

TEACHERS' FORUM

Generative AI for L2 Materials: Disruptor or Perpetuator of the Status Quo?

KRISTEN MICHELSON

Texas Tech University

kristen.michelson@ttu.edu

Generative AI has been called a disruptor of the status quo. However, what does status quo mean when it comes to World Language (WL) education? In this reflective essay, I describe my forays into generative AI tools wearing the hat of a language program director and foreign language teacher-educator to explore the affordances and constraints of generative AI tools for materials development in WL education. First, I discuss what I consider to be the status quo in WL education, followed by a discussion of how materials creation is a literacy process. I then describe several use cases with AI tools (Twee, Questionwell, ChatGPT, Claude.ai), demonstrating uses that might constrain innovation in WL education, either by producing materials that align with past practices or by not affording a dialogic process of co-construction between user and tool. Finally, I explore ways that generative AI can serve as a helpful collaborator in language teacher workflows and conclude with implications for policy guidelines for AI use among WL teachers.

Introduction

Scholars Kohnke, Moorhouse and Zou (2023) describe generative artificial intelligence (AI) as a potential “disruptor of the status quo” (p. 544). I agreed at first, as I thought about the status quo in the way they likely meant it: in everyday life, educational settings, and complex organizations and systems. However, when I began thinking about the status quo in terms of world language (WL) education, I realized that without careful, principled use, generative AI might in fact perpetuate the status quo by constraining innovation in materials development, and by extension, in the materials design literacy of WL instructors.

As a Language Program Director (LPD) overseeing the work of graduate instructors in French, and as a professor preparing to teach the WL methods course after a five-year hiatus, I felt it was as much a matter of responsibility as curiosity to understand the affordances and limitations of specific AI tools. How would I guide instructors—both novice and experienced—who would inevitably turn to me to ask whether AI use was acceptable in creating lesson plans, classroom tasks, or assessment tools, or in guiding their students' use of these technologies? As a generally enthusiastic early adopter of new technologies, I approached generative AI with much more caution given its infamy for hallucinations, biased output, stolen data, and enormous carbon footprint. Nevertheless, in summer 2024 I dove in for some situated practice with AI tools to make my own informed judgments about their utility in order to better guide instructors.

In the sections that follow, I describe my early forays into using AI tools in my capacity as a WL teacher-educator, and share critical reflections on the affordances of various tools and tasks from

the perspective of my pedagogical principles, largely informed by multiliteracies approaches (Kumagai & Lopez-Sanchez, 2015; New London Group, 1996; Paesani, Allen, & Dupuy, 2016). Through these experiences, I have derived my own working set of guidelines for AI use by teachers, which I outline in this article's conclusion and share via appendix. These guidelines align with my fundamental values of moving beyond the status quo in WL education, adapting to technological advances, and providing teachers with opportunities to learn about new approaches and refine their own pedagogic principles.

The Status Quo in WL Education

Despite decades of scholarship in applied linguistics arguing for advances in WL teaching (see Dupuy & Michelson, 2021; Levine, Melin, Crane, Chavez, & Lovik, 2008; MLA Report, 2007), a chasm persists between theory and practice, and practices persist that reflect outdated methods. For example, an instrumental view of language often leads to transactional language use in the classroom (Byrnes, Maxim, & Norris, 2010; Kern & Schultz, 2005). Analysis of the complexity of language as constitutive of identities and ideologies is largely reserved for upper-level literature courses (MLA Report, 2007). Language practice frequently overshadows language use, where decontextualized grammar instruction takes center stage, with an overemphasis on explicit rules to the detriment of approaches that treat grammar as a tool for meaning-making (see Larsen-Freeman, 2015). Reading instruction tends to focus on comprehension of surface level facts and details followed by conversations about the topic (Stoller, Anderson, Grabe, & Komiyama, 2013), rather than on comprehensive textual analysis (Michelson & Anderson, 2022). Communication is generally anchored in students' own life worlds, to the detriment of communication anchored in discourse worlds of target language cultures (Liddicoat, 2000; Magnan, 2008). Finally, language and culture are often still taught separately, with culture often being disseminated through isolated informational texts (Liddicoat, 2008).

Materials Creation as a Literacy Process

Designing materials for WL instruction can be viewed as a literacy practice whereby teachers as designers engage in a constant and iterative cycle of drawing on available designs¹ to create lesson plans, assessments, task prompts, worksheets, and so forth. These materials orient learning activities and guide teachers' actions in the classroom. Lesson plans can help set learning objectives, pace various learning tasks, reflect on the balance between teacher talk and student talk, script the language we might use to invite students' engagement, and so forth (Farrell, 2002). Teachers' design decisions are motivated by a pedagogic purpose, which Mishan and Timmis (2015) describe as "the defining characteristic of materials" (pp. 2-3). In the process of designing, teachers make agentic decisions about content, scope, sequencing, etc., to enact their pedagogic purpose. As Bouckaert (2019) has suggested, materials development "can strengthen teachers' commitment to the curriculum, and their sense of ownership and agency" (p. 442).

An Immersion in AI Tools

In setting out on this project, my organizing questions were how and whether various AI tools might afford teachers agency in the literacy practice of designing materials as well as how well they might fare in the creation of materials for use in WL educational contexts. To this end, I immersed myself

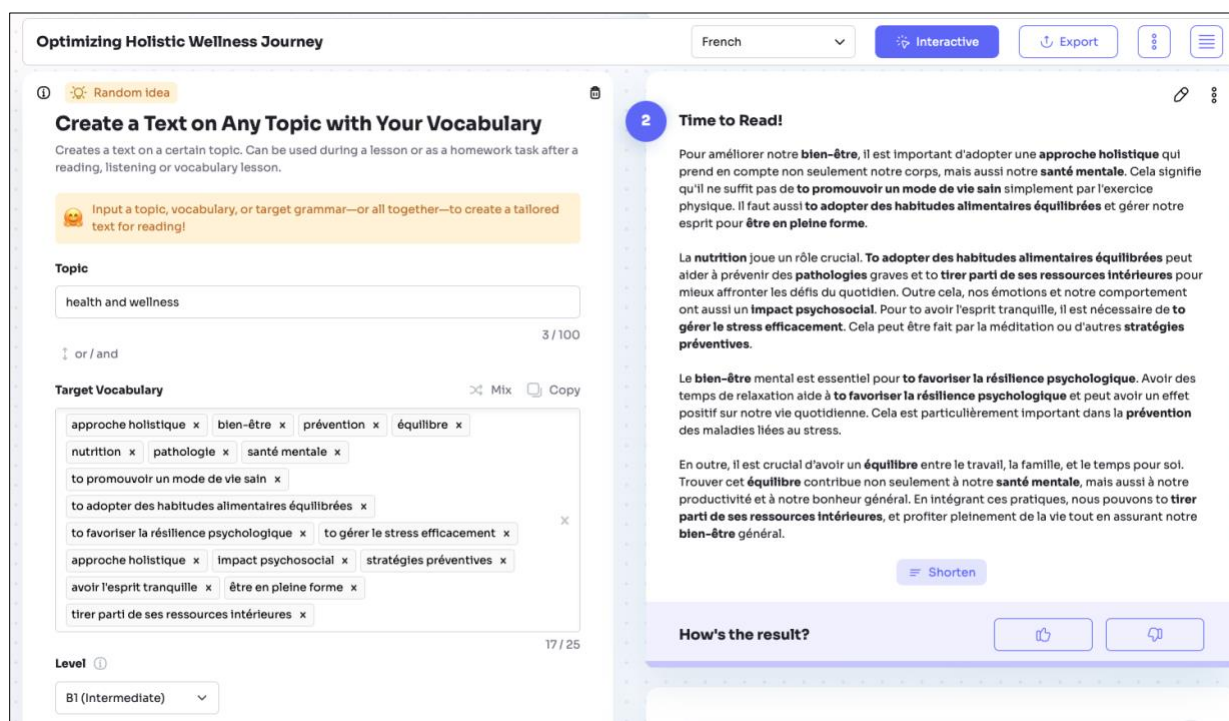
¹ I am using this term in the way it was proposed by the New London Group (1996) to mean everything brought to a literacy practice including, but not limited to, experiences of learning, conceptual knowledge, available material resources, linguistic resources and so forth.

