The GALL of it all: Grading and teaching in the age of GenAI-assisted language learning

Joo Jin Sim 🕞

National Institute of Education, Nanyang Technological University, Singapore

I watched the young man, dressed in scuffy denim and T-shirt, dark curly head bent in concentration, juggling a book and a stack of papers on his lap as the subway lurched through the darkness. At first glance, I thought that he was working on his assignment, writing quickly and intently, in a race with the subway and its inevitable stops. As a load of passengers spewed into the car, he shifted slightly to accommodate the new crush, and it was then that I saw varied and youthful handwriting on his pages. He was grading those papers, not composing his own ... (Miller, 1986, p. 26).

Much current exploration of GenAI-assisted language learning (yes, GALL) seems determinedly opposed to the idea of young men juggling books and papers in trains. Scholars want students to be free from the need to wait and the fear of embarrassment (Barrot, 2023; Hong, 2023; Godwin-Jones, 2022). The passage below echoes the optimism expressed by most scholars exploring GenAI's affordances:

Best of all, the feedback is instant, unlike teacher feedback, which understandably takes time ... by the time students receive their feedback, they might have completely forgotten everything from the previous writing. All in all, ChatGPT is a free and more efficient alternative to human tutors. (Hong, 2023, p. 40)

As a teacher myself, I too appreciate peaceful train rides unmolested by the need to grade papers. Any teacher you meet will tell you that he or she prefers a page from Murakami or Austen to a hastily scrawled 850-word composition on the pros and cons of smartphones. Every encounter with a subject-verb error, misplaced comma, and dangling modifier is another reminder that there's more work to do.

Yet I believe teachers will continue grading and students will continue to wait, not because we don't have faith in the powers of GenAI, but because we know the value of the apparently Sisyphean slog that so much of teaching appears to be. While often persuasive and clearly well-intentioned, exploration of GenAI's affordances has to date often left unexamined the rich and complex work that grading and other seemingly thankless tasks do for both teachers and learners. To develop strategies and principles for the integration of GenAI chatbots into instructional contexts, the current wave of GenAI research needs to engage in deeper dialogue with what second language acquisition (SLA) scholarship has learnt about vital language learning processes, including feedback, planning, and exposure to model texts.

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



1. Instant feedback whenever you want it!

Few would disagree with the notion that writers need feedback: it helps learners identify gaps in their repertoire and confirms the hypotheses that they make when they experiment with newly acquired language (Swain, 1995). With these principles in mind, many scholars have pushed for greater adoption of GenAI as a source of instant feedback.

Yet feedback can come in many forms, and not all of them enjoy a history of uncontested beneficence. Most scholars hail GenAI's ability to correct grammatical errors as one of its chief affordances (e.g., Barrot, 2023; Hong, 2023; Su et al., 2023), but this may not necessarily benefit learners. Written corrective feedback (WCF) can be comprehensive or selective, depending on how many error types a teacher chooses, and they can be selected before or after grading. In general, evidence from the field suggests that focusing on a smaller set of error types is both more manageable in terms of the cognitive load learners incur and the impact it has on motivation (Lee, 2020). In contrast, comprehensive WCF may overwhelm and discourage learners, especially beginning writers already lacking in confidence and skill. Besides the question of how much feedback to give, there has also been extensive debate over whether WCF should come in the form of explicit correction or hints without correction. Here, results have been less conclusive, and suggest that various factors, including task type, student proficiency, and error type, determine which might have more positive impact on learners. Broadly speaking, hints work when students can correct errors independently, whereas explicit WCF works best for more complex errors (Lee, 2013). Since hints encourage greater cognitive engagement, providing explicit correction indiscriminately - which is almost always likely to be what a chatbot does, unless users request hints instead - deprives more advanced learners of the chance to work things out themselves.

Yet all this is not to suggest that GenAI feedback cannot benefit learners. Unlike a teacher, a GenAI chatbot cannot readily determine how likely a particular learner is to correct an error himself or herself; this problem can be solved if a learner first asks for indirect WCF before asking a chatbot to provide explicit corrections, thus allowing him or her to benefit from the higher cognitive engagement afforded by indirect WCF and the greater support from direct WCF. Such a proposal is an example of what can be developed when we evaluate and refine an affordance in the light of extant SLA research.

