Didaktik
Deutsch

Alyssa Kutzner

Cross-linguistic writing with AI

Investigating transfer effects in L1/L2 and FL contexts

Abstract

This study explores how multilingual 9th-grade students use generative AI as a writing assistant in German (L1/L2) and English (FL) writing contexts by focusing on strategy use, cross-linguistic transfer, and proficiency-related variation. Drawing on a mixed-methods design that combines qualitative and quantitative data, the analysis suggests that generative AI may introduce a new mediating layer in the writing process, enabling learners to externalize cognitive subtasks such as planning, drafting, and revising. Five main AI writing strategies were identified, ranging from planning and linguistic optimization to metacognitive reflection. Overall, findings indicate patterns consistent with cross-linguistic transfer of AI writing strategies in this exploratory sample, highlighting AI's potential to foster metacognitive awareness and cross-linguistic strategy development in multilingual writing. However, proficiency appears to moderate these patterns: More proficient (Level B2) learners of English as a foreign language show greater strategic stability in their AI use, whereas lower-proficiency learners engage with AI in a more situational and cautious manner, a pattern that is also reflected in two illustrative learner cases.

Keywords: AI-assisted writing • cross-linguistic transfer • AI writing strategies • multilingual education • language proficiency

Didaktik Deutsch
Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur
Special Issue (2026) – Writing with AI. S. 123–157
DOI: 10.21248/dideu.971

1 Introduction and objectives

Writing in school increasingly unfolds at the intersection of cognitive strategy use, multilingual repertoires, and emerging AI tools. The advent of generative artificial intelligence (AI)—particularly large language models (LLMs) such as GPT-3.5, GPT-4.0, the recently released GPT-5.0, or even GPT-5.2—has transformed how writing is conceptualized and taught. Learners now access systems that can generate coherent, context-sensitive text and provide feedback on structure, style, and content, thereby challenging established notions of authorship, strategy use, and language learning. This development necessitates a rethinking of writing instruction: How can AI be meaningfully embedded without undermining autonomy or critical awareness? This question has been discussed in both emerging international research on AI-mediated higher-order regulation and evaluative processes in writing (e.g., Gonsalves, 2024; Shen & Teng, 2024; Song & Song, 2024) as well as in German-speaking contexts (e.g., Fürstenberg & Müller, 2024; Rezat & Schindler, 2025; Steinhoff & Lehnen, 2025). Although research has long shown that effective writers plan, draft, and revise strategically (Harris et al., 2010; Lei, 2016; Manchón & Roca De Larios, 2007; Raoofi et al., 2017; Roca De Larios et al., 2001), and that such strategies can transfer across languages (Cummins, 1980, 2000; García & Wei, 2015; Herdina & Jessner, 2002; Odlin, 1989), less is known about how learners mobilize these resources when a generative AI assistant is available throughout the writing process.

The present contribution is part of a project on the transformative impact of generative AI on learner writing in secondary education. As the second of three interconnected studies, it shifts from learners' L1/L2 (German) writing to multilingual contexts and examines the potential of *AI writing strategies* to support foreign-language (FL) learning and cross-linguistic transfer. Under authentic classroom conditions in a Grade-9 German *Gymnasium* (academic track secondary school), learners were introduced to AI as a writing assistant and encouraged to deploy it autonomously and purposefully across planning, drafting, and revising phases of summary writing in German (L1/L2) and English (FL). Using a mixed-methods design, standardized reading assessments (*LGVT 5-12+*; *Cambridge English B1 Preliminary for Schools*) and student–AI chat logs were analyzed to identify patterns of AI writing strategies and to examine the extent to which these strategies transfer across languages.

Building on models of writing, cross-linguistic influence, and metacognition, AI was treated as a potential scaffold for planning, drafting, and revising—benefits that depend on learners' critical control of AI output. Accordingly, the study addresses three questions:

1. What observable patterns of AI usage emerge in L1/L2 and FL writing contexts?
2. To what extent does transfer of *AI writing strategies* occur across languages?
3. How does the use of *AI writing strategies* relate to language proficiency in L1/L2 and FL contexts?

Preliminary findings from the first study within this research project (Kutzner, in press) indicate that learners develop emerging *AI writing strategies*: flexible repertoires that learners use to strategically decompose writing tasks and use AI for planning, drafting, or revision support. These strategies appear to resemble established writing scaffolds and hold potential for fostering writing development, particularly in multilingual or foreign language contexts. Comparable observations have also been reported in recent international research on AI-supported English as a foreign language (EFL) writing (e.g., Xu et al., 2025). Furthermore, findings suggest that interaction with AI may enhance learners' metalinguistic awareness, aligning with principles of the *noticing hypothesis* (Schmidt, 1990) proposing that *noticing* facilitates language learning by helping learners become aware of gaps between their own output and target-language input. For instance, learners report paying closer attention to lexical and syntactic patterns suggested by AI (Kutzner, in press); and this, in turn, appears to affect their linguistic output in both their first and foreign language writing.

By mapping strategy profiles onto language proficiency indicators and quantifying cross-linguistic overlap, the present study contributes: (a) an empirically grounded typology of AI-supported writing behaviors in school contexts; (b) indications on their transferability between L1/L2 and FL; and (c) implications for teaching AI-informed, process-oriented writing. The study raises important questions for the integration of AI in language teaching, particularly in fostering AI literacy (Allen & Kendeou, 2023; Ng et al., 2021) and strategic competence in multilingual learners. It also contributes to understanding how artificial interlocutors can scaffold writing in ways that support noticing, transfer, and cross-linguistic development.

2 Theoretical framework

This study is primarily grounded in process-oriented writing theory and models of cross-linguistic transfer. Concepts such as *noticing* and *AI literacy* serve as complementary perspectives to contextualize AI-mediated strategy use. Writing is widely conceptualized as a complex cognitive activity involving planning, text production, and revision (Hayes, 2012; Hayes & Flower, 1980). Skilled writers employ flexible writing strategies that enable them to manage cognitive load and maintain coherence. Research in writing pedagogy demonstrates that explicit instruction in such strategies significantly improves writing quality and self-regulation (Harris et al., 2010; Harris et al., 2011). Writing strategies support the writing process across its various phases (Hayes, 2012; Hayes & Flower, 1980), helping writers to plan, compose, and revise their texts more effectively.

Research on cross-linguistic influence and metacognition provides a comprehensive framework for understanding how linguistic and cognitive processes interact in multilingual learners. Cummins's (1980, 2000) *common underlying proficiency* (CUP) model assumes that bilingual individuals draw on a shared reservoir of cognitive and linguistic abilities, enabling knowledge and skill transfer across languages (see also Swain & Lapkin, 1995) once a sufficient proficiency threshold is reached (*threshold hypothesis*: Cummins, 1976). Complementing this view, Odlin (1989) distinguishes between positive (facilitative) and negative (interfering) transfer effects arising from similarities and discrepancies between languages. Later models such as Herdina and Jessner's (2002) *dynamic model of multilingualism* (DMM) and García and Wei's (2015) concept of translanguaging reconceptualize multilingual competence as an adaptive and fluid system in which learners flexibly draw on their full linguistic repertoire to construct meaning. Within this dynamic perspective, writing in multilingual contexts activates an integrated set of linguistic and metacognitive resources. Strategic skills developed in one language—particularly those related to planning, organization, and revision—can thus transfer to others. This process is mediated by metalinguistic awareness (Jessner, 2006) and *noticing* (Schmidt, 1990) that enable learners to consciously monitor linguistic form, compare output with target input, and refine their writing across languages. Schmidt's (1990) *noticing hypothesis* provides a key theoretical link between awareness and learning: Language acquisition occurs when learners notice the gap between their output and target-language input. In writing, this noticing can be facilitated through comparison, reflection, and feedback.

Generative AI introduces a new mediating layer in writing processes. Through prompting, learners can externalize subtasks such as planning, drafting, and revising (Putri et al., 2025). In doing so, AI reshapes the ecology of writing (Warschauer et al., 2023), and holds potential for fostering strategy development and metacognitive reflection. However, these benefits depend on learners' ability to regulate and evaluate AI output critically—a central component of AI literacy (Allen & Kendeou, 2023; Ng et al., 2021). Although research on AI-assisted writing has increased since the release of ChatGPT 3.5 in late 2022, significant gaps persist. Most existing studies focus on university students and emphasize the advantages of AI such as its support for idea generation, feedback, and language enhancement (Kim et al., 2025; Li, 2025; Siregar & Hz, 2025; Werdiningsih et al., 2024). Studies involving school-aged learners are fewer, but likewise tend to address isolated stages of the writing process—for example, AI-supported feedback rather than holistic writing development (Castro et al., 2025; Chen et al., 2024; Kangasharju et al., 2022; Zheldibayeva et al., 2025; or, for German L1 contexts, Fürstenberg, 2025; Pitzen, 2025). However, only a few of these studies explore how learners strategically integrate AI across the entire writing process; and this holds particularly for elementary and secondary education (Alghasab, 2025; Chan & Lee, 2025; or, for German L1 contexts, Bodora et al., 2025). Existing research often examines isolated stages such as revision or editing, or raises concerns about potential overreliance on AI without addressing broader, process-oriented strategies (Putri et al., 2025). Yet, such practices may foster noticing and strategic flexibility, especially among multilingual learners who navigate between linguistic systems. This study therefore seeks to advance the understanding of how these strategies are applied and potentially transferred across L1/L2 and FL contexts. Because empirical research in this area is still limited, the study adopts an exploratory design aimed at identifying recurring patterns.

3 Context and study design

The study was conducted at a public German *Gymnasium* (academic track secondary school) in the state of North Rhine-Westphalia during the 2024/25 school year with 20 ninth-grade students (aged 14–15). German was the primary language (L1) for most learners, although 13 students reported bilingual backgrounds (e.g., Turkish, Russian, Arabic). English was taught as a compulsory foreign language (FL) from Grade 2 onward, and, as stipulated by the Ministry of School and Education (*Ministerium für Schule und Bildung*, 2019), students in Grade 9 at German *Gymnasien* (G9 track) are expected to achieve at least level B1 of the Common European Framework of Reference for Languages (CEFR). All students had access to personal laptops provided by the school. Although they were familiar with digital learning tools, the integration of generative AI into the writing process was a rather novel experience, at least in classroom contexts. Written consent was obtained from students and parents.

The research followed a comparative, cross-linguistic, mixed-methods design with one assessment each in German (L1/L2) and English (FL) conducted under authentic classroom conditions. Following an introduction to AI's potential as a writing assistant in the writing process, learners were encouraged to use AI autonomously and purposefully throughout the writing process. This design allowed learners to build individual scaffolds to support their planning, composing, and revising activities. Each assessment included a standardized reading test, an AI-supported writing task, and a survey on learners' experiences and attitudes. The entire assessment was conducted over the course of three consecutive days.

