

Pedagogy First: Collegial Conversations about Large Language Models, Cognitive Debt, and Language Teaching

Trina Philpot-Montaña, EdD

Evaluator, Office of Institutional Effectiveness

The growing emphasis on using generative artificial intelligence (GenAI) at DLIFLC has led to experimentation, discussions, and presentations to develop appropriate practices for adopting this technology for teaching and learning. The 2025 Holiday Professional Development Program (HPDP) was an opportunity for teachers to build our capacity to integrate large language models (LLMs), such as ChatGPT and Copilot, and to discuss other matters of importance to the institute. This faculty forum article is a collegial conversation arising from one HPDP presentation that led to deeper discussions about the problem of minimally guided approaches to integrating GenAI into second language learning (SLL). As we work towards developing guiding principles for the productive use of GenAI, teachers are instrumental in supporting proficiency gains, cognitive development, and intellectual growth. From broad guiding principles, we may eventually generate LLM microstrategies of engagement grounded in a deep understanding of second language pedagogy and macrostrategies for language teaching. The challenge of meeting the changing needs of our students requires collegial dialogue and the continuous collection of empirical and anecdotal evidence to support the next generation of military linguists.

Keywords: Artificial Intelligence, Cognitive Debt, Large Language Models, Experiential Learning, TPACK, AI Literacy, GenAI, Machine Learning

INTRODUCTION

During the Holiday Professional Development Program (HPDP) in December 2025, I attended the presentation “Artificial Intelligence and Cognitive Debt” by Dr. Jawad Khan, Faculty Development Specialist in the Persian Farsi School. The presentation about cognitive debt and large language models (LLMs) led to insightful conversations with Dr. Khan about the implications of using ChatGPT and Copilot for language learning and teaching. As explained in the presentation,

cognitive debt is “a cumulative deficit in critical thinking, ethical reasoning, and professional judgement that arises from offloading cognitive tasks” (Watts, 2025, Abstract). In an article about the cost of LLM-augmented knowledge work (such as machine-generated summaries), Ruttenberg (2026) argued that cognitive debt results from attentional erosion and effort displacement and is more likely to occur when AI use is chronic, unreflective, and rewarded by institutional policies that promote AI use for the sake of efficiency.

Contextualizing AI Within a Broader Social Context

After the HPDP, I embarked on a commitment to read seven books on artificial intelligence (AI) to increase my knowledge of the promises and pitfalls of these tools. These books address critical concerns about generative AI (GenAI) and predictive AI and aim to balance the hype surrounding the tech industry’s claims about the capabilities of “anthropomorphized machines” (Bender & Hanna, 2025). Just as Dr. Khan’s discussion of cognitive debt caused me to think about the consequences of an over reliance on AI in second language learning (SLL), my new book collection (*The AI Con*, *Atlas of AI*, *Unmasking AI*, and *Empire of AI*, among other titles) gave me reasons to develop an informed skepticism regarding predictive and GenAI integration in society. Among other topics, the authors addressed issues related to algorithmic bias, copyright infringement, and intellectual property violations (Bender & Hanna, 2025; Buolamwini, 2023; Crawford, 2021; Hao, 2025). While there is substantial research on these broader social concerns of AI use, there is room for more discussion on the cognitive and ethical implications for using large language models (LLMs) in second language learning (SLL).

Elevating Pedagogy Over Technology

The purpose of this article is to further the conversation on common pitfalls of LLMs, like ChatGPT and Copilot in SLL. By focusing on developing broad guiding principles for LLM adoption in SLL, teachers can feel empowered to innovate microstrategies for using GenAI. I will ground my discussion in a principle I learned from the late Dr. Leo van Lier at the Middlebury Institute of International Studies (MIIS). Pedagogy first. Curriculum second. Technology last. By relating this principle (illustrated in Figure 1) to insights gained from two articles included in Dr. Khan’s presentation: *Paying the Cognitive Debt: An Experiential Learning Framework for Integrating AI in Social Work Education* (Watts, 2025) and *Your Brain on ChatGPT* (Kosmayna et al., 2025), I offer some considerations for teaching with LLMs. I do not propose a panacea of macro or microstrategies for LLMs in SLL, only a thoughtful contribution to the ongoing dialogue about AI in SLL.

