesse an KI-gestütztem Unterrichten zeigen und entsprechende Anwendungen bereits nutzen, allerdings eher zur Vor- und Nachbereitung von Unterricht als im Unterrichtsgeschehen selbst. Im Gegensatz dazu verfügen Französisch- und Spanischlehrkräfte – ähnlich wie Lehrkräfte nicht-fremdsprachlicher Fächer (Nazaretsky et al., 2022) – über eher rudimentär ausgeprägtes KI-bezogenes Wissen und fühlen sich mehrheitlich nicht ausreichend auf den Einsatz von KI im Fremdsprachenunterricht vorbereitet. Insbesondere die Konfrontation der Lernenden mit KI-Tools im Unterricht nimmt bislang nur eine Randstellung ein. Gründe dafür könnten, neben kaum vorhandenen didaktischen und empirisch validierten Konzepten, fehlende Ausstattung, mangelhafte Bereitstellung von in der Schule nutzbaren KI-Tools und Probleme bei der Adaption bestehender Tools (z.B. ChatGPT) aufgrund datenschutzrechtlicher Vorgaben sein.

Neben einer stärkeren Berücksichtigung von KI-bezogenen Themen im Rahmen der Lehramtsausbildung verdeutlichen die Befunde die Notwendigkeit, gezielte Fortbildungsangebote für die dritte Phase der Lehrkräftebildung zu entwickeln. Wie die Ergebnisse der vorliegenden Studie zeigen, sollten diese v.a. KI-bezogenes Grundlagenwissen vermitteln, den Abbau von KI-bezogenen Fehlkonzepten begünstigen und zentrale KI-Kompetenzen, z.B. die „Prompt Literacy“, fördern (vgl. Antonenko & Abramowitz, 2023; Nazaretsky et al., 2022). Da nur wenige Lehrkräfte mit konkreten KI-gestützten Unterrichtsaktivitäten vertraut sind, scheint es zudem geboten, im Rahmen von Weiterbildungsangeboten konkrete Einsatzszenarien zu fokussieren, da ansonsten keine umfassende Aufnahme in die unterrichtliche Praxis erwartet werden kann. Es ist ebenso von zentraler Bedeutung, über Makrothemen wie ethische Dimensionen bei der KI-Nutzung zu informieren, allerdings sollte stets ein Praxisbezug vorhanden sein. Dabei ist insbesondere darauf zu achten, die heterogenen digitalen Ausgangsbedingungen von Lehrkräften und Schulen zu berücksichtigen. Die Fortbildungen „Der Einsatz Künstlicher Intelligenz beim Lehren und Lernen der französischen/spanischen Sprache“ (vgl. Kap. 3.1) adressieren die skizzierten Bedarfe und können – trotz der noch ausstehenden Wirksamkeitsuntersuchung – als Ausgangsbasis für weitere Angebote dienen.

Auf Basis der bisher durchgeführten Fortbildungen (vgl. Kap. 3.1) konnte die Erkenntnis gewonnen werden, dass insbesondere Online-Fortbildungen oder Blended-Learning-Konzepte inkl. synchroner Anteile via Videokonferenz mit den Teilnehmenden sinnvoll scheinen. Ein- oder halbtägige Bildungsformate erweisen sich bei einem derart neuen und komplexen Thema, das zudem in die eigene Unterrichtspraxis überführt werden soll, als wenig zielführend. Aufgrund der Anlage der durchgeführten Fortbildungen ist eine Teil-Auslagerung in Form von Selbstlernmodulen möglich, strukturierte und angeleitete Praxisphasen bleiben dabei jedoch ein essenzieller Bestandteil.

Die im Rahmen dieser Studie entstandenen Bildungsmodule sind den jeweiligen Landesinstituten zugänglich und auf der Homepage des Forschungsprojektes<sup>4</sup> nach Anmeldung auf der Meta-Videoplattform [www.unterrichtsvideos.net](http://www.unterrichtsvideos.net) abrufbar. Zusätzlich profitiert die Dissemination von dem Umstand, dass alle Lerneinheiten auch in atomisierter Form als sog. Learning-Nuggets vorliegen und durch SCORM-Kompatibilität in jedem moodle-basierten LMS eingesetzt werden können. Die Fortbildungen selbst werden im Rahmen der GLA fest verankert.