in several apps I encountered that seemed to offer efficiencies for WL materials creation (Twee and Questionwell) as well as two more ubiquitous AI tools (ChatGPT and Claude).

Twee.com for Generating Texts

Twee.com operates on a freemium model and encourages users to “Create, share and grade language lessons in minutes.” It offers the possibility of creating custom materials, downloading ready-made lessons, and leveraging a variety of tools to create discussion questions, texts, and more, to align with the levels of the Common European Framework of Reference (CEFR) for language learning. Given my multiliteracies-informed interest in teaching through texts, I opted first for creating a text. Twee prompted me for the language (French), topic (health and wellness), part of speech (any), and level, which I could define by CEFR, grade, or age. I opted for the latter, selected 18+, then clicked “Do the magic!” Within seconds, I had before my eyes 16 entries consisting of verbs, nouns, and short phrases. Twee then provided options ranging from creating discussion questions to a dialogue on the topic, famous quotes, a supplemental text, and more. I went for the text, selected the B1 level and left the genre selection blank. I selected a length and clicked “do the magic!” Twee generated the following text (see Figure 1).

Figure 1
Screen Capture of Vocabulary List and Text Generated by Twee.com



In keeping with my pedagogical orientation, I reviewed this text through a multiliteracies lens. For me, a core tenet of multiliteracies pedagogies is that all language use is personal, cultural, and situational; we choose language to reflect our meanings based on our repertoires of grammar and vocabulary as well as expected social conventions of a particular communicative situation. A second important perspective guiding my pedagogical decisions is that texts provide access to discourse worlds and cultural conversations and therefore should be interpreted in their broader social contexts,

including considerations of identity and social purpose in their creation. Because this text was not human-created, it defied the very essential qualities of a text: there were no agentive design decisions in its creation; rather, it was generated as a result of statistical probabilities. When we outsource the process of generating a text to AI, we diminish human agency, and deprive learners of the opportunity to engage in a collaboration (Kern, 2003) with the text's creator in a dialogic process of meaning making. There is no human creator with whom to collaborate. When we ask our students to engage with AI generated texts, we are asking them to engage with vapid vessels of vocabulary, or, as Cope and Kalantzis (2023) might say, a "boring robo-text" (p. 7).

However, I decided not to walk away just yet and instead began thinking about how this AI-mediated materials development activity could present an interesting opportunity for learners. I thought about what could happen if we guided learners in critiquing this output with such questions as: If this were a human-generated text, in what kind of venue might this appear? Whose perspectives are missing? Is the message one-sided or are multiple perspectives offered? If YOU were writing a piece on *le bien-être*, what else would you include? However, without instructor-guided critical analysis of the shortcomings of this text, we are merely having students engage linguistically with empty ideas, thereby depriving them of opportunities to engage deeply with meaningful discourse and situated cultural messages. As Katharine Arens has said, "Texts are not only sources of linguistic knowledge and cultural facts but are also representations of how culture is created and how it functions; they must be engaged not only for language and elements of content, but also as strategic interventions on the field of culture—as examples of how knowledge is produced, circulated, and managed" (Arens, 2009, p. 162). Asking students to engage with AI-generated texts can entail limiting their engagement with culture and critical interpretations of how culture is created.

Putting my concerns aside, I succumbed to Twee's offer to create more, and I chose a dialogue this time (see Figure 2). As I scrutinized the text, I asked myself the kinds of questions I might ask students: "What do you notice about the formality of the language? Does this appear to be language that would be used between friends? Do they address each other by name or with a pronoun? Is the pronoun formal or informal?" I found the dialogues to be internally inconsistent, with unrealistic register shifts between informal ("*Salut, Paul!*" [Hey, Paul], which uses the familiar second person pronoun *tu*) and formal language (e.g., "*Oui, et il est également important de prendre en compte l'impact psychosocial sur notre bien-être*" [Yes, and it is equally important to consider the psychosocial impact on our well-being.]). In short, I found this to be a very unrealistic dialogue, masking as authentic conversation. I realized again that what I had was an excellent springboard for critical analysis of AI outputs, but not a cultural text that would invite my students into L2 discourse worlds. Upon further analysis, it also occurred to me that the very suggestion of a dialogue as a text type harkened back to audio-lingual methods.

Figure 2
Dialogue Generated by Twee.com

Optimizing Holistic Wellness Journey

Create a Dialogue on Any Topic
Creates a dialogue based on any text or topic. Give this dialogue to your students to read aloud or discuss!

Input a context or vocabulary, or both, to create a tailored dialogue for reading!

Text Image Website

Topic or text
health and wellness

3 / 1000

or / and

Target Vocabulary Mix Copy

bien-être x prévention x équilibre x nutrition x pathologie x
santé mentale x to promouvoir un mode de vie sain x
to adopter des habitudes alimentaires équilibrées x
to favoriser la résilience psychologique x to gérer le stress efficacement x
approche holistique x impact psychosocial x stratégies préventives x
avoir l'esprit tranquille x être en pleine forme x
tirer parti de ses ressources intérieures x

16 / 25

Level B1-B2

Do the magic!

4 Let's Dive into a Dialogue!

L Léa
Salut, Paul ! Tu as l'air très détendu aujourd'hui. Qu'est-ce qui se passe ?

P Paul
Hé, Léa ! Oui, j'essaie de vraiment me concentrer sur mon **bien-être** et ma **santé mentale**.

L Léa
C'est super important ! Est-ce que tu fais quelque chose en particulier pour ça ?

P Paul
Je travaille sur une **approche holistique** de la santé. Je pense qu'il faut favoriser la résilience psychologique. Ça aide à gérer le stress efficacement.

L Léa
Cela a du sens. J'essaie aussi d'adopter des habitudes alimentaires équilibrées. La **nutrition** joue un grand rôle dans notre **équilibre** général.

P Paul
Exactement ! Et j'ai lu que promouvoir un mode de vie sain est essentiel pour notre santé globale.

L Léa
Oui, et il est également important de prendre en compte l'**impact psychosocial** sur notre **bien-être**.

P Paul
Tout à fait. Parfois, je m'assois juste et je prends un moment pour **avoir l'esprit tranquille**.

L Léa
C'est une excellente idée. J'essaie aussi de tirer parti de mes ressources intérieures quand je me sens dépassée.