Figure 1

Pedagogy First, Curriculum Second, Technology Last

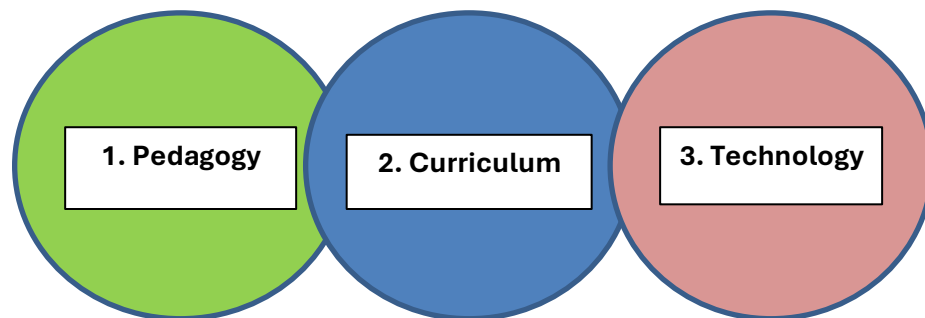


Figure 1 highlights the importance of teaching and learning by placing technology in service to curriculum and pedagogy, not the reverse.

UNDERSTANDING AI CAPABILITIES AND ACKNOWLEDGING PEDAGOGICAL CHALLENGES

Given the versatility and multiplicity of AI, it is essential to define key terms relevant to the current discussion to ensure a shared understanding of specific AI tools. GenAI refers to a type of technology that, via user prompts, can create new content (e.g., images, text, code) by mimicking patterns from large data sets. Under the umbrella of GenAI are LLMs, which are advanced GenAI systems that predict words in a sequence and can perform a range of human-like language tasks; for example, they can interpret text or summarize an article “with varying degrees of autonomy ... [performing] tasks that would otherwise be performed by humans” (Stryker, 2024, para. 3). Examples of LLMs include public interfaces such as Microsoft’s Copilot, Google’s Gemini assistant, or Open AI’s ChatGPT.

Pedagogical problems abound in the age of AI. Perhaps the most pressing challenge is the minimally guided integration of LLMs in higher education. According to the Digital Education Council (2025), 80% of faculty reported a lack of clarity in applying AI to teaching. In addition, many institutions lack clear AI policies (Ng et al., 2025), which has led to a pedagogical vacuum where teachers lack expert guidance or the ability to create structured learning objectives (Batista et al., 2024). Research indicates that a lack of guidance and criticality in AI educational use leads to skill erosion, cognitive atrophy, decreased critical thinking, diminished problem-solving, reduced empathy, and compromised professional judgment (Kosmayna et al., 2025; Perkinson, 2025; Watts, 2025). Moreover, the consequences of AI misuse in SLL represent a critical risk factor for developing competence as a (military) linguist, but risks can be mitigated by developing productive AI engagement strategies to support language proficiency gains, cognitive development, and intellectual growth.

As Dr. Khan emphasized, when teachers model responsible use of AI, the pitfalls of misguided use can be minimized. Teacher leadership in LLM adoption is critical because students tend to choose the path of least cognitive resistance (Kosmyna et al., 2025). For teachers to lead in the age of AI-assisted learning, we need macrostrategies (broad guiding principles) for AI engagement to guide students and ourselves in leveraging LLMs to support learning rather than detract from it. We may also need to consider curricular (or programmatic), modifications to achieve this goal.