Abschließend bleibt festzustellen, dass der datenschutzkonforme Einsatz von KI-Modellen in der dritten Phase der Lehrkräftebildung bislang durch institutionelle Rahmenbedingungen erschwert ist. Es eröffnen sich jedoch erste Perspektiven, etwa durch das AIToolLab der Goethe-Universität Frankfurt, das datenschutzkonforme Anwendungen selbstgehosteter oder kommerzieller Modelle in der Hochschullehre ermöglicht –

---

<sup>4</sup> <https://vigor.studiumdigitale.uni-frankfurt.de/vifonet-lerneinheiten/>

bislang jedoch vorwiegend für Studierende. Es wäre wünschenswert, diese Ansätze weiterzuentwickeln, um künftig auch Lehrenden und Lernenden strukturierte und curricular eingebettete Erfahrungen mit LLMs zu ermöglichen.

## Literatur und Internetquellen

- Al-khresheh, M.H. (2024). Bridging Technology and Pedagogy from a Global Lens: Teachers' Perspectives on Integrating ChatGPT in English Language Teaching. *Computers and Education: Artificial Intelligence*, 6, 100218. <https://doi.org/10.1016/j.caeai.2024.100218>
- Antonenko, P. & Abramowitz, B. (2023). In-service Teachers' (Mis)conceptions of Artificial Intelligence in K-12 Science Education. *Journal of Research on Technology in Education*, 55 (1), 64–78. <https://doi.org/10.1080/15391523.2022.2119450>
- Bajaña Ordoñez, M.A. & Espinal Bailón, A.D. (2024). *Exploring Teacher's Perceptions of Artificial Intelligence to Enhance English Speaking Skills*. <https://repositorio.upse.edu.ec/handle/46000/12031>
- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenzen von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Waxmann.
- Blömeke, S., Gustafsson, J.E. & Shavelson, R.J. (2015). Beyond Dichotomies: Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223 (1), 3–13. <https://doi.org/10.1027/2151-2604/a000194>
- Dimitriadou, E. & Lanitis, A. (2023). A Critical Evaluation, Challenges, and Future Perspectives of Using Artificial Intelligence and Emerging Technologies in Smart Classrooms. *Smart Learning Environments*, 10, 12. <https://doi.org/10.1186/s40561-023-00231-3>
- Grum, U. & Legutke, M.K. (2022). Sampling. In D. Caspari, F. Klippel, M.K. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch* (2. vollst. überarb. & erw. Aufl., S. 85–96). Narr Francke Attempto.
- Hartmann, D. (2021). Künstliche Intelligenz im DaF-Unterricht? Disruptive Technologien als Herausforderung und Chance. *Informationen Deutsch als Fremdsprache*, 48 (6), 683–696. <https://doi.org/10.1515/infodaf-2021-0078>
- Hinrichsen, J. & Coombs, A. (2013). The Five Resources of Critical Digital Literacy: A Framework for Curriculum Integration. *Research in Learning Technology*, 21 (1), 1–16. <http://dx.doi.org/10.3402/rlt.v21.21334>
- Hornberger, M., Bewersdorff, A. & Nerdel, C. (2023). What Do University Students Know about Artificial Intelligence? Development and Validation of an AI Literacy test. *Computers and Education: Artificial Intelligence*, 5, 100165. <https://doi.org/10.1016/j.caeai.2023.100165>
- Horz, H. & Schulze-Vorberg, L. (2017). *Digitalisierung in der Hochschullehre*. Konrad-Adenauer-Stiftung.
- Huwer, J., Irion, T., Kuntze, S., Schaal, S. & Thyssen, C. (2019). Von TPaCK zu DPaCK – Digitalisierung im Unterricht erfordert mehr als technisches Wissen. *MNU Journal*, 72 (5), 356–364.
- Hwang, Y., Lee, J.H. & Shin, D. (2023). What Is Prompt Literacy? An Exploratory Study of Language Learners' Development of New Literacy Skill Using Generative AI. *arXiv*. <https://doi.org/10.48550/arXiv.2311.05373>
- Irion, M., Korell, J.L. & Ißler, R. (2025). Kreatives Schreiben und Künstliche Intelligenz – eine empirische Untersuchung zu Arbeits- und Reflexionsprozessen von Französischlernenden der Sekundarstufe II am Beispiel des inneren Monologs. *Zeitschrift für Romanische Sprachen und ihre Didaktik*, 19 (1), 93–123. [https://elibra.ry.utb.de/doi/10.24216/ZRomSD-2025-1-9783838280424\\_005](https://elibra.ry.utb.de/doi/10.24216/ZRomSD-2025-1-9783838280424_005)