Reading proficiency served as a cognitive indicator of language proficiency and was measured using the *LGVT 5-12+* (Schneider et al., 2017) for German (L1/L2)—assessing reading comprehension, speed, and accuracy—and the *Cambridge English* (2018) *B1 Preliminary for Schools* reading test for English (FL). Reading proficiency was selected as an indicator of broader language competence, because AI-assisted writing requires learners to read, interpret, and evaluate AI-generated output. Reading comprehension integrates decoding, vocabulary knowledge, and higher-order inferencing processes that overlap with cognitive demands of writing. Writing competence, by contrast, is difficult to operationalize independently of surface-level accuracy in small exploratory samples. The Cambridge Reading component was chosen for English due to its CEFR alignment and its separate assessment of reading, ensuring construct clarity and cross-linguistic comparability with the LGVT.

Following the reading component, learners completed a summary-writing task based on a short story (*USE 2030* by Bob Blume, 2025, 1,328 words, for German; *The Text* by Brendan Dunne, 2012, 894 words, for English). Students wrote a summary (in German, an *Inhaltsangabe*)—that is, the English and German realizations of the same genre-specific text pattern. In line with current pedagogical practice in Germany, EFL instruction follows a *task-supported language learning* (TSLL) approach. Unlike fully task-based approaches, TSLL embeds communicative tasks within a syllabus that focuses on particular topics or linguistic structures (Müller-Hartmann & Schocker-von Ditfurth, 2011, p. 24). Within this framework, tasks serve as additional pedagogic methods that can be utilized to achieve various curricular goals such as developing intercultural, literary, and communicative competencies (Summer, 2012, p. 10). Similarly, writing instruction in German (L1/L2) education promotes *profilierte Schreibaufgaben* ([authentic and contextually embedded writing tasks]; Bachmann & Becker-Mrotzek, 2010) that situate writing within authentic, functional, communicative, and social contexts to enhance motivation and writing quality. In both the German and English assessments, the task design and accompanying instructions were aligned with these pedagogical principles, ensuring that the writing tasks were authentic, purpose-oriented, and communicatively meaningful (in accordance with Breen, 1985, pp. 61–69; Ellis, 2009, p. 6; Guariento & Morley, 2001, pp. 349–351). In both contexts, the summary task was embedded in a simulated classroom conference scenario in which students should present a short story to their peers. The summary (in German, an *Inhaltsangabe*) thus functioned not merely as a self-directed writing exercise but as an audience-oriented informational text intended to inform and engage classmates who had not read the story. This framing situates the writing task within a socially and communicatively meaningful classroom practice. Learners were encouraged to use GPT-4o via the GDPR-compliant educational platform *fobizz.com* to assist in planning, drafting, and revising their texts. All interactions were automatically logged.

The dataset includes 40 AI chat logs (20 for German and 20 for English), the corresponding student texts, as well as reading test results. All AI prompts were successfully categorized into the identified AI writing strategies; no prompt fell outside the analytical framework. Quantitative data were analyzed using Excel and JASP for descriptive and correlational statistics, whereas qualitative data were analyzed in MAXQDA using qualitative content analysis (Kuckartz & Rädiker, 2023). The analysis focuses on the relationship between reading performance and prompt-based strategy use in AI-assisted writing.

4 Methods and data analysis

The study was designed as an exploratory mixed-methods investigation combining quantitative and qualitative analyses to examine the relationship between reading proficiency and AI-assisted writing behavior. The study employed standardized reading tests *LGVT 5-12+* (Schneider et al., 2017) for German (L1/L2) and *Cambridge English* (2018) *B1 Preliminary for Schools Reading Test* for English (FL) alongside qualitative content analysis (Kuckartz & Rädiker, 2023) using MAXQDA in order to examine the relationship between patterns of AI strategy use and linguistic competence using reading proficiency as an indicator. All quantitative results are strictly descriptive and serve exploratory pattern identification rather than inferential validation. Given the small sample size, statistical precision should not be overinterpreted.

The *LGVT* reading test was scored according to the normative tables for Grade 9 *Gymnasium* (academic track secondary school). Raw scores were adjusted using the test-specific timing correction factor; corresponding percentile ranks were determined and subsequently converted into standardized *T* scores. Although percentile ranks represent an individual's relative standing within the normative sample, they are nonlinear and compress differences at the distributional extremes. In contrast, *T* scores (*M* = 50, *SD* = 10) are based on the normal distribution, allowing for meaningful statistical operations, cross-test comparability, and precise interpretation of distance from the mean. According to the *LGVT* manual, values below 30 indicate *very low* performance (PR 0–2.3); 30–39, *below average* (PR 2.3–15.9); 40–49, *lower average range* (PR 15.9–50); 50–59, *upper average range* (PR 50–84.1); 60–69, *above average* (PR 84.1–97.7); and 70–76, *very high* performance (PR 97.7–100).

The *Preliminary* raw results were converted into *Cambridge English Scale Scores*, enabling direct comparability across test versions. According to the *Cambridge English Scale* (2025), scores of 160–170 correspond to *Grade A* (CEFR level B2); 153–159, to *Grade B* (B1); 140–152, to *Grade C* (B1); and 120–139, to *Fail* (A2 level).

Table 1 presents the descriptive statistics (mean, median, standard deviation, range, minimum, and maximum) for the three *LGVT* subtests—reading comprehension (RC), reading speed (RS), and reading accuracy (RA)—alongside the *Cambridge English Scale Score*.

Table 1

Descriptive Statistics for LGVT and Cambridge English Scale Score

	Reading comprehension (RC) (*Leseverständnis (LV)*) *T* score	Reading speed (RS) (*Lesegeschwindigkeit (LGS)*) *T* score	Reading accuracy (RA) *(Lesegenauigkeit (LGN)) T score*	*Cambridge Scale Score*
M	44.35	42.16	52.75	147.55
Mdn	43.50	41.00	50.50	150.00
SD	8.60	7.36	10.85	14.26
Range	30.00	33.00	40.00	50.00
Min	30.00	27.00	33.00	120.00
Max	60.00	60.00	73.00	170.00

As shown in Table 2, all *LGVT* subtests (reading comprehension, reading speed, and reading accuracy) and the *Cambridge English Scale Score* are significantly intercorrelated, indicating that they assess closely related dimensions of reading competence. This pattern suggests that the tests capture overlapping cognitive processes involved in higher-level reading such as text comprehension, inference generation, and the integration of information across sentences. The strong associations between comprehension, speed, and accuracy imply that efficient decoding and fluent word recognition support overall understanding, whereas comprehension performance in turn reflects the successful coordination of these subskills. The close relationship between the *LGVT* performance and the *Cambridge English Scale Score* further supports the validity of both instruments as complementary measures of general reading proficiency across L1/L2 and FL contexts. Accordingly, all subsequent analyses considered the three *LGVT* subtests alongside the *Cambridge English Scale Score* to provide a comprehensive and multidimensional account of learners' reading competence.

Table 2

Spearman Correlations (rs) Between LGVT Subtests and Cambridge English Scale Score

Variable		Reading comprehension (RC)	Reading speed (RS)	Reading accuracy (RA)	Cambridge Scale Score
1. Reading comprehension (RC)	r_s	—			
	p	—			
2. Reading speed (RS)	r_s	.834***	—		
	p	< .001	—		
3. Reading accuracy (RA)	r_s	.837***	.511*	—	
	p	< .001	.021	—	
4. Cambridge Scale Score	r_s	.702***	.600**	.630**	—
	p	< .001	.005	.003	—

* p < .05. ** p < .01. *** p < .001.

A detailed overview of the results is provided in Appendix A in which participants' *Cambridge Preliminary* reading scores are sorted according to their performance and accompanied by the corresponding *LGVT* T-scores for reading comprehension (RC), reading speed (RS), and reading accuracy (RA) for each learner.

Whereas the quantitative measures provided a detailed profile of learners' reading proficiency, the qualitative analysis focused on their interactional and strategic behavior when engaging with AI during writing. The qualitative content analysis was conducted combining inductive and deductive coding procedures. The analytical framework drew on AI writing strategies that were inductively developed in a previous study of similar design (Kutzner, in press). In that earlier research, student–AI chat logs (*N* = 20) were openly reviewed, and prompts were coded inductively in order to identify recurring functional actions in learners' interaction with AI. From this coding process, five categories of AI writing strategies were derived. AI writing strategies are understood as an extension of established writing strategies (e.g., Harris et al., 2010; Lei, 2016; Manchón & Roca De Larios, 2007; Raoofi et al., 2017; Roca De Larios et al., 2001). They regulate not only the learner's own writing process but also the deliberate integration of AI into specific phases of that process. Importantly, the prompts themselves were not conceptualized as strategies. Rather, they were treated as observable operationalizations of underlying strategic decisions made by learners. The strategy labels build directly on the previously coded strategy categories. Each label combines: (a) an orientation descriptor reflecting the learner's strategic orientation (e.g., focused, holistic, functional, differentiated, metacognitive) with (b) a term referring to the function of the strategy in the writing process (e.g., structuring, optimizing, pragmatizing, limiting, questioning), thus describing how and for what purpose learners use AI in writing. The identified strategy categories included (the example prompts are illustrative model prompts that represent typical patterns identified in the data):

1. *Focally-structuring strategies* aimed at planning and organizing the writing process—for example, by generating outlines or writing plans ("Give me a writing plan to write a summary of a short story").
2. *Holistically-optimizing strategies* supporting linguistic and stylistic elaboration—for instance, through requests for useful phrases or synonyms ("Tell me useful phrases to write a summary of a short story").
3. *Functionally-pragmatizing strategies* oriented toward efficient task completion, such as delegating larger text operations to AI (e.g., generating a full summary).
4. *Differentially-limiting strategies* characterized by selective AI use, often limited to one or very few prompts.
5. *Metacognitively-inquiring strategies* serving reflective purposes including feedback-seeking ("Is this a good summary?"), revision support ("Give me a revision plan to revise a summary of a short story." "Is the following sentence correct: '. . . '?"), and clarification of content or language ("What does X mean?").

Because individual students employed multiple strategies, strategy use was not treated as mutually exclusive or as fixed learner types. Rather, learners shifted between different strategies over the course of the writing process. For example, planning may be characterized by Strategy-1-typical structuring practices; drafting, by Strategy-2-typical linguistic optimization; and revision, by Strategy-5-typical metacognitive inquiry. In line with process-oriented writing models (Hayes, 2012; Hayes & Flower, 1980), learners oscillated between planning, drafting, and revising, while AI assumed different functional roles—such as cognitive offloading, linguistic support, or reflective scaffolding—within this dynamic process. Thus, the strategies are conceptualized as situational and phase-sensitive usage patterns in AI-supported writing rather than as stable user types. In this sense, "phase-sensitive" does not imply a one-to-one correspondence with the classical three-phase writing model (planning, drafting, revising). Although certain strategies can be associated with particular writing phases, not all learners used AI in every phase, and some phases were completed entirely without AI support. Rather, the data indicate that learners selectively integrated AI depending on individual preferences, because the use of AI was neither obligatory nor uniformly distributed across phases. The strategies therefore do not map directly onto the three-phase writing model, but instead reflect selective and context-dependent configurations of AI integration within the broader writing process.