P Paul
En plus, c'est bon de connaître quelques **stratégies préventives** pour éviter les pathologies liées au stress.

L Léa
Absolument ! **Être en pleine forme** dépend aussi de notre état d'esprit et de nos choix quotidiens.

P Paul
Oui, on devrait s'encourager mutuellement à rester motivés.

L Léa
Bien sûr ! On pourrait même faire une activité ensemble pour renforcer notre **bien-être**.

Questionwell.org for Generating Reading Questions

Questionwell.org describes itself as, “[a] design studio that puts teachers in the driver’s seat while AI does the heavy lifting. Create, analyze, customize, and export standards-aligned materials to your favorite platforms in minutes.” Advertising a “free forever” model, Questionwell allows users to enter a source text, from which it generates comprehension and discussion questions as well as vocabulary. I pasted the text of an open access article we use frequently with third semester French students:

“*Biocoop investit dans du fast-food bio*” [Biocoop invests in organic fast food] (Détröyat, 2018). Rich with carefully composed and assembled images, this text also breaks stereotypes about fast food. Questionwell generated ten comprehension questions, five discussion questions, and five learning outcomes. I reviewed the comprehension questions (in the form of multiple-choice questions) to determine what was being assessed and saw that the questions seemed to be assessing recall of facts and details and could be answered by learners through skimming and scanning for key information.

Wanting to be more systematic in my approach to reviewing the output, I turned to one of my go-to frameworks: Afflerbach, Cho, and Kim’s (2015) hierarchical framework of reading tasks, which was modeled after Bloom’s Revised Taxonomy (Krathwohl, 2002). Using that lens, I found that seven of 10 questions addressed more basic skills such as remembering and understanding, while the other three questions addressed skills related to analyzing, corresponding to levels one through three in Afflerbach et al.’s seven-level framework. Discussion questions tended to be anchored in level four (applying) but relied on a high degree of knowledge about organic consumption trends in France and discourses surrounding Biocoop’s image, for which students would need much more context beyond this single text. Absent were questions about word choice and authorial intent, multimodal designs in the text, and extra-textual factors such as social contexts of production.

In short, these two AI tools seem to reflect instructional techniques of the past. Twee created texts that appear to exist for the sole purpose of decoding language. Questionwell’s reading questions focused primarily on surface-level comprehension to the detriment of questions that invite deeper textual analysis around multiple modes of communication and that account for authorship and audience factors.

An additional concern beyond cognitive complexity or the lack of attention to multimodal communication is AI’s general capacity to generate multiple choice questions. In the rapidly proliferating research on AI in language learning, several studies have explored the affordances and limitations of generative AI in this regard. Chun and Barley (2024) compared human versus AI-generated listening practice test items for Korean as a second language. They prompted ChatGPT to generate some multiple-choice questions (MCQs) using the Interlanguage Roundtable skill level descriptors, while also asking human experts to generate test items. They randomly combined the human- and AI-generated test items and asked three independent raters to rate these items overall. They found that the ratings given to the AI-generated and the human generated items were comparable, but through a more qualitative analysis, they found some weaknesses in the AI-generated items. Specifically, many of the distractors in the MCQs were often perceived as less plausible. Also, in certain cases other questions served as clues to the correct response. Sometimes there was a linguistic or structural contrast; for example, the key was often negative, and the distractors were often positive.

Finally, and perhaps most interesting: more than 50% of the time the correct answers were identifiable using general or world knowledge rather than using the listening passage. While the correct answer might have indeed been derived from the passage, the AI would have no sense about what constitutes world knowledge. As Kern (2024) asserts, “GPT does not really ‘know’ anything about language—or the world—rather, it predicts sequences of characters (and then larger chunks of language) based on their statistical likelihood of occurrence within the corpora on which it has been trained” (p. 519).

Rossi & Escubairó (2025) tested the affordances of a custom GPT for generating MCQs. They prompted ChatGPT Plus to generate MCQ items for a given text, targeting different levels of cognitive complexity from identifying surface level facts and details to more complex inferencing. Supplying their custom GPT with extensive documentation, including sample tasks and test specifications, they also asked it to provide metadata on each question. While Rossi and Escubairó actually found that GPT-4 did produce questions aimed at higher order processes, they also pinpointed many flaws such

as double key, lexical overlap between the stem and the response, misinterpreting of the text, and so forth. After having human test-takers trial the items, they found that items tapping inferencing of global comprehension were often deemed unclear or extremely difficult. Their overall conclusion was that GPT-4's performance on this task was "unpredictable." From these studies it seems that using AI to generate MCQs for reading yields unpredictable results that require extensive human refinements.

ChatGPT for Generating Lesson Plans

In another sitting, I explored how ChatGPT² would fare in creating lesson plans based on inputs of a grammatical and cultural focus. My initial prompt read: "I'm a French instructor at the college level teaching a course at the level of third semester French. I need a lesson plan for two 50-minute class sessions focused on the language function of "expressing hypothetical situations" using conditional verb forms. The cultural theme of the module in which these lessons sit is "la technologie" [sic]." The [generated plan](#) appears clean and organized, yet is highly problematic. The plan claims to "balance grammar practice (conditional verb forms for hypothetical situations) with cultural exploration of technology" (Open AI, 2025). Day 1's lesson features a predominant focus on form with a heavy emphasis on controlled practice. The language with which students are interacting is highly decontextualized: there is no integration of culture, nor is there any meaningful or shared discourse context in which students' responses in this controlled practice exercise might be situated. Day 2's lesson integrates an authentic text; however, the plan recommends a comprehension check and a cultural comparison with students' own experiences around technology habits, with no textual analysis nor critical thinking. When students *are* invited to enter into other life worlds, it is through imagined scenarios, not through a pre-established shared discourse context related to the L2. Focus on form is carried out through a matching activity focusing on recognition. The "communicative" task is centered around students' own uses of technology, and communication is anchored in the exchange of personal preferences and practices. The assessment suggestions again center on students using language in a cultural void, practicing language forms but not necessarily communicating.

I wondered what would happen if I pushed back against this lesson and continued the chat session with this follow up prompt³: "These lesson plans are not innovative and actually perpetuate the status quo in language teaching, which has been oft problematized. Can you see the problems with these lessons?" ChatGPT produced an analysis worthy of attention. Particularly striking were the following critiques:

- **"Form-as-endpoint, not meaning-as-action:** The conditional is treated as the target to practice rather than a resource to accomplish something consequential. This "PPP" arc (presentation–practice–production) reifies grammar drills and short, low-stakes talk."
- **"Thin interactional competence:** Hypotheticals in the wild rely on stance and mitigation (*on pourrait, il vaudrait mieux, je me demanderais si...*), turn-taking, disagreement strategies, and register choices (*tu/vous*). The plan centers verb morphology, not the interactional work."
- **"Tech as topic," not "culture-as-practice":** "La technologie" appears as a neutral theme. Absent are Francophone perspectives (e.g., CNIL & data privacy norms, platform alternatives, digital divides in DROM-COM, Minitel's legacy) or critical questions (surveillance, labor, accessibility)."

² In this sitting I was using ChatGPT 5.0, for which I had a trial subscription.