- Korell, J.L. & Albrecht, P. (2025). Artificial Intelligence in French and Spanish Foreign Language Education: A Systematic Review and Future Perspectives. *Language Education and Technology*, 5 (1), 21–41.
- Korell, J. L., Nijhawan, S., Ibler, R. & Viebrock, B. (2025). Fremdsprachenlernende und Künstliche Intelligenz – Eine empirische Untersuchung zu Kenntnissen, Meinungen und Nutzungsweisen von Englisch-, Französisch- und Spanischschüler:innen der Sekundarstufen I und II. *Zeitschrift für Fremdsprachenforschung (ZFF)*, 36 (1), 135–156. <https://doi.org/10.3278/ZFF2501W009>
- Le, M. (2023). Einstellungen und medienbezogene Kompetenzen von Französischlehrkräften zum Einsatz digitaler Medien im Fremdsprachenunterricht – eine empirische Untersuchung. *Zeitschrift für romanische Sprachen und ihre Didaktik*, 17 (2), 39–70.
- Lindner, A. & Berges, M. (2020). Can You Explain AI to Me? Teachers' Pre-Concepts about Artificial Intelligence. *IEEE Frontiers in Education Conference (FIE)*, 1–9. <https://doi.org/10.1109/FIE44824.2020.9274136>
- Lintner, T. (2024) A Systematic Review of AI Literacy Scales. *npj Science of Learning*, 9, 50. <https://doi.org/10.1038/s41539-024-00264-4>
- Lorenz, U. & Romeike, R. (2023). What Is AI-PACK? – Outline of AI Competencies for Teaching with DPACK. In J.-P. Pellet & G. Parriaux (Hrsg.), *Informatics in Schools. Beyond Bits and Bytes: Nurturing Informatics Intelligence in Education* (S. 13–25). Springer. [https://doi.org/10.1007/978-3-031-44900-0\\_2](https://doi.org/10.1007/978-3-031-44900-0_2)
- Meyasa, K.J., Artini, L.P. & Marsakawati, N.P.E. (2024). EFL Teachers' Perspectives on the Integration of ChatGPT in Writing Assessment. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 12 (1), 28–34. <https://doi.org/10.23887/jpbi.v12i1.78618>
- Mohammad Ali, A. (2023). *An Intervention Study on the Use of Artificial Intelligence in the ESL Classroom: English teacher perspectives on the Effectiveness of ChatGPT for Personalized Language Learning*. <https://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-61339>
- Nazaretsky, T., Ariely, M., Cukurova, M. & Alexandron, G. (2022). Teachers' Trust in AI-powered Educational Technology and a Professional Development Program to Improve It. *British Journal of Educational Technology*, 53 (4), 914–931. <https://doi.org/10.1111/bjet.13232>
- Ng, D.T.K., Leung, J.K.L., Chu, S.K.W. & Qiao, M.S. (2021). Conceptualizing AI Literacy: An Exploratory Review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Redecker, C. & Punie, Y. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Publications Office of the European Union. <https://publications.jrc.ec.europa.eu/repository/handle/JRC107466>
- Reinders, H. (2016). *Qualitative Interviews mit Jugendlichen führen: Ein Leitfaden* (3. durchges. u. erw. Aufl.). Oldenbourg. <https://doi.org/10.1515/9783110469561>
- Robert Bosch Stiftung (2025). *Deutsches Schulbarometer: Befragung Lehrkräfte. Ergebnisse zur aktuellen Lage an allgemein- und berufsbildenden Schulen*. Robert Bosch Stiftung. [https://www.bosch-stiftung.de/sites/default/files/documents/2025-06/DeutschesSchulbarometer\\_Lehrkräfte\\_2025.pdf](https://www.bosch-stiftung.de/sites/default/files/documents/2025-06/DeutschesSchulbarometer_Lehrkräfte_2025.pdf)
- Schramm, K. (2022). Empirische Forschung. In D. Caspari, F. Klippel, M.K. Legutke & K. Schramm (Hrsg.), *Forschungsmethoden in der Fremdsprachendidaktik: Ein Handbuch* (2. vollst. überarb. & erw. Aufl., S. 50–59). Narr Francke Attempto.
- Strasser, T. (2023). Schwache KIs, starke Performanz? Form und Wirkung von KI-gestützten Erklärvideos im Englischgrammatikunterricht der Unterstufe. *Zeitschrift für Fremdsprachenforschung*, 34 (1), 75–102.
- Weinert, F.E. (2014). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (3. aktual. Aufl., S. 17–31). Beltz.