In addition to the five main strategies, the previous study identified three supplementary categories that, while not representing independent strategy types, nonetheless revealed meaningful patterns of interaction with AI. These were *greeting/politeness* (phatic communication such as greetings or expressions of gratitude), *translation* (requests primarily for linguistic support), and *off-task activities* (playful prompts, e.g., asking for jokes or pop-cultural stories, playing Battleship, writing poems, or posing absurd tasks such as "2 + 2 = 5"[1]).

In the present study, the five main strategies are applied deductively to an independent sample in order to analyze their occurrence as well as the cross-linguistic transfer of AI writing strategies

[1] Similar patterns of playful prompting have been observed among Grade 4 elementary school students (Kutzner & Schindler, 2026).

between L1/L2 (German) and FL (English). In doing so, the study not only extends a previously developed typology of AI-assisted writing, but also provides an initial, careful validation of this typology by demonstrating its applicability to a new, independent dataset (albeit still rather small at N = 20). Whereas the original study combined multiple data sources to contextualize these strategies (student–AI chat logs, questionnaire responses [Likert-scale and open-ended responses], and characteristics of the corresponding student texts; see Kutzner, in press), the present analysis focuses exclusively on strategy use in relation to learners' reading proficiency. This allows for a more transparent and analytically focused examination of how learners employ AI writing strategies across languages. Selected student texts were additionally consulted to illustrate how these strategies are reflected in actual written products.

The supplementary categories identified in the earlier study were adapted for the present dataset. Because the current data did not include *off-task activities*, this category was dropped. Instead, several learners directly addressed the AI in a derogatory, confrontational, or frustrated manner, leading to the inductive identification of a distinct *AI rejection* category. To maintain conceptual coherence, the earlier *greeting/politeness* category was reframed as *AI affiliation*, capturing the affiliative dimension of anthropomorphic interaction. Additionally, the former *translation* category was integrated into the main strategy *metacognitively-inquiring*, because translation requests typically represent clarification and knowledge retrieval within the writing context in which AI acts as a reflective partner supporting comprehension and self-regulated writing.

5 Results and interpretation

5.1 AI writing strategies across languages

To gain deeper insights into how learners interacted with AI during the writing process, all prompts were analyzed using qualitative content analysis (Kuckartz & Rädiker, 2023) in MAXQDA. Coding was conducted by the author following an iterative procedure of category development and refinement. To enhance analytical rigor and interpretative reliability, category definitions and selected coding decisions were discussed repeatedly within the research group. Given the exploratory and theory-building character of the study and the limited dataset of the analysis, emphasis was placed on conceptual coherence and transparency rather than statistical intercoder agreement. All AI prompts were successfully categorized into the identified AI writing strategies ranging from text planning and linguistic refinement to metacognitive reflection and social interaction with AI. No interactions fell outside the analytical framework. All prompt examples presented in the following sections represent authentic excerpts from student–AI chat logs. To contextualize these qualitative patterns, strategy use was related to quantitative reading measures from *LGVT* (L1/L2, German) and *Cambridge Preliminary* (FL, English) tests. For each strategy, mean *LGVT* T-scores in reading comprehension, speed, and accuracy, as well as average *Cambridge English Scale Score* were considered. Additionally, the cross-linguistic transfer from L1/L2 (German) and FL (English) of each strategy type was quantified using the Jaccard index, representing the overlap of learners who employed the same strategy in both languages according to the formula:

$$J = \frac{|\, A \cap B \,|}{|\, A \cup B \,|}$$

where A and B represent the sets of *individual learners* who used a given strategy in L1/L2 (German) and FL (English) respectively. The intersection $|A \cap B|$ was determined based on the actual overlap of learners across both language contexts. The resulting Jaccard values therefore reflect empirically observed individual-level overlap from the actual intersection of learners. For full transparency, Appendix B reports all learners alongside their reading test results, their respective strategy use in both languages, as well as the resulting individual Jaccard values. Higher values indicate that a larger proportion of the same learners applied a given AI strategy in both languages, reflecting a higher degree of cross-linguistic transfer and strategic stability, whereas lower values suggest language- or context-specific patterns. This measure thus captures how consistently individual learners drew on similar AI-assisted writing strategies across linguistic contexts. The index ranges from 0 (*no overlap*) to 1 (*complete overlap*). Intermediate values can be interpreted as follows: Values below .3 indicate low transfer (mostly distinct strategy use), values between .3 and .6 indicate moderate transfer (some overlap in strategy use), and values above .6 indicate high transfer (substantial consistency in strategy use across languages). Tables 3 and 4 present the resulting strategies and their respective transfer values. For reasons of space and readability, the tables report aggregated frequencies of prompts and users per strategy. Nonetheless, the Jaccard indices reported in the tables are based on the actual learner-level data rather than on marginal frequencies alone. The underlying learner-level distributions and individual strategy constellations across both language contexts are documented in full detail in Appendix B.

Table 3

Five Main AI Writing Strategies

	Characteristics	Student prompt examples (FL)	Prompts and users		
			L1/L2 (German)	**FL (English)**	**Transfer (Jaccard index)**
Strategy 1: focally-structuring	• AI primarily used for text planning and structuring • AI serves as an organizing tool and a thinking aid	• “Help me plan a summary of the short story ‘The Text’ by Brendan Dunne (2012)” (o5b-gpc) • “Give me a writing plan to write a summary of a short story” (lqm-cml)	11 prompts by 8 users *LGVT*: • RC average ⌀ = 48 • RS average ⌀ = 44 • RA average ⌀ = 60	16 prompts by 13 users *Cambridge Scale Score* average ⌀ = 149 (C–B1)	.62

Strategy 2: holistically-optimizing	• AI used for linguistic fine-tuning: choice of words, style, grammar, sentence structure, and useful phrases • Prompts asking for synonyms and rephrasing • Iterative adjustment of text length and complexity	• "Can you give me useful phrases for the summary" (kzo-wng) • "another woord for but" (sld-dpz)	9 prompts by 7 users *LGVT*: • RC average ⌀ = 45 • RS average ⌀ = 41 • RA average ⌀ = 55	17 prompts by 9 users *Cambridge Scale Score* average ⌀ = 150 (C–B1)	.60
Strategy 3: functionally-pragmatizing	• AI used to increase efficiency: (supposedly) saves time and reduces effort • Focus on "quickly getting results"	• "write me a summary for the story of brendan dunne" (sld-dpz) • "write a summary about 'the text' by brendan dunne" (c6a-vgp)	14 prompts by 7 users *LGVT*: • RC average ⌀ = 48 • RS average ⌀ = 46 • RA average ⌀ = 54	10 prompts by 4 users *Cambridge Scale Score* average ⌀ = 156 (B–B1)	.57
Strategy 4: differentially-limiting	• AI use ends after one or two prompts; the text is largely written independently	• "is this summary good?" (auy-bq3)	7 users *LGVT:* • RC average ⌀ = 41 • RS average ⌀ = 42 • RA average ⌀ = 47	6 users *Cambridge Scale Score* average ⌀ = 138 (fail–A2)	.44
Strategy 5: metacognitively-inquiring	• Clarification and knowledge retrieval within the writing context (including translations) • AI as a reflection partner: assistance with evaluation, obtaining formative feedback	• "What does 'to jeer' mean?" (b5b-il8) • "I have now created a summary of the story, would you be able to read it and point out any mistakes?" (dat-k6n) • "this is my summary can you please mke it better" (kzo-wng) • "how is this?" (mm4-ysy)	44 prompts by 15 users *LGVT*: • RC average ⌀ = 45 • RS average ⌀ = 42 • RA average ⌀ = 53	56 prompts by 15 users *Cambridge Scale Score* average ⌀ = 146 (C–B1)	.58

Table 4

Two Supplementary Categories of AI Interaction

	Characteristics	Student prompt examples (FL)	Prompts and users		
			L1/L2 (German)	FL (English)	Transfer (Jaccard index)
AI affiliation	• Phatic communication in which the AI is greeted, thanked, or addressed in a socially affiliative manner	• "hello" (rv5-gjh) • "Thank you very much. I am grateful for the advice, it is of great use" (dat-k6n)	8 prompts by 6 users *LGVT*: • RC average ⌀ = 46 • RS average ⌀ = 43 • RA average ⌀ = 54	4 prompts by 2 users *Cambridge Scale Score* average ⌀ = 164 (A–B2)	.14
AI rejection	• Derogatory, confrontational, or frustrated remarks directed at the AI.	• "N0? what are you on vro. never trust ai" (fsl-dda) • "youre useless" (iyv-svk)	8 prompts by 1 user *LGVT*: • RC average ⌀ = 60 • RS average ⌀ = 60 • RA average ⌀ = 54	4 prompts by 2 users *Cambridge Scale Score* average ⌀ = 157 (B–B1)	.50

Across both languages, the five main strategies reflect functionally distinct forms of AI engagement throughout the writing process, as shown in Table 3. The most prevalent strategy type was *metacognitively-inquiring* (Strategy 5), employed by the majority of learners in both L1/L2 and FL. This indicates that learners primarily perceived AI as a reflective partner that supports comprehension, clarification, and formative feedback. This is reflected in prompts such as "What does 'to jeer' mean" (user b5b-il8), "I have now created a summary of the story, would you be able to read it and point out any mistakes?" (user dat-k6n), or "this is my summary can you please make it better" (user kzo-wng). This shows that through *metacognitively-inquiring* behavior, students became aware of both linguistic and conceptual knowledge gaps. By deliberately asking AI for clarification or elaboration, they engaged in *noticing processes*, using AI as a trigger for reflection and knowledge expansion. This aligns with Schmidt's (1990) *noticing hypothesis* and suggests that AI can act as a catalyst for awareness and self-regulated learning rather than merely as a feedback provider. The moderate Jaccard index (.58) suggests that such reflective strategies are relatively stable across languages, pointing to their generalizability as higher-order cognitive tools that transcend linguistic boundaries. The *holistically-optimizing* strategies (Strategy 2) showed similarly moderate transfer values (.60) as illustrated by prompts such as "Can you give me usefull [*sic*] phrases for the summary" (user kzo-wng) or "another woord [*sic*] for but" (user sld-dpz) that target lexical and stylistic refinement. The transfer of *focally-structuring* strategies