So the prospect of instant feedback may not be as simple as it first appears, at least when it comes to students. But what about teachers? Wouldn't less grading free up time, allowing them to design better lessons, and stronger curricula?

Grading, not composing: this is of course literally true – the young man Miller observed was not writing an essay of his own. Yet through annotation, recast, and highlighting, teachers in fact construct a portrait of the learner richer and more vivid than whatever report GenAI can provide in the form of statistics and generalizations. We understand more deeply and retain longer whatever knowledge we've arrived at through deep processing; the information we acquire quickly without analysis and effort, however, evaporates quickly (Marton & Säljö, 1976a; Marton & Säljö, 1976b). We know that when the script is returned to them, our students do not slowly work their way through every tick, cross, or comment, but leap immediately to find out what grade they got. But the work we have done has not been for naught, for through our reading and rereading of each student's script we come to a deeper awareness of where he or she is, and where he or she needs to go. Consciously or not, we pay special attention to aspects of performance that correspond to the lessons we have taught, and use it to evaluate and refine our lessons. In effect, grading is a core part of the reflective practitioner's dialogue with the experiences he has designed:

... the designer may take account of the unintended changes he has made in the situation by forming new appreciations and understandings and by making new moves. He shapes the situation, in accordance with his initial appreciation of it, the situation "talks back," and he responds to the situation's back-talk. (Schön, 1983, p. 79) Without the hard slog of grading behind the scenes, it is doubtful that a teacher's classroom performance can be as relevant and urgent as what her students need. For besides the portraits of individual learners, a kind of unconscious tabulation and cross-comparison is at work; the more patient practitioner may well keep a running record of common errors and clear strengths, but it may be more reasonable to expect most teachers beleaguered by fatigue and dogged by deadlines to leave such work to the less conscious, more implicit processes of the mind. Teachers emerge from the process clearer about what to do next and where to steer their students. Perhaps the class has shown that they are ready to move on to a new genre. More likely, there are clear errors that the majority has demonstrated, and longstanding offences that smaller groups have yet to eradicate from their performance. Such conclusions, lodged in the teacher's mind through the quiet, constant hammer of analysis, comparison, and rereading, drive a teacher to make important decisions. And we can make these decisions with conviction because we have seen the evidence ourselves, experienced the confusion wrought by poor syntax, or chortled at the rich bathos of inappropriate register. We know that we might lose precious curriculum time because of this deviation, but we make the decision nonetheless because it is urgent. Our students need the detour.

CURRICULUM essentially means a race course. It shares the same root as CORRIDOR, CURRENT, COURSE, words that evoke straight lines or circles. But the curricula drawn up before the start of the year are really tentative drafts; the actual course of study is inscribed through the decisions made by teachers from day to day. Instead of a smooth untroubled path, the circle of a curriculum might perhaps be more accurately envisioned as a line that is smooth only in places, to be interrupted hither and thither by a recursive series of loops whenever teachers need to go over old material and reinforce old messages. If we do end up offloading a substantial part of our students' written output to GenAI for grading, what we gain in terms of time for lesson planning may ultimately come to nothing if, without close encounters with our students through grading, we lack the understanding needed to rethink our teaching.

2. Ideas organized the way you like it!

Besides instant feedback, scholars have also discussed the tantalizing prospect of learners offloading the planning process to GenAI. Planning in general involves goal setting, idea generation, and organization (R. T. Kellogg, 1996); if attention is a limited resource, investing too much of it in these processes might leave little left for choosing the right word, and constructing a more complex sentence (R. T. Kellogg, 1994). In fact, trade-offs occur not only between planning and language but within the latter as well: learners have to choose between extending the complexity of their language, maintaining accuracy, or focusing on fluency (Skehan, 1998). These trade-offs make the idea of GenAI as a writing assistant highly attractive. Barrot (2023, p. 57), for instance, argues that a GenAI tool like ChatGPT can not only suggest "essay topics based on the user's area of interest" but also create or "transform any outline into a sentence, topic, alphanumeric, or decimal system format" that help learners develop their own outlines.