³At this moment I changed the setting to the "Thinking" option, which was offered with ChatGPT 5.0, where it "thinks longer for better answers" as compared with the "Instant" option where it "answers right away." (OpenAI, 2025)

- **“Assessment \approx attendance:** Participation and one-off sentences measure compliance, not performance. No success criteria, no product for an audience, no feedback cycle, and no evidence of transfer beyond the activity.” (OpenAI, 2025).

Upon pushing back, ChatGPT acknowledged the limiting view of grammar for grammar’s sake, a lack of attention to interactional work in authentic communication, the superficiality of the treatment of culture, and the lack of meaningful assessments. However, this was only upon pushing back. This experience corroborates work by Cope and Kalantzis (2023) who have advised: “This is the lesson for learners and teachers: don’t ask it to do more than it can. Give it facts, but don’t ask for them. Give it disciplinary concepts rigorously defined elsewhere if you are going to ask it to work with them. Supply specific context when needed.” (p. 22).

While my experience here is limited to one session, it is also affirmed by research by Dornburg and Davin (2024), who conducted an empirical study using AI to generate lesson plans. They used a zero-shot prompting approach to ask ChatGPT to create lesson plans aligned with the UNCC world language teacher licensure program and analyzed the resulting lesson plans. They repeated this approach with greater prompt specificity each time. Among other findings related to prompt specificity and output consistency, they found that some lesson plans reinforced behaviorist approaches by suggesting learners practice scripted role-plays and pronunciation through flashcards, also perpetuating the language-culture divide by focusing solely on language.

Discussion

Cope and Kalantzis (2023) have suggested: “[AI] can’t be innovative in the sense of creating new vectors of meaning. To this extent, it will tend always to average its responses to norms of the past” (p. 24). In all three of these cases, initial output from AI tools perpetuated the status quo by reifying traditional methods of language instruction. Twee created texts that appear to exist for the sole purpose of decoding and practicing language. In the case of Twee, the ease with which a user can toggle between creating texts, questions, dialogues, and vocabulary lists and customize these to learner levels makes this a tempting option for an overworked WL instructor with multiple course preps, overfilled classrooms, and/or simultaneous graduate study. In fact, apps like Twee eclipse the critical step of defining pedagogic purpose, which becomes obscured amidst a set of tempting shortcuts toward what appear to be useful WL materials. Questionwell’s reading questions focused on surface-level comprehension to the detriment of more cognitively challenging questions inviting deeper textual analysis. Finally, Chat GPT’s initial two-day lesson sequence reified a focus on form in decontextualized settings to the detriment of meaningful communication embedded in a particular cultural discourse context. When used in isolation from human designed pedagogies, these tools have the potential to constrain innovation in language teaching and perpetuate the status quo.

However, these tools have the potential to be useful collaborators in materials creation with thoughtful human engagement. Twee does appear to have some features (such as converting audio or video to text) that could be useful in crafting materials. ChatGPT could be strategically prompted with more specific criteria anchored in the pedagogical approaches of the user. While tools like Twee and Questionwell are limited because they do not afford dialogic and recursive prompting, chatbots like ChatGPT or Claude allow for more purposeful human-centered collaboration in materials design.

AI as Collaborator in Language Teacher Workflows

While these experiences testing the affordances of AI give me pause in light of the tools' potential constraints on innovation in language teaching, in other cases these tools have shown me their potential to simplify and shorten workflows. One of these tasks is creating rubrics. I have often used Claude.ai to create rubrics and have refined this practice through a prompt strategy in which I specify the criteria and the performance levels. Because performance descriptors generally consist of variations on a linguistic pattern, AIs are fairly good at generating tables of them.

Another helpful use of AI in simplifying language teacher workflows is in creating alternative versions of writing or speaking prompts. I tested this out last year in creating new prompts for a student taking a make-up exam, prompting Claude.ai with the following: "Beginning French students have an exam with the following two prompts. Please make an alternative version of the exam with different prompts which tap similar linguistic structures." I specified that "The topics do not need to be exactly the same," and gave a list of available topics. The next time I needed alternative versions of an exam, I returned to Claude.ai with a slightly more nuanced request. Instead of asking Claude.ai for an alternative version, I asked it to analyze the writing and speaking prompts for their implied communication functions, and the language forms needed to complete the tasks. Claude.ai generated five communication functions, seven different language forms needed, and some assessment considerations. I then asked Claude.ai to use the same parameters and generate a new writing prompt that assesses knowledge of the same communication functions and language forms. It produced a set of new prompts along with a justification (See Appendix A).

Although I had turned to Claude.ai for support in creating assessment materials, I immediately realized that the output could also serve as teaching tools for students. So, I asked Claude to align the communication functions and associated language forms needed to carry them out and it supplied a grid. (See Appendix B). While requiring some modifications, such a grid could be a helpful training tool for students to reflect explicitly on the fact that to engage in narrative storytelling, they might need the *passé composé*, the *imparfait*, and perhaps sometimes the present. To express personal future goals, they might need expressions of purpose such as "*pour, afin de, dans le but de*, etc." A grid like this can help students notice links between language forms and communication functions. Over time, after explicit use of such a grid, students might also become skilled at viewing a writing or speaking prompt and implicitly analyzing the communication functions inherent in the prompt. Such a practice could eventually help reinforce the notion that grammar alone is not the object of study but rather a tool for communication.

Implications for Policy Guidelines for Novice WL Instructors

My experiences testing the affordances and constraints of some available AI tools have led me to believe that the default outputs of many AI tools reflect the status quo in WL education. Given that "[m]aterials wield power in educational contexts, by shaping action, conveying knowledge, and collaborating with other human and nonhuman entities" (Guerrettaz, Engman, & Matsumoto, 2021, p. 9), the stakes are high. Beyond my concern about perpetuating the status quo in WL education, I remain concerned that over-reliance on AI limits instructors' agency and engagement in teaching. Bouckaert (2019) asserts that materials design is "craftsmanship" and is "defined by complexity, messiness and recursiveness" (p. 444). It "involves the teacher's engagement, anticipation and imagination, and results in a unique object in the shape of the newly created materials" (Bouckaert, 2017, p. 17). By outsourcing the materials design process to AI, teachers limit their opportunities to engage in this complexity and to bring their voice and interest to the process.

The conundrum remains: how does a WL teacher go about leveraging these tools to produce lessons and materials aligned with more contemporary approaches to language instruction if they have not yet had the opportunity to learn such approaches? This suggests to me the importance of joint pursuits with novice instructors in using AI to create materials, where teacher-educators model a process of leading with pedagogical principles and using AI as a collaborator, rather than leading with AI and tweaking the outputs to suit.

Consequently, for the short term, I have proposed a set of guidelines for use that facilitate our workflows, such as creating rubrics and alternative versions of assessments or task prompts and simplifying teacher language for lesson plan scripts. I recommend against the use of AI for generating texts, full reading question sets or lesson plans, and have provided rationales for each use case. (See Appendix C.) Longer term, I am designing a custom GPT that will be trained in the pedagogical principles and frameworks that are foundational to our program. Beyond merely stating rules, I encourage instructors to play with AI tools and discuss their experiences with the team so that everyone can collaboratively assess affordances and constraints and instructors can ultimately exert their agency in judiciously using AI. For now, general principles include:

- joint, mentored engagement with AI
- use of AI as collaborator rather than primary generator
- critical evaluation of AI output through pedagogical principles and student learning outcomes
- attention to ethical and environmental considerations

Conclusion

Returning to the original question about whether ChatGPT and AI are disruptors of the status quo, I would agree that AI has disrupted the status quo in our workflows as educators; however, if we want to disrupt the status quo in our pedagogical practices in WL education, AI might not help us but instead might constrain innovation if its outputs are used as an endpoint without principled, pedagogically-informed evaluation and refinement. Furthermore, certain AI uses could be detrimental to novice language teachers' development.