(Strategy 1) can be regarded as high (.62), as reflected in prompts such as “Help me plan a summary of the short story ‘The Text’ by Brendan Dunne (2012)” (user o5b-gpc) or “Give me a writing plan to write a summary of a short story” (user lqm-cml). Together, these three strategies represent procedural, linguistic, and cognitive scaffolding functions within AI-assisted writing. Their balanced occurrence in both L1/L2 and FL contexts suggests that learners are able to flexibly integrate AI across different stages of the writing process from planning and linguistic refinement to metacognitive reflection. In contrast, *functionally-pragmatizing* strategies (Strategy 3) reflect an instrumental orientation toward efficiency and task completion. Learners employing this approach delegated writing tasks to the AI, aiming to minimize effort or time investment. This orientation becomes visible in prompts such as “write me a summary for the story of brendan dunne” (user sld-dpz) or “write a summary about ‘the text’ by brendan dunne” (user c6a-vgp). The moderate transfer value (.57) indicates that such instrumental use patterns were relatively stable across languages, suggesting that the tendency to exploit AI for efficiency may represent a general behavioral disposition rather than a language-specific adaptation. However, although both metacognitive and pragmatic strategies transfer across languages, they differ fundamentally in depth of engagement: The former involves reflective regulation of the writing process with queries such as “how is this?” (user mm4-ysy), whereas the latter primarily targets output optimization through direct task delegation such as “write me a summary . . .” (user sld-dpz). *Differentially-limiting* strategies (Strategy 4) were associated with the lowest transfer value (.44) and the weakest reading proficiency (CEFR A2 for English, lowest *LGVT* means for German). Learners employing these strategies used AI only sparingly, producing only one or two prompts that were functionally assignable to other strategy types. This pattern is reflected in minimal interaction sequences—for example, in prompts such as “is this summary good?” (user auy-bq3), after writing the summary without any AI support. Learners relied primarily on their own writing, indicating cautious engagement and limited trust in AI output. The combination of limited AI use and low transfer implies that weaker learners may not yet have developed a flexible or adaptive repertoire for integrating (AI) writing strategies into their writing process.

The two supplementary categories, *AI affiliation* and *AI rejection*, shown in Table 4, highlight anthropomorphization (e.g., Chaturvedi et al., 2025; Kim & Im, 2023; Li et al., 2025) as a cross-cutting phenomenon in AI use. In both cases, AI is treated as a social interlocutor mirroring human-like communication patterns rather than as a neutral tool. Expressions of greeting, gratitude, or playfulness address constitute affiliative bonding and phatic communication as reflected in prompts such as “hello” (user rv5-gjh) or “Thank you very much. I am grateful for the advice, it is of great use” (user dat-k6n), whereas confrontational or rejecting remarks similarly presuppose a human-like counterpart capable of intention or failure, for example “youre [*sic*] useless” (user iyv-svk) or “N0? what are you on vro [*sic*]. never trust ai” (user fsl-dda). Although *AI affiliation* was rare and showed little cross-linguistic transfer (.14), it reveals moments of social alignment and gratitude toward AI. Conversely, *AI rejection* reflects explicit distancing from the system, typically arising from frustration or distrust and showing moderate transfer across languages (.50). Both categories illustrate the interpersonal framing of AI interaction: Affiliative and antagonistic responses alike presuppose a degree of social agency and interpersonal relevance for the AI. Such interactions may reduce affective barriers and foster motivation, but also risk diminishing critical distance and reinforcing naïve assumptions about the AI’s epistemic competence. Consequently, learners may perceive AI not merely as a functional tool but as a seemingly authentic interlocutor—a finding of particular relevance for pedagogical approaches to AI-assisted writing.

Overall, the pattern of Jaccard indices across strategies reveals a moderate degree of cross-linguistic transfer, indicating that learners generally employed comparable AI-supported writing strategies in both L1/L2 and FL. However, the variability across strategies reflects differences in functionality and cognitive depth. Strategies involving planning, linguistic optimization, or metacognitive reflection (Strategies 1, 2, and 5) demonstrate consistent use across languages, suggesting that such cognitively integrated forms of AI engagement operate as transferable writing strategies. *Functionally-pragmatizing* strategies (Strategy 3), although comparable in transfer rate, reflect a more instrumental engagement with AI, indicating a general behavioral tendency toward efficiency rather than a language-specific adaptation. By contrast, *differentially-limiting* strategies (Strategy 4) remain the least transferable, pointing to cautious or inconsistent AI use among less proficient learners. This suggests that the transferability of (AI) writing strategies depends less on language or linguistic factors and more on the cognitive complexity and degree of learner autonomy. Strategies grounded in reflective regulation or deliberate process management generalize more easily across languages, whereas those shaped by avoidance or minimal involvement remain more context-bound and situational.

To complement the strategy-level analysis, an additional measure was applied to capture the degree of strategic transfer between L1/L2 and FL at the group level. A *generalized Jaccard similarity coefficient* based on strategy frequency vectors was calculated to quantify the overall structural similarity of strategy use across languages. This aggregate measure compares the distribution of learners across strategies in both languages and reflects the degree to which similar patterns of AI-supported writing behavior emerged at the group level.

The generalized Jaccard index was calculated using the number of learners who employed each strategy in L1/L2 (German) and FL (English). The index is defined as:

$$J_g = \frac{\sum_i min(L1_i, FL_i)}{\sum_i max(L1_i, FL_i)}$$

where $L1_i$ and FL_i denote the number of learners who used strategy i in L1/L2 (German) and FL (English) respectively. Importantly, this measure operates purely on aggregated frequency counts and does not take into account whether the same learners employed a given strategy in both language contexts. Rather than capturing individual-level overlap, it compares the overall distribution of strategy use across languages at the group level. Because learners could employ multiple strategies, the analysis reflects multilabel strategy use rather than mutually exclusive group membership. The distribution of learners across strategies in both language contexts is summarized in Table 5.

Table 5

Number of Learners Employing Each AI Writing Strategy in L1/L2 (German) and FL (English)

Strategy	L1/L2 users	FL users	Min (L1; FL)	Max (L1; FL)
Strategy 1: focally-structuring	8	13	8	13
Strategy 2: holistically-optimizing	7	9	7	9
Strategy 3: functionally-pragmatizing	7	4	4	7
Strategy 4: differentially-limiting	7	6	6	7
Strategy 5: metacognitively-inquiring	15	15	15	15
AI affiliation	6	2	2	6
AI rejection	1	2	1	2
Σ			**43**	**59**

Based on the aggregated distributions presented in Table 5, the calculation yields:

$$J_g = \frac{\sum_i \min(L1_i, FL_i)}{\sum_i \max(L1_i, FL_i)} = \frac{43}{59} = .73$$

This value indicates a high degree of structural similarity in the overall distribution of AI writing strategies across languages. It suggests that, at the group level, comparable numbers of AI writing strategy occurrences are observed in both contexts, with only moderate differences in relative strategy use.

5.2 Cross-linguistic transfer of AI writing strategies across CEFR levels

To further contextualize these quantitative findings on strategic transfer, the distribution of AI writing strategies was examined in relation to learners' CEFR proficiency levels in English (FL). This step allows for a more nuanced understanding of how linguistic competence is associated with the presence of particular strategies across languages. The results of this analysis are later complemented by grouping the individual Jaccard indices by proficiency level, thereby linking the qualitative typology of AI writing strategies with quantitative measures of linguistic proficiency to account for cross-linguistic transfer. When examined across CEFR proficiency levels, distinct patterns emerged in the distribution of AI writing strategies. The values reported in Table 6 represent the number of strategy uses across learners in the respective language context. Given the small subgroup sizes, however, these analyses are strictly descriptive and exploratory and do not allow inferential generalization. Percentages should be interpreted with caution, because small absolute differences correspond to relatively large percentage shifts due to the limited subgroup sizes.

Table 6

Distribution of AI Writing Strategies by English CEFR Proficiency Level and Language Context (L1/L2, FL)

Strategy	B2		B1		A2	
	L1/L2	FL	L1/L2	FL	L1/L2	FL
Strategy 1: focally-structuring	2 (18.18%)	2 (20%)	5 (17.24%)	9 (29.03%)	1 (9.09%)	2 (20%)
Strategy 2: holistically-optimizing	1 (9.09%)	2 (20%)	4 (13.79%)	6 (19.35%)	2 (18.18%)	1 (10%)
Strategy 3: functionally-pragmatizing	1 (9.09%)	1 (10%)	6 (20.69%)	3 (9.68%)	0	0
Strategy 4: differentially-limiting	1 (9.09%)	0	3 (10.34%)	3 (9.68%)	3 (27.27%)	3 (30%)
Strategy 5: metacognitively-inquiring	3 (27.27%)	3 (30%)	8 (27.59%)	8 (25.81%)	4 (36.36%)	4 (40%)
AI affiliation	2 (18.18%)	1 (10%)	3 (10.34%)	1 (3.23%)	1 (9.09%)	0
AI rejection	1 (9.09%)	1 (10%)	0	1 (3.23%)	0	0
Total (strategy uses across learners)	11 (100%)	10 (100%)	29 (100%)	31 (100%)	11 (100%)	10 (100%)

At the B2 level, *metacognitively-inquiring* strategies accounted for 3 out of 11 strategy uses in L1/L2 and 3 out of 10 in FL, making this the most frequently observed strategy in both contexts and suggesting a reflective engagement with AI. *Focally-structuring* strategies accounted for 2 occurrences in both L1/L2 and FL; indicating a balanced integration of planning-related support across languages, whereas *holistically-optimizing* strategies appeared slightly more often in the FL (2 occurrences) than in L1/L2 (1 occurrence). This pattern suggests that strategy use at higher proficiency spans a range of functions with a somewhat stronger reliance on linguistic refinement in the foreign language. Notably, both *AI affiliation* (2 occurrences in L1/L2 and 1 in FL) and *AI rejection* (1 occurrence in both contexts) were observed most frequently at this proficiency level. Whereas affiliative interactions point to moments of social engagement with the AI, the presence of rejection may indicate a more critical stance toward AI output, including the ability to distinguish between helpful and less helpful responses.

At the B1 level, *metacognitively-inquiring* strategies again accounted for the highest number of strategy uses (8 out of 29 in L1/L2 and 8 out of 31 in FL). In the FL, however, *focally-structuring* strategies appeared more frequently (9 occurrences) than in L1/L2 (5 occurrences), suggesting a stronger reliance on AI for structuring support in the foreign language. *Holistically-optimizing* strategies were also observed in multiple instances across both contexts (4 in L1/L2 and 6 in FL), indicating that linguistic refinement plays a relevant role at this proficiency level, particularly in the FL. *Functionally-pragmatizing* strategies, by contrast, were more frequent in L1/L2 (6 occurrences) than in FL (3 occurrences), which may suggest that more instrumental uses of AI are less consistently transferred across languages. Overall, these patterns indicate that at intermediate proficiency levels, learners tend to rely more strongly on AI for text organization and linguistic refinement in the foreign language.