PRIORITIZING TARGET LANGUAGE PRODUCTION

Dr. Khan used Kolb's experiential learning theory to propose six AI-assisted activities grounded in concrete experience, reflective observation, abstract conceptualization, and active experimentation (Watts, 2025). He aimed to highlight that machine output has faults and that students must be trained to view synthetic content with skepticism to minimize the risk of drawing incorrect inferences. The suggested activities involved vocabulary building, pronunciation, descriptions, and speech acts. One vocabulary activity involved giving students AI-generated sentences to review and improve spelling, synonyms, and context. The description activity involved students uploading a picture to a GenAI tool and getting a synthetic description of facial features that students would improve and rewrite. In a post-presentation conversation, Dr. Khan and I discussed how these AI-centered activities could be improved to align with the foundational tenets of second language acquisition, such as providing opportunities for students to create original language, to personalize vocabulary and grammar learning, and to negotiate interaction with peers, teachers, and the target language community.

Although Dr. Khan's intent was to provide opportunities for students to recognize flaws in synthetic language, we can broaden the application of cognitive-debt-reducing activities to ensure pedagogical alignment with principles of SLL. One strategy is to instruct learners to perform a proficiency-oriented activity, such as describing a person/image, before prompting the machine to do the task. Subsequently, students can compare their original descriptions with synthetic texts by categorizing comparisons in a multi-column table to analyze linguistic features, such as lexical choice, verb forms, discourse features, and socio-cultural elements. Additionally, as language-specific LLMs are starting to emerge, it will be interesting to see how these regional LLMs made by and for a specific cultural and linguistic audience compare to an English-based LLM when describing a person.

Regardless of the activity, teachers can consider having students contrast their original descriptions with those of classmates and target language (TL) speakers in the broader community to preserve the importance of human-centered, meaningful interactions to describe a person/image. Communicative exchanges with speakers of the TL community to describe physical characteristics of people related to skin color, hair textures, and facial features would typically involve idiosyncratic language use, cultural nuances, and negotiated interaction. Analyzing differences between human-generated and synthetic descriptions promotes cognitive engagement without sacrificing the benefits of creating with the TL. Furthermore, for lower

proficiency learners, the language generated by LLMs is often beyond $i + 1$, meaning that the language is not considered comprehensible input, which is a cornerstone of effective language teaching. However, engaging with peers of similar proficiency levels and experienced teachers creates an environment more conducive to level appropriate language learning and cognitive engagement. Although teachers and learners can adjust prompts to vary the level of complexity of machine output for specific imagined interlocutors, synthetic descriptions lack unique individual experiences and idiosyncratic cultural knowledge that TL speakers and language learners bring to communicative encounters.

Aspiring military linguists need to engage in the creative process of combining newly acquired and recycled vocabulary and painstakingly choosing verb conjugations to convey original meaning rooted in lived experiences. Evaluating synthetic content before creating meaning is a passive exercise. Furthermore, the ability to evaluate mechanical language is limited if learners do not possess the linguistic ability to perform the task unassisted. Pairing Dr. Khan's cognitive-debt-reducing activities with front-loaded student-centered language production gives teachers and students more options for using LLMs effectively. This approach to SLL in the age of AI has the potential to increase language proficiency and learner awareness of the promises and limitations of LLMs.

ALIGNING AI-ASSISTED STRATEGIES WITH LEARNING OBJECTIVES

Although it is imperative to experiment with different strategies for integrating AI into SLL, it is equally important to align strategies with sound pedagogical decisions. For example, AI-assisted (rather than AI-centered) activities can be aligned with learning objectives in the following ways:

- *Learning Objective*: Students will be able to accurately orally describe a person, using basic grammatical structures and vocabulary for clothing, accessories, facial features, and other physical characteristics in culturally appropriate ways at the ILR level 2.
- *Activities*: Working in pairs or small groups, students will orally describe a person and then write down their description. Students will then evaluate descriptions produced by LLMs, by their classmates, and by TL speakers based on lexical choice, verb forms, discourse features, and socio-cultural elements. Students will discuss differences and write an improved description incorporating the best elements of human and machine language. As a culminating activity, students will give an oral description of a different person without any notes. Feedback provided by the teacher should focus on accuracy as described in the learning objective. Teachers should explain the rationale for the activity, emphasizing the benefits of preserving adequate cognitive load throughout the learning process and the risks of over-reliance on AI tools.