Developing instructional materials is a literacy practice that requires a clear sense of pedagogic purpose for each text and task. In WL education, it requires a keen understanding of foundational concepts in second language acquisition (SLA) and WL pedagogy, knowledge of evidence-based practices, an ability to set objectives that reflect contemporary approaches and to align materials with objectives. However, rather than ban AI use among instructors, I believe we should help WL instructors develop in using AI tools strategically, all while developing their conceptual knowledge of WL pedagogies. We must model a practice of leading with pedagogic principles and using AI as a helpful collaborator. Finally, it would benefit us to recognize that using AI for instructional purposes is a developmental process: as we try new techniques, we learn to optimize the tool. Ultimately, we must remain cautious that our AI uses do not perpetuate the status quo in our pedagogical practices in WL education.

Acknowledgements

I would like to thank Nisrine Slitine El Mghari and colleagues in the Department of Modern and Classical Languages, Literatures and Cultures at the University of Kentucky for the invitation to present an earlier version of this work at the 78th annual Kentucky Foreign Language Conference,

which provided an important forum for thinking through many of the issues raised in this article. I also wish to thank the anonymous reviewers for their insightful and helpful comments on this manuscript. Finally, I wish to thank Emily Hellmich and Kimberly Vinall and the editorial staff of L2 Journal for maintaining a venue for this kind of scholarly exchange and for their strong editorial support.

References

- Afflerbach, P., Cho, B.-Y., & Kim, J.-Y. (2015). Conceptualizing and assessing higher-order thinking in reading. *Theory Into Practice*, 54(3), 203–212. <https://doi.org/10.1080/00405841.2015.1044367203>
- Anthropic. (2025). *Claude (Claude 3.5)* [Large language model]. <https://claude.ai/>
- Arens, K. (2009). Teaching culture: The Standards as an optic on curriculum development. In V. Scott & C. Blyth (Eds.), *AAUSC 2009: Principles and Practices of the Standards in College Foreign Language Education* (pp. 160–180). Boston: Heinle. <https://www.slrjournal.org/item/259>
- Bouckaert (2017) Teachers' development of critically reflective practice through the creation of classroom materials. Unpublished doctoral thesis. London: University of Roehampton.
- Bouckaert, M. (2019). Current perspectives on teachers as materials developers: Why, what, and how? *RELC Journal*, 50(3), 439–456. <https://doi.org/10.1177/0033688218810549>
- Byrnes, H., Maxim, H. H., & Norris, J. M. (2010). The changing dynamics in collegiate foreign language education. *The Modern Language Journal*, 94(Supplement), 1-235
- Chun, J. Y., & Barley, N. (2024). A comparative analysis of multiple-choice questions: ChatGPT-generated items vs. human-developed items. In C. A. Chapelle, G. H. Beckett, & J. Ranalli (Eds.), *Exploring artificial intelligence in applied linguistics* (pp. 118-136). Iowa State University Digital Press. <https://doi.org/10.31274/isudp.2024.154.08>
- Cope, B., & Kalantzis, M. (2023). A multimodal grammar of artificial intelligence: Measuring the gains and losses in generative AI. *Multimodality & Society*, 4(2), 123-152. <https://doi.org/10.1177/26349795231221699>
- Détroyat, O. (2018, October 4). *Biocoop investit dans du fast-food bio*. Le Figaro. <https://www.lefigaro.fr/societes/2018/10/04/20005-20181004ARTFIG00227-biocoop-investit-dans-du-fast-food-bio.php>
- Dornburg, A., & Davin, K. J. (2024). ChatGPT in world language lesson plan creation: Trends, variability, and historical biases. *ReCALL*, 1–16. <https://doi.org/10.1017/s0958344024000272>
- Dupuy, B. & Michelson, K. (Eds.). (2021). *Pathways to Paradigm Change: Critical Examinations of Prevailing Discourses and Ideologies in Second Language Education*. Boston, MA: Cengage Learning.
- Farrell, T. S. C. (2002). Lesson planning. In J. C. Richards & W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 30–39). Cambridge University Press. <https://doi.org/10.1017/cbo9780511667190.006>
- Guerrettaz, A. M., Engman, M. M., & Matsumoto, Y. (2021). Empirically defining language learning and teaching materials in use through sociomaterial perspectives. *The Modern Language Journal*, 105(S1), 3–20. <https://doi.org/10.1111/modl.12691>
- Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *The Modern Language Journal*, 108(2), 515–533. <https://doi.org/10.1111/modl.12924>
- Kern, R., & Schultz, J. M. (2005). Beyond orality: Investigating literacy and the literary in second and foreign language instruction. *Modern Language Journal*, 89(3), 381–392. <https://doi.org/10.1111/j.1540-4781.2005.00312.x>
- Krathwohl, D. (2002). A revision of Bloom's Taxonomy: An overview. *Theory Into Practice*, 41, 121–218. https://doi.org/10.1207/s15430421tip4104_2
- Kumagai, Y., & López-Sánchez, A. (2015). Advancing multiliteracies in world language education. In Y. Kumagai, A. López-Sánchez, & S. Wu (Eds.), *Multiliteracies in world language education* (pp. 1-25). Routledge. <https://doi.org/10.4324/9781315736143>
- Larsen-Freeman, D. (2015). Research into practice: Grammar learning and teaching. *Language Teaching*, 48(2), 263–280. <https://doi.org/10.1017/s0261444814000408>
- Levine, G. S., Melin, C., Crane, C., Chavez, M., & Lovik, T. (2008). The language program director in curricular and departmental reform. *Profession*, 240–254. <https://doi.org/10.1632/prof.2008.2008.1.240>
- Liddicoat, A. (2000). Everyday speech as culture: Implications for language teaching. In A. Liddicoat & C. Crozet (Eds.), *Teaching languages, teaching cultures* (pp. 51–63). Melbourne, Australia: Applied Linguistics Association of Australia.
- Liddicoat, A. J. (2008). Pedagogical practice for integrating the intercultural in language teaching and learning. *Japanese Studies*, 28(3), 277–290. <https://doi.org/10.1080/10371390802446844>