At the A2 level, *metacognitively-inquiring* strategies again showed the highest number of strategy uses (4 out of 11 in L1/L2 and 4 out of 10 in FL). At the same time, a comparatively high number of learners employed *differentially-limiting* strategies (3 occurrences in each context), indicating a tendency toward cautious or minimal engagement with AI support that was more pronounced than at higher proficiency levels. In contrast, strategies associated with planning and linguistic elaboration were less frequent: *Focally-structuring* strategies accounted for only 1 occurrence in L1/L2 and 2 in FL, whereas *holistically-optimizing* strategies appeared in 2 occurrences in L1/L2 and 1 in FL. This pattern suggests a comparatively limited use of planning and refinement-oriented strategies, which may indicate a less developed strategic repertoire for managing the writing process (with AI support). This contrasts with the broader and more flexible distribution of strategies observed among more proficient learners and aligns with previous findings suggesting that lower-proficiency writers tend to employ fewer and less elaborated writing strategies (Harris et al., 2010; Lei, 2016; Manchón & Roca De Larios, 2007; Raoofi et al., 2017; Roca De Larios et al., 2001).

Overall, these observations tentatively suggest an association between linguistic proficiency and the breadth as well as the qualitative nature of AI-supported writing strategies. In addition, the language context appears to influence AI use. In the FL, a slightly higher number of learners rely on structural and linguistic scaffolding, whereas in L1/L2 contexts, reflective and evaluative uses of AI are more widely represented.

The observed distributional patterns across proficiency levels suggest potential differences in how learners engage with AI during writing. However, although these frequency-based results capture which strategies are most characteristic for each group, they do not yet account for the stability of individual strategy use across linguistic contexts. To complement the group-level findings and to obtain a more fine-grained perspective, individual Jaccard indices were calculated for each learner to assess intra-individual transfer of AI writing strategies between L1/L2 (German) and FL (English). The index is defined as:

$$J = \frac{|A \cap B|}{|A \cup B|}$$

where A and B represent the sets of strategies used by a single learner in each language. This measure quantifies how consistently individual learners employed similar strategic patterns across linguistic contexts, thereby reflecting the stability or variability of their AI-assisted writing behavior. A complete overview of all individual Jaccard indices for each user is provided in Appendix B. Across all learners, the mean Jaccard index was .57 (see Table 7), indicating a moderate overall degree of strategic transfer between L1/L2 and FL. To further refine this analysis, the individual Jaccard indices were grouped by English CEFR proficiency levels (see Table 7). Linking these transfer measures with proficiency levels allows for a more differentiated perspective on how linguistic competence and language background relate to the stability, variability, and transferability of AI writing strategy use.

Table 7

Descriptive Statistics for the Transfer of AI Writing Strategies (Jaccard Indices)

	Overall Jaccard indices	B2 Jaccard indices	B1 Jaccard indices	A2 Jaccard indices
M	.57	.67	.54	.55
Mdn	.50	.59	.50	.50
SD	.30	.24	.27	.45
Range	1.00	.50	.75	1.00
Min	.00	.50	.25	.00
Max	1.00	1.00	1.00	1.00

Learners at the B2 level exhibited the highest mean transfer (M = .67), indicating a greater degree of consistency in their strategic engagement with AI across languages. In contrast, learners at the B1 (M = .54) and A2 (M = .55) levels demonstrated moderate transfer, suggesting more situational and language-specific variability in their strategy use.

Taken together, the distributional patterns of strategy occurrences, the group- and individual-level transfer measures, and the comparison of individual Jaccard indices across CEFR proficiency levels suggest that cross-linguistic consistency in AI-assisted writing behavior appears more pronounced when strategies involve planning, linguistic refinement, or metacognitive reflection (i.e., focally-structuring, holistically-optimizing, and metacognitively-inquiring strategies) rather than efficiency-driven or cautious modes of engagement (i.e., functionally-pragmatizing and differentially-limiting strategies). Within this sample, learners with higher reading competence tended to transfer such strategies more consistently and flexibly across languages, which may reflect that proficiency enables more deliberate and self-regulated (Harris et al., 2010; Harris et al., 2011) forms of AI use in writing. In contrast, lower-proficiency learners showed a comparatively greater degree of variability and situational adaptation, relying on narrower, more context-dependent patterns of interaction. Overall, the findings point to a possible relation between linguistic proficiency and strategic stability, suggesting that linguistic proficiency may be linked not only to the qualitative nature of the strategies employed but also to the stability and flexibility of their cross-linguistic application. Nevertheless, given the exploratory nature and limited sample size of the study, these patterns should be interpreted with caution.

5.3 Exemplary learner cases of AI writing strategy use

To complement the quantitative and strategy-based analyses, two exemplary student cases are presented to illustrate how AI-supported writing strategies manifest in actual written products across languages. The selected cases are intended to illustrate contrasting patterns of cross-linguistic transfer and proficiency levels and should be interpreted as exploratory, nonrepresentative examples.

An illustrative example of a high-proficiency learner is provided by Student mm4-ysy, whose texts and interaction patterns reflect a consistent and strategically stable use of AI across languages. The learner achieved a Cambridge English Scale Score of 160 (Grade A, CEFR B2), along with LGVT *T* scores in the lower average range for reading comprehension (RC = 46) and reading speed (RS = 41) and in the upper average range for reading accuracy (RA = 56). Notably, the learner shows complete cross-linguistic transfer of AI writing strategies (Jaccard index = 1.0), indicating identical strategy use in L1/L2 and FL contexts.

In the German (L1/L2) task, the learner employed *holistically-optimizing* and *metacognitively-inquiring* strategies throughout the writing process. The interaction with AI is characterized by iterative refinement, including requests for stylistic improvement such as "*nenne mir textbausteine um eine inhaltsangabe für eine kurzgeschichte zu schreiben*" [*sic*] [Tell me useful phrases to write a summary of a short story] and evaluative feedback on draft versions "*Wie ist das bis jetzt?*" [How is this so far?]. The resulting German summary demonstrates a clear structure and coherent progression of ideas, with a concise presentation of the dystopian scenario and its central conflict (see Figure 1).

Figure 1

German Summary by Student mm4-ysy

Kurzgeschichte von mm4-ysy

Die Kurzgeschichte „USE 2030" von Bob Blume, veröffentlicht im Jahr 2025, thematisiert die Konsequenzen politischer Korruption in einer dystopischen Zukunft, in der Elon Musk die Kontrolle über die USA übernommen hat.

Zu Beginn der Geschichte wird der Ich-Erzähler vorgestellt, der aufgrund eines sogenannten „Visual Echo" gezwungen ist, mit Kuli auf Papier zu schreiben, da er nicht freisprechen kann. Er schildert die Ereignisse der letzten Jahre: Der damalige Präsident wurde vermutlich mithilfe von Ketamin vergiftet, was Elon Musk ermöglichte, die Macht zu übernehmen. Die USA werden nun als USE – Vereinigte Staaten von Elon Musk bezeichnet.

In dieser neuen Ordnung gibt es keine Freiheit mehr. Die Menschen werden überwacht, unter anderem durch Jeeps mit schwarzen Flaggen, die ein weißes „X" tragen, und dazu ermahnt, in ihren Unterkünften zu bleiben. Alles, was als „unamerikanisch" gilt, ist streng verboten, um angeblich die Unabhängigkeit des Landes zu sichern. Der Erzähler beschreibt, wie die Abschiebung stark zugenommen hat und dass „regierungsfeindlichen" Personen ihre Staatsbürgerschaft verlieren.

Er selbst besitzt ein paar europäische Konservendosen, die er wie Schätze hütet, und plant, seinen Brief zu verstecken, um späteren Generationen zu zeigen, wie es wirklich war. Die Spannung steigt, als der Erzähler plötzlich Schritte im Treppenhaus hört.

Die Geschichte endet damit, dass der Erzähler kurz davor ist, wegen den Visual Echo, dass ihn „erwischt" hat, entdeckt zu werden. Er erinnert die Leser daran, dass es einst eine Zeit gab, in der die Menschheit frei war.

A similar pattern can be observed in the English (FL) task. Here, the learner again relied on *holistically-optimizing* and *metacognitively-inquiring* strategies, using AI to refine linguistic expression and to evaluate the developing text. Examples include "tell me useful phrases to write a summary of a short story" or "how is this?" The final English summary is likewise well-structured, coherent, and linguistically precise, additionally including an interpretative element that connects the narrative to the fable *The Boy Who Cried Wolf* (see Figure 2).

Figure 2

English Summary by Student mm4-ysy

Summary mm4-ysy

The short story *"The Text"* by Brendan Dunne, published in 2012, follows Amy, a girl with asthma who develops feelings for a boy named Kamal. However, when Kamal takes one joke too far, no one believes him in a moment of real danger, which tragically leads to his death.

The short story follows Amy, a girl with asthma who isn't good with technology, and on that day she somehow forgot her inhaler, her best friend Tara, Kamal and his best friend Grant whom neither of the girls really like

Kamal invites Amy and Tara to visit an abandoned house behind the factory to see his art project. However, Kamal and Grant play a prank by pretending Grant is a zombie, which shocks Amy and triggers her asthma. Although Amy brushes it off as paint fumes, the girls are upset and lose trust in the boys.

The following evening at the school disco, Kamal is nowhere to be found. Grant searches for him, but the girls dismiss him. When Tara receives a text from Kamal saying he is trapped in the abandoned house and needs help urgently, she assumes it is another prank and tells others to ignore it. The next morning, the news reports that Kamal had been seriously injured and died after being trapped under rubble.

Overall, the short story emphasizes the dangers of losing trust, much like in the classic fable "The Boy Who Cried Wolf" once trust is broken, even a true cry for help may be ignored, with devastating consequences.

Across both language contexts, the learner demonstrates a highly consistent approach to AI-supported writing. The learner applies the same combination of refinement- and feedback-oriented strategies in both German and English while maintaining primary authorship of the texts. AI is used selectively as a supportive scaffold rather than as a substitute for text production. This stable pattern is reflected in the maximal Jaccard value (1.0) and suggests that the learner draws on a transferable repertoire of AI writing strategies. The case thus exemplifies how higher-proficiency learners may engage with AI as a tool for iterative optimization and reflective evaluation across languages, supporting the broader finding that strategic stability increases with linguistic proficiency.

In contrast, Student auy-bq3 represents a lower-proficiency learner whose AI use appears more selective and less strategically integrated across languages. The learner achieved a Cambridge English Scale Score of 134 (Fail, CEFR A2), alongside LGVT *T* scores in the lower average range for reading comprehension (RC = 42), below average for reading speed (RS = 39), and in the upper average range for reading accuracy (RA = 50). The learner's low cross-linguistic transfer of AI writing strategies (Jaccard index = .25) indicates only limited overlap in strategy use between L1/L2 and FL contexts.