As a basic guiding principle for using LLMs in SLL, I offer a five-step model for responsible classroom use. See Table 1.

Table 1

Five-Step Model for Responsible Classroom Use of GenAI Tools

Step	Student/Teacher Action	Purpose
1. Student Performance	Students engage in the cognitive process to produce language (in writing or a transcript of speaking) according to proficiency standards, such as describing a person or thing, narrating in the past, present, or future, reporting facts, and writing original sentences with unit vocabulary words	Students benefit from producing or interpreting language before they use an AI tool
2. Peer Comparison	Students compare their production with what a peer produced. This can be done in pairs or small groups. Teachers should set clear guidelines on what students should compare (e.g., grammatical accuracy, vocabulary, cultural nuances, etc.)	Engaging in peer interaction to negotiate meaning is where language learning occurs. The give and take of human interaction contributes to the development of fluency, grammatical competence, and lexical development as well as fostering trust and confidence building among learning communities.
3. GenAI Tool Comparison	Students work individually or in pairs/small groups to compare their product with what a GenAI tool such as Microsoft Copilot produces. Teachers can either share a GenAI-created product for students to compare, or they can train students on how to write prompts so they can compare products.	Having students compare their product with what a GenAI tool creates provides learners with an alternative source of TL input and develops not only their critical thinking skills but also their language proficiency.
4. Guided Linguistic Analysis	Students go back to their original product and analyze it, focusing on opportunities for revision based on the comparisons they made and feedback and discussions from peers and/or their teacher. Teachers should provide clear guidance on what students should focus on (e.g., linguistic accuracy, cultural nuances, lexical variations).	Having students re-visit, analyze, and revise their product contributes to greater depth in language learning.

5. Revised Performance without Notes	Students deliver a presentation or revise their original product without using notes. In other words, this should be spontaneous verbal or written production.	Providing opportunities for students to produce the target language by gradually reducing the number of scaffolds contributes to improved learning outcomes and helps teachers and students to identify progress made and gaps to be addressed in subsequent lessons.
--------------------------------------	--	---

These five steps represent an approach that preserves cognition, human interaction, critical thinking, and discourse analysis.

PRESERVING NEGOTIATED HUMAN INTERACTION

As we comprehend the idiosyncrasies of LLMs in multiple languages, we can benefit from refreshing our understanding of Kumaravadivelu's (2003) 10 macrostrategies for language teaching: 1) maximizing learning opportunities, 2) minimizing perceptual mismatches, 3) facilitating negotiated interaction, 4) promoting learner autonomy, 5) fostering language awareness, 6) activating internal heuristics, 7) contextualizing linguistic input, 8) integrating language skills, 9) ensuring social relevance, and 10) raising cultural consciousness. Of the 10 macrostrategies discussed in the book *Beyond Methods: Macrostrategies for Language Teaching*, a foundational and seminal text for anti-prescriptive language instruction, I will highlight facilitating negotiated interaction, the third macrostrategy, as being relevant to the example of using GenAI tools for practicing the task of describing a person.

Kumaravadivelu's chapter on facilitating negotiated interaction begins with the following quote from Dick Allwright (an influential applied linguist who co-authored a book with MIIS's former professor Kathleen Bailey) in 1984. "The importance of interaction is not simply that it creates learning opportunities; it is that it constitutes learning itself" (in Kumaravadivelu, 2003). This quote encapsulates the centrality of (human) interaction and (human) negotiation as key to (language) learning. As is well known in SLA and is also common practice at the DLIFLC, learning opportunities are maximized when there are multiple negotiated interaction points (between students, students and teachers, small groups, and with the broader community of target language speakers). To maximize interactional language learning opportunities, LLMs must be relegated to the role of assisting in language learning and teaching rather than being the direct focus of language production and analysis.