- Wienrich, C., Carolus, A., Markus, A. & Augustin, Y. (2022). *AI Literacy: Kompetenzdimensionen und Einflussfaktoren im Kontext von Arbeit*. Denkfabrik Digitale Arbeitsgesellschaft.
- Youn, S. & Jin, S.V. (2021). “In A.I. We Trust?” The Effects of Parasocial Interaction and Technopian versus Luddite Ideological Views on Chatbot-based Customer Relationship Management in the Emerging “Feeling Economy”. *Computers in Human Behavior*, 119, 106721. <https://doi.org/10.1016/j.chb.2021.106721>

## Beitragsinformationen

### Zitationshinweis:

Irion, M., Breitenbach, S., Korell, J.L. & Ibler, R. (2025). KI-bezogene Kompetenzen von Französisch- und Spanischlehrkräften – eine empirische Untersuchung. *HLZ – Herausforderung Lehrer\*innenbildung*, 8 (1), 304–321. <https://doi.org/10.11576/hlz-7961>

Eingereicht: 26.03.2025 / Angenommen: 08.08.2025 / Online verfügbar: 16.09.2025

ISSN: 2625–0675



Dieses Werk ist freigegeben unter der Creative-Commons-Lizenz CC BY-SA 4.0 (Weitergabe unter gleichen Bedingungen). Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Tabellen, Zitate etc.) sind von der CC-Lizenz ausgenommen. Für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen. <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>

## English Information

**Title:** AI-related Competences of French and Spanish Teachers: An Empirical Study

**Abstract:** The integration of artificial intelligence (AI) into foreign language teaching offers new didactic possibilities but also presents challenges for teachers in terms of critical reflection and the use of corresponding tools. This study examines the AI-related competences of  $N = 279$  French and Spanish teachers through knowledge tests and self-assessment questionnaires. Theoretically, the study is based on key AI competence models, such as the AI-PACK model (Lorenz & Romeike, 2023), AI Literacy (Ng et al., 2021), and Prompt Literacy (Hwang et al., 2023). Methodologically, it follows a quantitative-descriptive approach aimed at inventorying teachers’ AI-related competences, including their AI-related knowledge, opinions, and self-reported use of AI tools. The findings show that teachers rate their AI competences as rather low, which aligns with the results of the knowledge tests. However, there is a strong interest in AI, with both opportunities and challenges recognized. Concrete potential for further training is particularly evident in the didactic integration of AI tools in foreign language teaching, as their use in the classroom has so far been sporadic. The study emphasizes the need for targeted professional development programs to promote the critical and reflective use of AI among French and Spanish teachers.

**Keywords:** Artificial Intelligence (AI); teacher research; competence; French lessons; Spanish lessons