In the German (L1/L2) task, the learner employed *holistically-optimizing* (*anderes wort für repressiven* [another word for repressive]) and *metacognitively-inquiring* (*ist das gut für ein schlussteil* [sic] [Is that a good ending?]) strategies alongside *AI affiliation* (*DANKE* [thank you]), engaging in multiple interactions with the AI to refine wording, structure the text, and obtain evaluative feedback. The resulting German summary is coherent and well-structured, clearly presenting the central elements of the story. The learner effectively uses AI as a supportive scaffold for refining wording and structuring the text, resulting in a linguistically clear and organized summary (see Figure 3).

Figure 3

German Summary by Student auy-bq3

Kurzgeschichte von Bob Blume Inhaltsangabe

Der Ich-Erzähler sitzt in seiner Wohnung in Ohio am 23. Februar 2030 und beginnt, all seine Beobachtungen und Gedanken mit Papier und Stift zu schreiben – möglicherweise sein letzter Versuch, Zeugnis abzulegen.

Im Jahr 2030 lebt der Ich-Erzähler in einem totalitären Staat, den „United States of Elon Musk". Die Demokratie wurde abgeschafft, und die Regierung kontrolliert das Volk durch Überwachung, Ausgangssperren und Unterdrückung. Öffentliche Orte verfallen, viele Produkte fehlen, und die Bevölkerung lebt in Angst. Der Erzähler berichtet von der Abschiebung seiner Nachbarin und sieht sich selbst bedroht. Trotz dieser Umstände versucht er, seine Gedanken schriftlich festzuhalten, um die Geschehnisse für die Zukunft zu dokumentieren.

Am Ende der Geschichte hört der Erzähler Schritte im Treppenhaus und fürchtet, entdeckt zu werden. Obwohl er nicht sicher ist, ob jemand zu ihm kommt, bleibt er entschlossen, mit seinem Text ein Zeichen zu setzen. Er möchte die Erinnerung an eine Zeit bewahren, in der Freiheit noch ein selbstverständliches Gut war.

A different pattern emerges in the English (FL) task. Here, AI use is markedly reduced and limited to only one prompt, focused on general feedback ("is this summary good?"), without extended revision or optimization cycles. This corresponds to a *differentially-limiting* strategy in which AI interaction remains minimal and does not substantially shape the writing process. The resulting English summary, while covering the main plot, is comparatively verbose, less condensed, and shows structural inconsistencies (see Figure 4).

Figure 4

English Summary by Student auy-bq3

The Text

The Message, titled *The Text* by Brendan Dunne, tells the story of two boys who enjoy playing jokes but take them too far. In the end, no one believes them anymore. There is a girl named Amy who really likes a boy named Kamal. Kamal invites Amy and her friend Tara to come with him so he can show them some graffiti he made for school. However, the girls don't know that Kamal only wants to prank them with his friend Grant. When they arrive at an abandoned house, Kamal tells them that Grant will come later because he has to buy some things. After a few minutes, Amy starts hearing noises like groaning and moaning, similar to a zombie. Kamal insists it's nothing and that he doesn't hear anything. Suddenly, a bloodstained zombie appears. Amy thinks it's a real zombie, which causes her breathing to worsen due to her asthma. Meanwhile, Tara knows it is actually Grant in disguise and notices Amy's breathing difficulties. The next evening was Halloween, but Tara was still worried about Amy because of what happened the day before. However, Amy insisted she was okay and that it wasn't serious. She stopped caring about Kamal and just enjoyed dancing with Tara. Suddenly, Grant approached them looking worried. He said they should text Kamal because his phone had been stolen. The girls didn't take it seriously and thought it was just another one of Kamal's jokes. But then, Tara's phone beeped with a message from Kamal saying he was trapped and asking her to call an ambulance. Despite this, Tara didn't believe him and told everyone to ignore the message. The next day, Amy's parents were listening to the news when they heard a report about a sixteen-year-old boy who had died. The news asked if anyone had any information about the incident. That's when Amy's dad remembered that Amy had gone to that abandoned house with her friends on Monday. The final news was that the boy who died had lost a lot of blood and was buried under a pile of rubble. He was identified as Kamal Naseer. This story teaches an important lesson about honesty and trust. Because the boys took their jokes too far, no one believed Kamal when he really needed help, which led to a tragic outcome.

This suggests that the learner relies predominantly on their own writing with limited integration of AI as a scaffold for revision or improvement.

Across both language contexts, the learner demonstrates a less stable and only weakly transferable pattern of AI-supported writing behavior. Whereas AI is used more actively in the German task, its role in the English task remains limited and largely peripheral. The low Jaccard value reflects this discrepancy and indicates that AI writing strategies are not applied consistently across languages. A tentative explanation for this pattern may lie in differences in linguistic competence: The learner appears more confident and engaged in the German (L1/L2) context in which the proficiency is in the average range and AI is integrated more frequently and productively. In contrast, lower proficiency in the English (FL) context (Cambridge: fail; CEFR A2) may lead to more cautious and reduced AI use. This case thus exemplifies how lower-proficiency learners may engage with AI more cautiously and inconsistently, with limited integration into the writing process and reduced cross-linguistic transfer of AI writing strategies.

Taken together, the two cases illustrate contrasting patterns of AI-supported writing across proficiency levels and language contexts. Whereas the high-proficiency learner (mm4-ysy) demonstrates consistent and transferable use of AI as a scaffold for refinement and reflection in both languages, the lower-proficiency learner (auy-bq3) shows a more fragmented and context-dependent pattern with uneven integration of AI across L1/L2 and FL. This contrast tentatively suggests that cross-linguistic transfer of AI writing strategies may be associated not only with strategic awareness but also with linguistic competence; and this, in turn, may enable learners to engage with AI more flexibly, consistently, and productively across languages. However, given the illustrative nature of these two cases, such interpretations remain exploratory and require further validation in larger samples.

6 Discussion and outlook

The findings of this study provide new insights into how learners engage with AI in the writing process and how these patterns vary across linguistic and proficiency levels. As an exploratory study based on a small-scale classroom sample, the aim was to develop an initial, data-driven understanding of AI-supported writing behavior. The results should therefore be interpreted as preliminary indications rather than generalizable claims, and they warrant further validation in larger and more diverse cohorts.

The qualitative analysis revealed that learners employed a diverse repertoire of AI writing strategies ranging from *focally-structuring* and *holistically-optimizing* approaches to *metacognitively-inquiring* uses in which AI served as a reflective and dialogic partner. Importantly, these strategies were not fixed user types, but rather flexible, phase-dependent writing behaviors, thereby aligning with process-oriented models of writing (e.g., Hayes, 2012). Notably, this process-related character was not imposed by the analytical framework but emerged inductively from the interaction data. Learners moved dynamically between planning, drafting, and revising, using AI as both a cognitive and linguistic scaffold.

The quantitative findings complement this perspective by indicating that such strategic patterns are not confined to a single language context. A high generalized Jaccard similarity coefficient (J_g = .73) points to a similarity in the distribution of learners across strategy types in L1/L2 and FL, suggesting that comparable strategic repertoires emerge at the group level. However, this aggregate measure does not capture individual-level transfer, so this was examined separately using individual Jaccard indices. On the individual level, especially the cognitively demanding strategies for planning, drafting, and revising (i.e., *focally-structuring*, *holistically-optimizing*, and *metacognitively-inquiring* strategies) appear to be more stable and transferable across languages than *pragmatic* or *limiting* strategies. This is reflected in their comparatively higher Jaccard values (.62, .60, and .58 respectively vs. .57 and .44 respectively), suggesting patterns consistent with partial cross-linguistic transfer in AI-supported writing within this exploratory sample. At the same time, meaningful variation between learners was revealed: The language context appears to shape AI engagement, with learners relying more on structural and linguistic scaffolding in FL writing, whereas reflective and evaluative uses prevail in L1/L2 tasks. These differences also align with proficiency levels, as individual Jaccard indices reveal: Students with higher English proficiency (B2) demonstrate greater strategic stability and transferability (M = .67), whereas those at B1 (M = .54) or A2 levels (M = .55) exhibit more situational and context-dependent use. This observation is consistent with prior research on transfer and language awareness (Cummins, 1980, 2000; García & Wei, 2015; Herdina & Jessner, 2002; Odlin, 1989), and may indicate that higher linguistic proficiency is associated with more strategic stability across contexts. However, the observed associations between proficiency and strategic stability should be interpreted with caution, because causal directionality cannot be established, and alternative explanations such as motivational differences or prior AI experience remain possible.

This pattern is further illustrated by the two exemplary cases presented in Section 5.3. The high-proficiency learner demonstrates consistent and transferable use of refinement- and feedback-oriented strategies across both languages, using AI as a scaffold for iterative optimization while maintaining primary authorship. In contrast, the lower-proficiency learner shows a more fragmented and context-dependent pattern, with more frequent and productive AI integration in the L1/L2 context but minimal and peripheral use in the FL. This contrast highlights how differences in linguistic competence may shape not only the extent of AI use but also its functional integration into the writing process, supporting the observed variation in cross-linguistic transfer. Because these cases are illustrative, however, they serve to contextualize rather than generalize the broader quantitative findings.

Interestingly, the identification of *AI affiliation* and *AI rejection* as supplementary categories highlights the anthropomorphic dimension of human–AI interaction. In this dataset, *AI affiliation* involved greeting, thanking, or politely addressing the system, thereby positioning it as a socially responsive interaction partner. Conversely, *AI rejection* did not merely involve refraining from AI use but included direct, confrontational, or evaluative remarks addressed to the system (e.g., insulting or blaming statements). In both cases, the AI was treated as a dialogic counterpart, suggesting that learners attributed a degree of social presence or agency to the AI rather than treating it as a purely instrumental and neutral tool. Although such social framing may lower affective barriers and enhance engagement and motivation, it may also diminish critical distance and foster epistemic naivety toward AI-generated content (see also Dang & Nguyen, 2025; Yan et al., 2025; Zhai et al., 2024). This dual role of AI as both a cognitive and socioaffective agent in the learning process underscores the need for explicit reflection and the fostering of critical AI literacy in educational contexts.

From a pedagogical perspective, these findings indicate that AI interaction may support metalinguistic awareness and self-regulated learning processes, suggesting its pedagogical potential for supporting language acquisition within task-supported and scaffolded learning environments. Furthermore, they point to several implications: First, AI-assisted writing instruction should focus not only on tool competence but also on strategic awareness: helping learners to consciously select, adapt, and evaluate AI use according to the task requirements and writing phase. Second, teachers can scaffold this process through promoting metacognitive awareness and explicit reflection on when and how AI is helpful, and this may strengthen learners' autonomy and critical engagement. Third, because higher proficiency appears to be associated with greater strategic transfer within this sample, targeted support for lower-proficiency learners may be necessary—for example, through modeling effective prompts, explicit instruction on evaluating AI outputs, and contrastive activities across L1/L2 and FL contexts.