- Magnan, S. S. (2008). The unfulfilled promise of teaching for communicative competence: Insights from sociocultural theory. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 349–379). London, UK: Equinox
- Michelson, K., & Anderson, A. (2022). Ideological views of reading in contemporary commercial French textbooks: A content analysis. *Second Language Research and Practice*, 3(1), 34–61.
- Mishan, F., & Timmis, I. (2015). *Materials development for TESOL*. Edinburgh: Edinburgh University Press. <https://doi.org/10.1515/9780748691371>
- MLA Ad Hoc Committee on Foreign Languages. (2007). Foreign languages and higher education: New structures for a changed world. *Profession*, 234–245. <https://doi.org/10.1632/prof.2007.2007.1.234>
- New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60–92. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>
- OpenAI. (2025). *ChatGPT (GPT-4)* [Large language model]. <https://chat.openai.com/>
- Paesani, K., Allen, H.W., & Dupuy, B. (2016). *A multiliteracies framework for collegiate foreign language teaching*. Pearson.
- QuestionWell ((2025). *QuestionWell* [Web app]. 2025 QuestionWell AI. <https://questionwell.org>
- Rossi, O., & Escubairó, J. M. M. (2025). Generating multiple-choice items for a B2 English reading test with GPT-4: targeting higher-order cognitive processing. [Preprint]. https://doi.org/10.35542/osf.io/qjfpb_v1
- Stoller, F. L., Anderson, N., Grabe, W., & Komiyama, R. (2013). Instructional enhancements to improve students' reading abilities. *English Teaching Forum*, 1, 2–11.
- Twee. (2025). *Twee* [Web app]. 2025 Twee AI LLC. <https://twee.com>

Appendix A

Chat Interaction with Claude.ai Around Creating Alternate Writing Prompts

Here is a prompt for my French students for their written exam. Analyze this prompt for the communication functions it elicits and the language forms needed in order to complete the tasks:

"You are planning on studying abroad next semester and intend to live with a host family. The exchange program requires that you describe your general daily routine in 10-12 sentences so that they can match you up with a host family that will best suit you."

I'll analyze this French writing prompt to identify the communication functions it elicits and the language forms students would need to use effectively.

Communication Functions

1. **Describing habitual actions** - Students need to outline their typical daily activities
2. **Expressing time and sequence** - Explaining when activities occur and in what order
3. **Personal description** - Sharing personal preferences and habits
4. **Future plans** - Contextualizing the routine in relation to upcoming study abroad
5. **Implied comparison** - Indirectly suggesting compatibility with potential host families

Language Forms Needed

1. **Present tense** - For describing regular/habitual actions (je me lève, je prends, etc.)
2. **Reflexive verbs** - For daily personal care (se lever, se coucher, se brosser les dents)
3. **Time expressions** - To sequence the day (le matin, à midi, après le dîner, etc.)
4. **Frequency adverbs** - To indicate how often activities occur (souvent, parfois, toujours)

5. **First-person narration** - The entire response will be in "je" form
6. **Transition words** - To connect ideas and create flow (ensuite, puis, après)
7. **Vocabulary related to** - Daily routines, meals, school/work, leisure activities

Assessment Considerations

- The 10-12 sentence requirement provides clear parameters for assessment
- Students should demonstrate variety in vocabulary and structures to avoid repetition
- The practical context makes the task authentic and purposeful
- Time management is important as students need to cover a full daily routine within the sentence limit

This prompt effectively tests students' ability to describe routine activities in French while maintaining a clear communicative purpose that simulates a real-world language task.

Using these same parameters, please write a new writing prompt that assesses knowledge of the same communication functions and language forms.

Here's a new writing prompt that assesses the same communication functions and language forms:

Prompt:

"You have been selected to participate in a cultural exchange program with a French high school. Before your arrival, your French 'correspondent' (pen pal) wants to know about your lifestyle and preferences. Write 10-12 sentences describing your typical weekend activities, including what you enjoy doing, when you do these activities, and any special habits or preferences you have that would help your correspondent prepare for your visit."

This prompt maintains the same assessment parameters as the original by:

1. **Targeting the same communication functions:**
 - Describing habitual actions (weekend routines)
 - Expressing time and sequence
 - Personal description (preferences and habits)
 - Implied future context (the upcoming visit)
 - Implied compatibility (helping the correspondent prepare)

2. Requiring the same language forms:

- Present tense for habitual actions
- Reflexive verbs for personal activities
- Time expressions to organize the weekend timeline
- Frequency adverbs to indicate regularity of activities
- First-person narration
- Transition words to connect ideas
- Vocabulary related to leisure, meals, and personal habits

3. Maintaining similar assessment parameters:

- 10-12 sentence requirement
- Authentic communicative purpose
- Clear audience (the French correspondent)
- Need for varied vocabulary and structures

This variation shifts the focus from weekday routines to weekend activities, which might elicit a different range of vocabulary while still requiring the same grammatical structures.

Can you do the same analysis of the following prompt:

"You are preparing to go on the job market for a job in France (it can be any sector you like). In preparation you are drafting a cover letter that you could submit for an application, once you see an ad for a job that interests you. Even if you don't have the specific job yet, write a general cover letter of 10-12 sentences in which you discuss: previous work experience, strengths and weaknesses, your personality, etc. Seeing as this is for a future job application, your answer must also be professional and well-structured to demonstrate your proficiency with the language."

I'll analyze this French writing prompt to identify the communication functions it elicits and the language forms students would need to use effectively.

Communication Functions

1. **Professional self-presentation** - Students need to present themselves in a positive, professional manner

2. **Describing past experiences** - Outlining previous work experience
3. **Self-assessment** - Evaluating personal strengths and weaknesses
4. **Character description** - Explaining personality traits relevant to employment
5. **Persuasive communication** - Convincing a potential employer of suitability
6. **Formal communication** - Adhering to professional correspondence conventions

Language Forms Needed

1. **Past tenses** - Passé composé and imparfait for describing previous work experiences
2. **Present tense** - For describing current skills and personality traits
3. **Formal register** - Using appropriate politeness formulas and formal vocabulary
4. **Conditional structures** - Potentially for expressing future possibilities (je pourrais, je serais)
5. **Professional vocabulary** - Field-specific terminology related to chosen career
6. **First-person narration** - Written from "je" perspective
7. **Connective phrases** - To create a cohesive, well-structured document
8. **Evaluative adjectives** - To describe personal qualities and skills

Assessment Considerations

- The 10-12 sentence requirement provides clear parameters while forcing concision
- The professional context requires attention to formality and register
- Students must demonstrate both narrative skills (past experiences) and descriptive abilities (personal qualities)
- The open nature of the job sector allows students to incorporate vocabulary relevant to their interests
- The emphasis on structure and professionalism adds an additional assessment dimension beyond pure linguistic accuracy
- Students must balance positive self-promotion with realistic self-assessment (strengths and weaknesses)

This prompt effectively tests a more formal and professional communication style while still requiring students to describe themselves and their experiences, representing a clear progression in communicative complexity from the daily routine prompt.

Thank you. Now can you write a new prompt?

Here's a new writing prompt that assesses the same communication functions and language forms as the job application cover letter prompt:

Prompt: "You have been selected as a finalist for a prestigious scholarship to study at a French university. As part of the final selection process, you must submit a personal statement to the scholarship committee. Write 10-12 sentences in which you present your

academic background, highlight your strengths and areas for growth, describe your character traits that make you an ideal candidate, and explain how this scholarship aligns with your future goals. Remember that this is a formal document that will be evaluated by academic professionals, so your writing should be well-structured and demonstrate your advanced French language abilities."