Halliday (1985, as cited in Kumaravadivelu, 2003) proposed three types of interaction: textual, interpersonal, and ideational. Textual interaction refers to modifying messages based on the interlocutor's responses and reactions, while interpersonal interaction refers to sociolinguistic aspects such as roles and relationships among speakers. Ideational interaction refers to the ideas and emotions participants bring to a communicative encounter based on lived experiences (past and present), cognitive awareness, and sociocultural sensitivity. In synthetic exchanges with the

machine, true interaction does not take place because the three types of interaction are not fully present. GenAI cannot differentiate between how humans interact in different contexts. Although through crafty prompting, users can simulate interactions from various perspectives and with multiple imagined interlocutors.

EMBEDDING AI LITERACY IN PROGRAM PLANNING

Given students' tendency to use GenAI as a shortcut rather than as a support tool, language learning materials could be expanded to include AI literacy. These digital skills would be grounded in augmenting cognitive, intellectual, and linguistic development and reinforcing academic integrity. According to Watts (2025), academic policies and guidance are central to preserving academic integrity and safeguarding the learning process in the age of GenAI. Suggested policies include requiring students to justify AI use and offering guiding principles for AI adoption in the curriculum or program goals, for example considering any risks related to academic integrity, learning, data privacy, or human interaction. Teaching students to recognize these risks when using AI in various learning scenarios increases their AI literacy and opportunities for learning. In addition to teaching AI literacy, we can lean on evidence-based approaches for using LLMs to support learning in various contexts while also documenting our experiential knowledge. Valuing active experimentation rather than hurried implementation will likely lead to more favorable outcomes. While some teachers may prefer to teach without LLMs, others may be overly enthusiastic about these tools. Perkinson (2025), a critic of LLMs in education, posited that AI rejection is not realistic because students are using AI to complete even the lowest of stakes assignments requiring minimal effort. He demonstrated that teachers could design AI exercises that guide students to uncover limitations of these tools. His AI literacy exercise was successful because it led students to experience the fallacies of machine output and to develop "an informed skepticism with which to confront utopian claims about the AI-driven future" (p. 35).

ELEVATING TEACHER EXPERTISE

With a firm pedagogical foundation and learning materials that include AI literacy, we can focus on broad guiding principles for teachers using LLMs to enhance learning. Watts (2025) discussed effective technology integration based on the Technological Pedagogical Content Knowledge (TPACK) framework, which demonstrates that pedagogical goals and content knowledge must guide the integration of new educational technologies. In other words, teachers must understand how the tools work, how people learn, as well as possess deep content knowledge. As we continue to experiment with LLM integration into SLL, we will understand more about how this technology works and, better yet, how it does not. By researching LLMs in multiple languages and developing experiential knowledge, we will comprehend the possibilities and limitations of AI in intensive foreign language programs. We can also identify pedagogical problems as the starting point for strategizing how to integrate LLMs in SLL. This approach aligns with Buolamwini's (2023) view on a problem-centric approach for using AI to solve issues that affect humanity, rather than creating new issues that arise from the pursuit of efficiency and maximizing tech use for tech's sake.

As we focus on identifying problems of practice, our language-specific and AI knowledge will grow. We will be equipped to guide learners to observe the fallacies of machines and to appreciate human expertise and thought processes. Applying knowledge of discourse analysis, sociolinguistics, critical language awareness, sociopolitical consciousness, and intercultural competence, our role as teachers, linguists, and SLL experts becomes more significant as we guide learners to leverage the possibilities and limitations of LLMs.

Upon reflecting on a DLIFLC teacher's comment about the impending obsolescence of the language teaching profession, I believe our expertise will become more important as we lean into our depth of pedagogical knowledge and scrutinize LLM output in our respective languages. We can ask ourselves if the waves of tech upgrades have significantly improved learning outcomes. Although we cannot think in terms of a binary relationship between technology and pedagogy, we can assert that pedagogy should guide the implementation of LLMs. We may be on the brink of the pendulum oscillating toward innovation of educational structures, pedagogical techniques, and curricular frameworks. This shift could lead to a productive integration of AI in SLL or to restrict its use to certain contexts.