Despite these contributions, the study has several limitations: The small sample size and classroom-based context limit the generalizability of the findings. Future research should therefore aim to validate the identified strategy types and transfer patterns in larger and more diverse samples in order to test their robustness and generalizability across different learner populations. Such studies could examine whether the observed relationships between strategy use, cross-linguistic transfer, and proficiency levels hold under varying educational conditions. Moreover, the focus on a single writing genre (summary writing) and a specific AI tool may have influenced the observed strategy patterns. Future research should therefore also explore how AI writing strategies can vary across genres, modalities (e.g., multimodal or argumentative writing), and AI systems with different affordances. Longitudinal designs could further illuminate how learners' AI literacy (Ng et al., 2021) and strategy repertoires evolve over time and whether cross-linguistic transfer increases with continued AI use.

In sum, the integration of qualitative and quantitative analyses suggests that AI-supported writing may function as a site of both linguistic transfer and metacognitive development. The findings underscore the importance of positioning AI not as a substitute for writing competence but as a catalyst for reflective engagement with language. In this sense, AI can serve as a "mirror" for learners' linguistic and metacognitive development, provided that its use is embedded in pedagogical frameworks that promote autonomy, critical awareness, and linguistic transfer. The challenge for future pedagogy therefore lies not merely in teaching learners how to use AI, but in fostering the reflective capacity to decide when, why, and to what extent AI should be used in the writing process.

References

Alghasab, M. B. (2025). English as a foreign language (EFL) secondary school students' use of artificial intelligence (AI) tools for developing writing skills: Unveiling practices and perceptions. *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2025.2505304

Allen, L. K., & Kendeou, P. (2023). ED-AI Lit: An interdisciplinary framework for AI literacy in education. *Policy Insights from the Behavioral and Brain Sciences, 11*(1), 3–10. https://doi.org/10.1177/23727322231220339

Bachmann, T., & Becker-Mrotzek, M. (2010). Schreibaufgaben situieren und profilieren. In T. Pohl & T. Steinhoff (Eds.), *Textformen als Lernformen* (pp. 191–209). Gilles und Francke.

Blume, B. (2025). *USE 2030*. https://bobblume.de/2025/02/11/kurzgeschichte-use-2030/

Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics, 6*(1), 60–70. https://doi.org/10.1093/applin/6.1.60

Bodora, A. L., Decker, L., Fuhlrott, M., Nolden, A., & Steinhoff, T. (2025). Wie schreiben Schüler:innen in der 8. Klasse mit ChatGPT? Einblicke in das Design-Based-Research-Projekt KI-Schreibarrangements. *Leseräume, 12*(11), 1–4.

Cambridge English (2018). *B1 Preliminary for Schools. Preliminary English Test: Reading Sample Test.* https://www.cambridgeenglish.org/exams-and-tests/preliminary-for-schools/preparation/

Cambridge English (2025). *B1 Preliminary for Schools results*. https://www.cambridgeenglish.org/exams-and-tests/preliminary-for-schools/results/

Castro, V., Nascimento, A. K. D. O., Zheldibayeva, R., Searsmith, D., Saini, A., & Kalantzis, M. (2025). Generative AI in K-12 education: The CyberScholar initiative. *Computers and Society.* https://doi.org/10.48550/arXiv.2502.19422

Chan, C. K. Y., & Lee, K. K. W. (2025). The balancing act between AI and authenticity in assessment: A case study of secondary school students' use of genAI in reflective writing. *Computers & Education, 238*. https://doi.org/10.1016/j.compedu.2025.105399

Chaturvedi, R., Verma, S., Srivastava, V., & Khot, S. S. (2025). Exploring the frontier of anthropomorphism in AI agents: Trends and way forward. *Business and Society Review, 130*(1), 42–80. https://doi.org/10.1111/basr.70002

Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S., & Li, M. (2024). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research, 61*(8). https://doi.org/10.1177/07356331231189294

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism 9*, 1–43.

Cummins, J. (1980). The construct of language proficiency in bilingual education. *Georgetown University Round Table on Languages and Linguistics*, 81–103.

Cummins, J. (2000). Language, power, and pedagogy: Bilingual children in the crossfire. *Bilingual Research Journal*, *25*(3), 405–412.

Dang, C. T., & Nguyen, A. (2025). Distinguishing fact from fiction: Student traits, attitudes, and AI hallucination detection in business school assessment. *arXiv:2506.00050*. https://doi.org/10.48550/arXiv.2506.00050

Dunne, B. (2012). *The text (B1)*. https://learnenglishteens.britishcouncil.org/study-break/reading-zone/b1-graded-reading/text-b1

Ellis, R. (2009). Task-based language learning and teaching. Oxford UP.

Fürstenberg, M. (2025). Zur Qualität KI-generierten Feedbacks. Ein explorativer Vergleich menschlicher und künstlicher Intelligenzen. In H. G. Müller & M. Fürstenberg (Eds.), *DeutschGPT 2.0.Deutschunterricht im Dialog mit Künstlicher Intelligenz* (pp. 163–190). Frank & Timme. https://doi.org/10.26530/20.500.12657/104475

Fürstenberg, M., & Müller, H. G. (2024). KI im Deutschunterricht. Funktionsprinzipien und kompetenzbezogene Einsatzmodelle. *Der Deutschunterricht, 5*, 2–13.

García, O., & Wei, L. (2015). Translanguaging, bilingualism, and bilingual education. In W. E. Wright, S. Boun, & O. García (Eds.), *The handbook of bilingual and multilingual education* (pp. 223–240). Wiley-Blackwell. https://doi.org/10.1057/9781137385765

Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. *Journal of Marketing Education, 48*(1), 4–19. https://doi.org/10.1177/02734753241305980

Guariento, W., & Morley, J. (2001). Text and task authenticity in the EFL classroom. *ELT Journal, 55*(4), 347–353. https://doi.org/10.1093/elt/55.4.347

Harris, K., Santangelo, T., & Graham, S. (2010). Metacognition and strategies instruction in writing. In H. S. Waters & W. Schneider (Eds.), *Metacognition, strategy use, and instruction* (pp. 226–256). Guilford.

Harris, K. R, Graham, S., MacArthur, C. A., Reid, R., & Mason, L. H. (2011). Self-regulated learning processes and children's writing. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 187–202). Routledge/Taylor & Francis.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Lawrence Erlbaum.

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism. Perspectives of change in psycholinguistics*. Multilingual Matters. https://doi.org/10.21832/9781853595547

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh UP. https://doi.org/10.1515/9780748626540

Kangasharju, A., Ilomäki, L., Lakkala, M., & Toom, A. (2022). Lower secondary students' poetry writing with the AI-based Poetry Machine. *Computers and Education: Artificial Intelligence, 3*. https://doi.org/10.1016/j.caeai.2022.100048

Kim, J., & Im, I. (2023). Anthropomorphic response: Understanding interactions between humans and artificial intelligence agents. *Computers in Human Behavior, 139*. https://doi.org/10.1016/j.chb.2022.107512

Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies, 30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Kuckartz, U., & Rädiker, S. (2023). Qualitative content analysis: Methods, practice and software (2nd ed.). SAGE.

Kutzner, A. (in press). KI-Schreibstrategien: Empirische Einblicke in das Schreiben mit Künstlicher Intelligenz. *ide – informationen zur deutschdidaktik, 4*.

Kutzner, A. & Schindler, K. (2026). Writing a fairy tale with a little help of ChatGPT – Prompting experiences of fourth-graders. *Journal für Medienlinguistik (jfml), 8*(2): Talking AI. Linguistic Approaches to Artificial Intelligence, 36–72. https://doi.org/10.21248/jfml.2026.73

Lei, X. (2016). Understanding writing strategy use from a sociocultural perspective: The case of skilled and less-skilled writers. *System, 60*(1), 105–116. https://doi.org/10.1016/j.system.2016.06.006

Li, S. (2025). Generative AI and second language writing. *Digital Studies in Language and Literature, 2*(1), 122–152. https://doi.org/10.1515/dsll-2025-0007

Li, T., Wang, M., & Wang, F. (2025). Anthropomorphism of artificial intelligence service agent and consumer responses: A systematic literature review and future research agenda. *International Journal of Consumer Studies, 49*(3), 1–51. https://doi.org/10.1111/ijcs.70066

Manchón, R. M., & Roca De Larios, J. (2007). On the temporal nature of planning in L1 and L2 composing. *Language Learning, 57*(4), 549–593. https://doi.org/10.1111/j.1467-9922.2007.00428.x

Ministerium für Schule und Bildung (2019). *Verwaltungsvorschriften zur Verordnung über die Ausbildung und die Abschlussprüfungen in der Sekundarstufe I (VVzAPO-S I).* https://bass.schule.nrw/Service/12691.htm

Müller-Hartmann, A., & Schocker-von Ditfurth, M. (2011). *Task-supported language learning*. Schöningh. https://doi.org/10.36198/9783838533360

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*, 1–11. https://doi.org/10.1016/j.caeai.2021.100041

Odlin, T. (1989). *Language transfer. Cross-linguistic influence in language learning*. Cambridge UP. https://doi.org/10.1017/CBO9781139524537

Pitzen, L. (2025). Peer- vs. KI-Feedback. Wie Schüler*innen der Sekundarstufe I Textfeedback zu einer Zusammenfassung rezipieren. In K. Schindler, N. Wilder, & Virtuelles Kompetenzzentrumfür Künstliche Intelligenz und Wissenschaftliches Arbeiten (VK:KIWA) (Eds.), *Kleine Kieler Beiträge zu Künstlicher Intelligenz. Kiel: Virtuelles Kompetenzzentrum für Künstliche Intelligenz und Wissenschaftliches Arbeiten (VK:KIWA)*. https://doi.org/10.38071/2025-01281-7

Putri, B. G., Suryati, N., Ivone, F. M., & Hayati, N. (2025). The role of AI in academic writing: Unveiling strategies among EFL students. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 12*(2), 774–792. https://doi.org/10.22219/celtic.v12i2.42039

Raoofi, S., Binandeh, M., & Rahmani, S. (2017). An investigation into writing strategies and writing proficiency of university students. *Journal of Language Teaching and Research 8*(1), 191–198. http://dx.doi.org/10.17507/jltr.0801.24

Rezat, S., & Schindler, K. (2025). Textgenerierende KI im Deutschunterricht: Texte lesen und schreiben mit KI. *Praxis Deutsch, 311*, 4–13.