This prompt maintains the same assessment parameters as the job application cover letter by:

Targeting the same communication functions:

- Professional/formal self-presentation
- Describing past experiences (academic instead of work)
- Self-assessment of strengths and weaknesses
- Character description relevant to the opportunity
- Persuasive communication to a selection committee
- Formal communication with appropriate register

Requiring the same language forms:

- Past tenses for describing academic history
- Present tense for current qualities and characteristics
- Formal register appropriate for academic professionals
- Conditional structures for future plans
- Academic vocabulary instead of professional terminology
- First-person narration
- Connective phrases for cohesion
- Evaluative language for personal qualities

Maintaining similar assessment parameters:

- 10-12 sentence requirement
- Emphasis on structure and formality
- Balance between past experiences and current qualities
- Need to be persuasive while remaining authentic
- Requirement for organized, cohesive writing

This variation shifts the context from job-seeking to academic advancement while preserving the formal nature and self-presentational aspects of the original prompt.

Can you analyze the related speaking prompt:
 Congratulations! You've been called in for an interview! The interviewer has asked you a number of questions and finishes by asking you to discuss a time you felt proud of something you did at work. A good response to this question will be lengthy (roughly a minute long) and will provide specific details about your job and this particular experience which could be a time you faced a challenge, resolved a problem or situation, took on a responsibility or leadership position, etc."

I'll analyze this French speaking prompt to identify the communication functions it elicits and the language forms students would need to use effectively.

Communication Functions

1. **Narrative storytelling** - Students need to recount a specific workplace experience
2. **Expressing feelings/emotions** - Describing the feeling of pride and potentially other emotions
3. **Describing past events in sequence** - Chronologically explaining what happened
4. **Self-promotion** - Highlighting personal accomplishments in a positive light
5. **Providing context** - Setting up the workplace scenario with sufficient detail
6. **Problem-solution structure** - Potentially describing a challenge and how it was overcome

Language Forms Needed

1. **Past tenses** - Primarily passé composé and imparfait to narrate past events
2. **First-person narration** - Speaking from personal experience using "je"
3. **Descriptive adjectives** - To convey emotions and characterize the situation
4. **Time markers** - To establish sequence of events (d'abord, ensuite, finalement)
5. **Cause and effect structures** - Explaining why certain actions were taken (parce que, donc)
6. **Professional vocabulary** - Job-specific terminology relevant to the scenario
7. **Emotional vocabulary** - Words to express pride, satisfaction, accomplishment
8. **Transition phrases** - To maintain flow during a minute-long response

Assessment Considerations

- The time requirement (approximately one minute) tests students' ability to sustain extended discourse
- The prompt requires spontaneous language production in an interview context
- Students must balance narrative detail with efficiency to meet the time requirement
- The professional context demands appropriate register and vocabulary

- Self-evaluation is embedded in the task (explaining why they felt proud)
- The prompt allows flexibility for students to choose experiences matching their language proficiency
- Successful responses will demonstrate both linguistic accuracy and effective storytelling

This speaking prompt effectively assesses students' ability to produce extended, cohesive discourse about past professional experiences while incorporating emotional reflection—skills that go beyond basic conversational French and enter the realm of professional communication.

Can you write a new prompt that is related thematically to the new writing prompt and achieves the same assessment purposes?

Here's a new speaking prompt that thematically connects to the scholarship application writing prompt while assessing the same communication functions and language forms:

Speaking Prompt: "Félicitations! You've advanced to the interview stage for the prestigious French university scholarship. After discussing your academic background, the scholarship committee asks you to describe a specific academic achievement or project that made you particularly proud. Your response should be approximately one minute long and include details about the context, the challenges you faced, the steps you took, and why this achievement was significant for your personal and academic development. Be prepared to explain how this experience demonstrates qualities that make you an ideal scholarship recipient."

This prompt maintains the same assessment parameters as the original speaking prompt by:

- 1. Targeting the same communication functions:**
 - Narrative storytelling about a specific experience
 - Expressing feelings of pride and accomplishment
 - Describing past events in sequence
 - Self-promotion in a relevant context
 - Providing sufficient background information
 - Problem-solution structure
- 2. Requiring the same language forms:**
 - Past tenses to narrate the experience
 - First-person narration
 - Descriptive adjectives for emotions and situations
 - Time markers to organize the chronology

- Cause and effect structures to explain reasoning
 - Academic vocabulary instead of professional terminology
 - Emotional vocabulary to convey significance
 - Transition phrases for a cohesive one-minute response
- 3. Maintaining similar assessment parameters:**
- One-minute length requirement for extended discourse
 - Spontaneous language production in an interview setting
 - Balance between detail and concision
 - Appropriate formal register for academic context
 - Self-reflection component
 - Flexibility to choose a topic matching language abilities

This new prompt creates a cohesive assessment pair with the scholarship application writing prompt, shifting the focus from workplace achievement to academic accomplishment while preserving the essential linguistic and communicative demands.

Can you align the communication functions and associated language forms needed to carry them out in a grid?

I'll analyze this French writing prompt to identify the communication functions it elicits and the language forms students would need to use effectively.

Communication Functions

1. **Professional self-presentation** - Students need to present themselves in a positive, professional manner
2. **Describing past experiences** - Outlining previous work experience
3. **Self-assessment** - Evaluating personal strengths and weaknesses
4. **Character description** - Explaining personality traits relevant to employment
5. **Persuasive communication** - Convincing a potential employer of suitability
6. **Formal communication** - Adhering to professional correspondence conventions

Language Forms Needed

1. **Past tenses** - Passé composé and imparfait for describing previous work experiences
2. **Present tense** - For describing current skills and personality traits
3. **Formal register** - Using appropriate politeness formulas and formal vocabulary

4. **Conditional structures** - Potentially for expressing future possibilities (je pourrais, je serais)
5. **Professional vocabulary** - Field-specific terminology related to chosen career
6. **First-person narration** - Written from "je" perspective
7. **Connective phrases** - To create a cohesive, well-structured document
8. **Evaluative adjectives** - To describe personal qualities and skills

Assessment Considerations

- The 10-12 sentence requirement provides clear parameters while forcing concision
- The professional context requires attention to formality and register
- Students must demonstrate both narrative skills (past experiences) and descriptive abilities (personal qualities)
- The open nature of the job sector allows students to incorporate vocabulary relevant to their interests
- The emphasis on structure and professionalism adds an additional assessment dimension beyond pure linguistic accuracy
- Students must balance positive self-promotion with realistic self-assessment (strengths and weaknesses)

This prompt effectively tests a more formal and professional communication style while still requiring students to describe themselves and their experiences, representing a clear progression in communicative complexity from the daily routine prompt.