CONCLUSION

May this collegial conversation between Dr. Khan and myself serve as a model for future dialogue focused on pedagogical and curricular innovation as well as critical language awareness in the age of GenAI. I hope this article encourages teachers to nurture that inner voice, warning us to proceed with caution during this pivotal moment of rapid acquisition of LLMs in SLL, and inspires us to continue experimenting, scrutinizing, researching, and generating creative solutions to pedagogical problems. These solutions can align with the following principles: prioritizing the human need to create language, comparing multiple linguistic features of human and synthetic content, aligning AI-assisted strategies with learning objectives, embedding AI literacy in program planning, preserving meaningful human interactions, reducing automation bias, and elevating teacher expertise. By using LLMs to help solve pedagogical challenges in SLL, we can indeed honor the ideology of Pedagogy First. Curriculum Second. Technology Last.

Acknowledgements

I wish to thank Dr. Jawad Khan for his insightful HPDP presentation, which inspired this faculty forum article. Our engaging conversations after his presentation were refreshing and energizing. I am especially grateful for the multiple rounds of feedback he provided during the writing process. I look forward to future collaborations with Dr. Khan. As a MIIS graduate of the class of 2014, I also wish to acknowledge the amazing language professors from the Middlebury Institute of International Studies who helped shape the pedagogical practices of numerous DLIFLC faculty over the years. Lastly, I celebrate every teacher's contribution to thoughtful discussions taking place on language teaching in the age of GenAI.

REFERENCES

- Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, 15(11), 676. <https://doi.org/10.3390/info15110676>
- Bender, E. M., & Hanna, A. (2025). *The AI con: How to fight the big tech's hype and create the future we want*. Harper Collins.
- Buolamwini, J. (2023). *Unmasking AI: My mission to protect what is human in a world of machines*. Random House.
- Crawford, K. (2021). *Atlas of AI. Power, politics, and planetary costs of artificial intelligence*. Yale University Press.
- Digital Education Council. (2025). DEC AI literacy framework. Available online: <https://www.digitaleducationcouncil.com/post/digital-education-council-ai-literacy-framework>
- Hao, K. (2025). *Empire of AI: Dreams and nightmares in Sam Altman's OpenAI*. Penguin Random House.
- Khan, J. (2025, December 23). *AI & Cognitive Debt* [PowerPoint slides]. Presentation delivered at the 2025 Holiday Professional Development Program (HPDP), Defense Language Institute Foreign Language Center.
- Kosmayna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.H., Beresnitzky, A.V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *Cornell University ArXiv*, 1. <https://doi.org/10.48550/arXiv.2506.08872>
- Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. Yale University Press.
- Ng, B. Y., Li, J., Tong, X., Ye, K., Yenne, G., Chandrasekaran, V., & Li, J. (2025). Analyzing security and privacy challenges in generative AI usage guidelines for higher education. *arXiv*. <https://doi.org/10.48550/arXiv.2506.20463>
- Perkinson, S. (2025, Spring). An AI exercise in the classroom. *The Newsletter of the International Center of Medieval Art* [Newsletter]. 29-35. https://static1.squarespace.com/static/53a4b792e4b073bf214c0e66/t/67ddcdb4e1ee531df076cb82/1742589366973/ICMA_MarchNewsletter_v7+FINAL.pdf
- Ruttenberg, D. (2026). Cognitive debt: The cumulative cognitive cost of AI-augmented knowledge work. *Center for Open Science*. https://osf.io/preprints/psyarxiv/yhvec_v1
- Stryker, C. (2024). *What are LLMs?* <https://www.ibm.com/think/topics/large-language-models>
- Watts, K. J. (2025). Paying the cognitive debt: An experiential learning framework for integrating AI in social work education. *Education Sciences*, 15(10), 1304. <https://doi.org/10.3390/educsci15101304>

DIALOG ON LANGUAGE INSTRUCTION (ISSN 1058-3858) is the journal of the Defense Language Institute Foreign Language Center, in the U.S. Department of Defense. The views expressed herein are those of the author(s), not the Department of Defense or its elements. Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the individual authors.