Roca De Larios, J., Marin, J., & Murphy, L. (2001). A temporal analysis of formulation processes in L1 and L2 writing. *Language Learning, 51*(3), 497–538. https://doi.org/10.1111/0023-8333.00163

Schneider, W., Schlagmüller, M., & Ennemoser, M. (2017). *LGVT 5–12+. Lesegeschwindigkeits- und -verständnistest für die Klassen 5–12* (2nd ed.). Hogrefe.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129–158. https://doi.org/10.1093/applin/11.2.129

Shen, X., & Teng, M.F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity, 52*. https://doi.org/10.1016/j.tsc.2024.101524

Siregar, A. R. F., & Hz, B. I. R. (2025). Undergraduate students’ perceptions on utilizing Deepseek AI as a tool in writing journal articles. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 12*(2), 713–729. https://doi.org/10.22219/celtic.v12i2.41375

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1260843

Steinhoff, T., & Lehnen, K. (2025). Schreiben mit Künstlicher Intelligenz: Das GPT-Modell (Ghost, Partner, Tutor). *Leseräume, 11*.

Summer, T. (2012). From method to postmethod. In M. Eisenmann & T. Summer (Eds.), *Basic issues in EFL teaching and learning* (pp. 1–15). Winter.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing, 62*. https://doi.org/10.1016/j.jslw.2023.101071

Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities, 11*(1). https://doi.org/10.1080/23311983.2024.2392388

Xu, T., Huang, F., & Teo, T. (2025). Understanding EFL learners' strategies in AI-assisted English writing: An activity theory perspective. *Porta Linguarum. An International Journal of Foreign Language Teaching and Learning, 8,* 69–87. https://doi.org/10.30827/portalin.viXIII.32849

Yan, L., Pammer-Schindler, V., Mills, C., Nguyen, A., & Gašević, D. (2025). Beyond efficiency: Empirical insights on generative AI's impact on cognition, metacognition and epistemic agency in learning. *British Journal of Educational Technology, 56*(5), 1675–1685. https://doi.org/10.1111/bjet.70000

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(37), Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zheldibayeva, R., Nascimento, A. K. D. O., Castro, V., Kalantzis, M., & Cope, B. (2025). The impact of AI-driven tools on student writing development: A case study. *Online Journal of Communication and Media Technologies, 15*(3), Article e202526. https://doi.org/10.30935/ojcmt/16738

Author's address:

Alyssa Kutzner, University of Wuppertal, Faculty of Humanities and Cultural Studies | German Studies, German Language Education, Chair of Prof. Dr. Kirsten Schindler
kutzner@uni-wuppertal.de

Appendix A

Reading Test Results (Sorted by Performance: Cambridge Scale Score/Grade/CEFR Level and Reading Comprehension)

Student code	*Cambridge Preliminary* Cambridge Scale Score	*Cambridge Preliminary* grade / CEFR level	Reading comprehension (RC) *(Leseverständnis (LV))* *T* score	Reading comprehension (RC) *(Leseverständnis (LV))* Classification	Reading speed (RS) *(Lesegeschwindigkeit (LGS))* *T* score	Reading speed (RS) *(Lesegeschwindigkeit (LGS))* Classification	Reading accuracy (RA) *(Lesegenauigkeit (LGN))* *T* score	Reading accuracy (RA) *(Lesegenauigkeit (LGN))* Classification
dat-k6n	170	A–B2	58	Upper average range	53	Upper average range	63–>73	Above average to very high
fsl-dda	167	A–B2	60	Above average	60	Above average	54	Upper average range
lqm-cml	160	A–B2	48	Lower average range	41	Lower average range	56	Upper average range
mm4-ysy	160	A–B2	46	Lower average range	41	Lower average range	56	Upper average range
1jj-px3	157	B–B1	55	Upper average range	50	Upper average range	59	Upper average range
b5b-il8	157	B–B1	54	Upper average range	48	Lower average range	63–>73	Above average to very high
z9c-nry	157	B–B1	48	Lower average range	41	Lower average range	63–>73	Above average to very high
c6a-vgp	153	B–B1	43	Lower average range	43	Lower average range	46	Lower average range
uqe-0qh	153	B–B1	40	Lower average range	37	Below average	50	Upper average range

8ds-ok0	150	C–B1	47	Lower average range	43	Lower average range	51	Upper average range
o5b-gpc	150	C–B1	42	Lower average range	39	Below average	50	Upper average range
sld-dpz	150	C–B1	41	Lower average range	41	Lower average range	46	Lower average range
iyv-svk	147	C–B1	50	Upper average range	49	Lower average range	52	Upper average range
s9c-vnm	143	C–B1	35	Below average	37	Below average	40	Lower average range
rv5-gjh	143	C–B1	30	Below average	<27	Very low	50	Upper average range
auy-bq3	134	fail–A2	42	Lower average range	39	Below average	50	Upper average range
ovz-kbj	130	fail–A2	31	Below average	34	Below average	37	Below average
ouf-wbf	128	fail–A2	31	Below average	42	Lower average range	33	Below average
qqs-yax	122	fail–A2	42	Lower average range	39	Below average	50	Upper average range
kzo-wng	120	fail–A2	44	Lower average range	39	Below average	56	Upper average range

Appendix B

Overview of Individual Results for Reading Performance, AI Writing Strategies, and Cross-Linguistic Transfer (Jaccard Indices)

(Sorted by Performance: Cambridge Scale Score/CEFR Level/Grade and Reading Comprehension)

<table>
<tr><th rowspan="2">Student code</th><th colspan="2">Cambridge Preliminary</th><th colspan="2">Reading comprehension (RC)
(Leseverständnis (LV))</th><th colspan="2">Reading speed (RS)
(Lesegeschwindigkeit (LGS))</th><th colspan="2">Reading accuracy (RA)
(Lesegenauigkeit (LGN))</th><th colspan="3">AI writing strategies</th></tr>
<tr><th>Cambridge Scale Score</th><th>Grade/ CEFR level</th><th>T score</th><th>Classification</th><th>T score</th><th>Classification</th><th>T score</th><th>Classification</th><th>L1/L2
(German)</th><th>FL
(English)</th><th>Transfer (Jaccard index)</th></tr>
<tr><td>dat-k6n</td><td>170</td><td>A–B2</td><td>58</td><td>Upper average range</td><td>53</td><td>Upper Average Range</td><td>63–>73</td><td>Above average to very high</td><td>• focally-structuring
• AI affiliation
• differentially-limiting</td><td>• focally-structuring
• metacognitively-inquiring
• AI affiliation</td><td>.5</td></tr>
<tr><td>fsl-dda</td><td>167</td><td>A–B2</td><td>60</td><td>Above average</td><td>60</td><td>Above Average</td><td>54</td><td>Upper average range</td><td>• functionally-pragmatizing
• metacognitively-inquiring
• AI affiliation
• AI rejection</td><td>• functionally-pragmatizing
• AI rejection</td><td>.5</td></tr>
</table>

lqm-cml	160	A–B2	48	Lower average range	41	Lower Average Range	56	Upper average range	• focally-structuring • metacognitively-inquiring	• focally-structuring • holistically-optimizing • metacognitively-inquiring	.67
mm4-ysy	160	A–B2	46	Lower average range	41	Lower Average Range	56	Upper average range	• holistically-optimizing • metacognitively-inquiring	• holistically-optimizing • metacognitively-inquiring	1.0
1jj-px3	157	B–B1	55	Upper average range	50	Upper Average Range	59	Upper average range	• holistically-optimizing • functionally-pragmatizing • metacognitively-inquiring	• focally-structuring • holistically-optimizing • metacognitively-inquiring	.5
b5b-il8	157	B–B1	54	Upper average range	48	Lower Average Range	63–>73	Above average to very high	• focally-structuring • metacognitively-inquiring	• focally-structuring • holistically-optimizing • metacognitively-inquiring • AI affiliation	.5
z9c-nry	157	B–B1	48	Lower average range	41	Lower Average Range	63–>73	Above average to very high	• focally-structuring • holistically-optimizing • functionally-pragmatizing	• focally-structuring • holistically-optimizing	.5

									• metacognitively-in-quiring		
c6a-vgp	153	B–B1	43	Lower aver-age range	43	Lower Aver-age Range	46	Lower aver-age range	• functionally-pragma-tizing • AI affiliation • differentially-limiting	• focally-structuring • functionally-pragma-tizing	.25
uqe-0qh	153	B–B1	40	Lower aver-age range	37	Below Aver-age	50	Upper aver-age range	• functionally-pragma-tizing • metacognitively-in-quiring • AI affiliation (1)	• focally-structuring • functionally-pragma-tizing	.25
8ds-ok0	150	C–B1	47	Lower aver-age range	43	Lower Aver-age Range	51	Upper aver-age range	• metacognitively-in-quiring • differentially-limiting	• metacognitively-in-quiring • differentially-limiting	1.0
o5b-gpc	150	C–B1	42	Lower aver-age range	39	Below Aver-age	50	Upper aver-age range	• focally-structuring • holistically-optimizing • metacognitively-in-quiring	• focally-structuring • holistically-optimizing • metacognitively-in-quiring	1.0
sld-dpz	150	C–B1	41	Lower aver-age range	41	Lower Aver-age Range	46	Lower aver-age range	• focally-structuring • functionally-pragma-tizing	• focally-structuring • holistically-optimizing • functionally-pragma-tizing • metacognitively-in-quiring	.5

iyv-svk	147	C–B1	50	Upper average range	49	Lower Average Range	52	Upper average range	• focally-structuring • functionally-pragmatizing • metacognitively-inquiring	• focally-structuring • metacognitively-inquiring • AI rejection	.5
s9c-vnm	143	C–B1	35	Below average	37	Below Average	40	Lower average range	• holistically-optimizing • differentially-limiting	• holistically-optimizing • metacognitively-inquiring • differentially-limiting	.67
rv5-gjh	143	C–B1	30	Below average	<27	Very Low	50	Upper average range	• metacognitively-inquiring • AI affiliation	• focally-structuring • metacognitively-inquiring • differentially-limiting	.25
auy-bq3	134	fail–A2	42	Lower average range	39	Below Average	50	Upper average range	• holistically-optimizing • metacognitively-inquiring • AI affiliation	• metacognitively-inquiring • differentially-limiting	.25
ovz-kbj	130	fail–A2	31	Below average	34	Below Average	37	Below average	• metacognitively-inquiring • differentially-limiting	• differentially-limiting	.5

ouf-wbf	128	fail–A2	31	Below average	42	Lower Average Range	33	Below average	• metacognitively-inquiring • differentially-limiting	• metacognitively-inquiring • differentially-limiting	1.0
qqs-yax	122	fail–A2	42	Lower average range	39	Below Average	50	Upper average range	• differentially-limiting	• focally-structuring • metacognitively-inquiring	0
kzo-wng	120	fail–A2	44	Lower average range	39	Below Average	56	Upper average range	• focally-structuring • holistically-optimizing • metacognitively-inquiring	• focally-structuring • holistically-optimizing • metacognitively-inquiring	1