Appendix B

Claude.ai Generated Grid of Communication Functions and Language Forms

Communication Function	Associated Language Forms
Narrative storytelling	<ul style="list-style-type: none"> • Past tenses (passé composé, imparfait) • Chronological sequencing • Descriptive details • First-person perspective ("je")
Expressing feelings/emotions	<ul style="list-style-type: none"> • Emotional vocabulary (fier/fière, satisfait(e), content(e)) • Descriptive adjectives • Idiomatic expressions of feeling • Intensity adverbs (très, vraiment, extrêmement)
Describing past events in sequence	<ul style="list-style-type: none"> • Time markers (d'abord, ensuite, puis, enfin) • Transition words (pendant ce temps, par la suite) • Temporal conjunctions (quand, lorsque, après que) • Varied past tenses for foreground/background events
Self-promotion	<ul style="list-style-type: none"> • Positive descriptive vocabulary • Phrases highlighting personal contribution • Strategic use of first-person pronouns • Balanced modesty and confidence expressions
Providing context	<ul style="list-style-type: none"> • Setting vocabulary (academic terminology) • Descriptive language for situation • Imperfect tense for background information • Prepositional phrases for location/setting
Problem-solution structure	<ul style="list-style-type: none"> • Cause and effect expressions (parce que, donc, par conséquent) • Contrastive conjunctions (cependant, malgré, pourtant) • Goal-oriented vocabulary • Process description terminology
Connecting to future goals	<ul style="list-style-type: none"> • Future and conditional tenses • Expressions of purpose (pour, afin de, dans le but de) • Aspirational vocabulary

Communication Function	Associated Language Forms
	<ul style="list-style-type: none"> • Linking phrases between past achievements and future plans
Formal academic register	<ul style="list-style-type: none"> • Sophisticated vocabulary • Complex sentence structures • Appropriate level of formality • Academic/discipline-specific terminology

Appendix C

Guidelines on using AI in Materials Creation for French Courses

GenAI = Generative artificial intelligence

LLM = Large Language Model

General principles:

Before using AI for materials creation for courses in our program, consider the following principles related to 1) human-led collaborative use; 2) pedagogical grounding; and 3) ethical and environmental considerations.

1. **Use AI as a collaborator, not a creator.**
Do not use AI as your starting point or primary generator of materials. Instead, engage in iterative, dialogic prompting (rather than zero-shot prompting).
2. **Engage in joint, mentored use.**
Bring AI use cases to team meetings for collective discussion and refinement. Share your experiences to improve prompting practices and our collective knowledge over time.
3. **Anchor AI use in research-based pedagogical principles.**
Critically examine all AI output through established theories and frameworks in second language acquisition (SLA) and world language (WL) pedagogy.
4. **Frame prompts with disciplinary knowledge and learning outcomes.**
When generating materials, supply AI with explicit pedagogical principles to ensure alignment with instructional goals.
5. **Ecological impact.**
Generative AI requires substantial computing power and has a significant carbon footprint. Each time you use AI, consider whether its pedagogical value justifies its environmental and human costs.
6. **Authenticity and authorship.**
AI-generated texts lack human authorship and identity, existing in a cultural void. Recognize the decontextualized nature of AI output and instead consider using texts created by human authors.
7. **Bias.**
Because LLMs reflect biases in their training data, approach all output critically and be prepared to revise, challenge, or reject it using specific analytical frameworks.
8. **Privacy.**
Do not input others' texts or personal data into generative AI tools. Assume that all user data and input may be used for further model training.
9. **Limitations on innovation.**
AI output is inherently past-oriented as it is trained on existing data. Recognize that AI use can limit innovation and novel ideas.

Specific use cases:

This document is a work in progress and will be updated as new cases arise. If you have a use case that is not listed below, please consult with me so that we can collectively evaluate your case and determine whether it meets acceptable uses and add it to the appropriate section in these guidelines.

Use cases	Sample prompt	OK?	Why/why not?
<p>Creating rubrics</p>	<ul style="list-style-type: none"> ▪ Please create a rubric. ▪ The criteria are: Purpose/Task: Degree to which the writing sample accomplishes the task and includes relevant information; Organization and coherence: Degree to which the writing sample exhibits a logical and coherent sequence throughout; provides a clear sense of a beginning, middle, and end. Makes smooth transitions between ideas; Vocabulary: Degree to which the writing sample includes a wide variety of vocabulary that expands the topic, even if there are minor inaccuracies; Language Control: Degree to which the writing sample demonstrates a high degree of control of structures and conventions such as subject-verb agreement, tense, noun-adjective agreement, correct word order, spelling/ accents; Comprehensibility: Degree to which the writing sample achieves communication with few or no errors hindering the overall comprehensibility of the sample. ▪ The performance levels and scores are: 4 = Meets/exceeds expectations 3 = Approaches expectations 2 = Approaches expectations 1 = Needs improvement 0 = Did not submit. 	<p>YES</p>	<p>The language of rubrics is patterned. LLMs operate on pattern recognition and statistical probabilities of characters and words appearing in a string.</p> <p>There is neither personal nor identifying nor proprietary data contained in the prompt.</p>
<p>Creating alternate versions of a prompt for a classroom task or an assessment</p>	<ul style="list-style-type: none"> ▪ Beginning French students have an exam with the following two prompts. Please make an alternative version of the exam with different prompts which tap similar linguistic structures. "Writing: You are visiting your family members or friends for the new year. In a paragraph of about 6 to 8 sentences, describe the leisure activities you and your family/friends are doing during your visit. Speaking: Imagine you are the <i>météorologiste</i> (weatherman/woman) for your school news channel. Give the weather report for the day. Make sure to say what day it is, what the temperature is going to be, what the weather is like, and give any advice one might give for that type of weather. Your weather report should be at least 1 minute long and should be spontaneously spoken (you should not be reading off your planning 	<p>YES</p>	<p>The language of rubrics is patterned. LLMs operate on pattern recognition and statistical probabilities of characters and words appearing in a string.</p> <p>There is neither personal nor identifying nor proprietary data contained in the prompt.</p>

	<p>paper). The topics do not need to be exactly the same. Available topics for either prompt are: family life, friends, university life, cafe culture, leisure time activities.”</p>		
<p>Creating simplified teacher talk</p>	<ul style="list-style-type: none"> Enter your teacher script into the chatbot and ask to simplify your language into “comprehensible input for beginning French learners.” 	<p>YES</p>	<p>The language of rubrics is patterned. LLMs operate on pattern recognition and statistical probabilities of characters and words appearing in a string. They are also good at modifying the pattern.</p> <p>There is neither personal nor identifying nor proprietary data contained in the prompt.</p>
<p>Asking AI to create your entire lesson plan</p>	<ul style="list-style-type: none"> I'm a French instructor at the college level teaching a course at the level of third semester French. I need a lesson plan for two 50-minute class sessions focused on the language function of "expressing hypothetical situations" using conditional verb forms. The cultural theme of the module in which these lessons sit is "<i>la technologie</i>". 	<p>NO</p>	<p>LLMs are trained on massive amounts of data. Therefore, the output is always oriented to the past, and as such, constrains innovation. Research has shown that ChatGPT tends to produce lesson plans for WL education that are oriented around outdated methods (see Dornburg & Davin, 2024). Personal experience has corroborated this research.</p>
<p>Generating texts for reading</p>	<ul style="list-style-type: none"> Please create a text with the following vocabulary on the topic of "<i>le bien-être</i>". 	<p>NO</p>	<p>All language use is personal, cultural, and situational; we choose language to reflect our meanings based on our repertoires of grammar and vocabulary expected social conventions of a particular communication situation. Texts provide access to discourse worlds and cultural conversations and therefore should be interpreted in their broader social contexts, including the role of identity in their creation.</p> <p>AI-generated texts do not have authorship or identity and exist in a cultural void.</p>
<p>Creating multiple</p>	<ul style="list-style-type: none"> Please create some multiple-choice questions around the following text... 	<p>NO</p>	<p>Reading is a complex process involving interpreting both surface level propositions</p>

<p>choice reading questions</p>		<p>and inferencing (interpreting propositions that are not explicitly stated). AI-generated questions tend to assess recall of facts and details but do not assess learners along higher levels of cognitive complexity. They only tap the linguistic mode and do not address multimodal designs in a text.</p>
--	--	---