When idea generation is no longer a challenge, a learner should have sufficient cognitive resources left over to devote to the translating process; in addition, if the ideas proposed by the chatbot are sophisticated enough, they may push the learner to retrieve the linguistic resources necessary for expressing these ideas (Robinson, 2001). Given that learners involved in writing are already adopting a more syntactic mode of processing (as compared to learners who are processing information), this could help them expand their linguistic repertoire. If they do not find the linguistic resources necessary for additional input, either from more knowledgeable others in the form of peers or human tutors or the chatbot itself (Swain, 1995).

However, research from a sociolinguistic perspective suggests that this may not always occur. Instead of deploying the available resources for enhancing the complexity and accuracy of their output, learners may well focus on simply completing their task; ironically, this may be because they treat the task as a real-world experience and focus on communicating meaning (Ortega, 1999; Batstone, 2005). In any case, how learners wish to complete a task can never be completely controlled by the task designer (Breen, 1987).

Batstone (2005) suggests that a learning orientation is needed; Skehan (1996) sees adventurousness and risk-taking as necessary preconditions for the productive use of freed-up resources. In both cases, what a learner needs is a culture or environment that encourages mistakes and "failing forward." It has been suggested that chatbots might create such preconditions by removing the anxiety learners experience in actual interaction with humans; yet freedom from embarrassment also means the absence of an authentic audience who might give meaning to the act of writing. No matter how warm or encouraging its tone might be, a chatbot is not another human being whose respect we cherish, whose laughter we would like a pun or quotation to elicit. Empirical investigation has consistently shown that students' writing quality rises when addressing authentic audiences other than teachers (Block & Strachan, 2019; Cohen & Riel, 1989; Wiggins, 2009), possibly due to greater motivation to process and store information about such audiences in long-term memory (Magnifico, 2010). How much motivation, then, can we expect to find in students when the intended audience is not even their human teacher?

Thus the impact of planning is not as direct as one might initially believe. Yet we can refine our use of GenAI to support planning by considering studies on task planning effects (Ellis, 2021; Johnson & Abdi Tabari, 2023). For instance, (Ingley and Pack, 2023) recommend requesting a chatbot to play the role of an instructor who can help a writer brainstorm for ideas via a dialogue; this echoes teacher-fronted or guided planning, which Skehan and Foster (1999) found to be effective in promoting gains across both complexity and accuracy. While their study focused on oral discourse, such findings suggest that guidance in the planning process can help learners overcome trade-off effects, the phenomenon where focusing on one dimension of language performance leads to concomitant dips in other dimensions due to limited attentional resources. Teachers would benefit from empirical investigation on whether a prompt like "You are an English language teacher – engage me in a dialogue to plan the ideas and language I can use for an essay on whether video games help or harm teenagers" could lead to the type of task performance Skehan and Foster, (1999) described.

3. Model texts at the drop of a hat!

Much of the enthusiasm that GenAI has inspired can be traced to its remarkable ability to produce what commentators often describe as uncannily humanlike texts. In the heady months after its launch, articles written by ChatGPT quickly went viral once they were published by trusted news providers. This capacity to immediately and almost flawlessly produce texts customized according to users' specifications has given rise to claims that GenAI can provide model texts that serve as input for language learning (e.g., Kohnke et al., 2023).

Yet if GenAI output appears increasingly humanlike, is it because writing has become steadily more robotic and mechanical? This might well be truer in the field of formal education than other sectors of modern life. For decades now, scholars have debated the value of the five-paragraph essay (Brannon et al., 2008) or the PEEL paragraph (Gibbons, 2019; McKnight, 2021); the genre approach, originally developed to enlighten second language learners about the organizational structures that first language learners implicitly possess (Cope & Kalantzis, 1993), has in many cases been oversimplified into a "recipe" approach where students are assessed on their ability to reproduce texts with a specific number of paragraphs for each part of an essay (Derewianka, 2003). This formulaic approach

sounds exactly like what Barrot (2023) has observed: a GenAI composition "typically starts with a definition and a brief history of the concept", "discusses the effects", and concludes with a "summary of main points, final thoughts and call to action (p. 4)."

As a teacher, I have often found it difficult to find authentic texts that match the features and organization of a particular genre. The regularities and patterns that analysts report (e.g., Derewianka & Jones, 2016) have been distilled from a great number of texts; one would be hard pressed, however, to find any one text that might exemplify all the defining traits of the genre it belongs to. A teacher who curates texts from the wild is obliged to explain to students that these texts represent what writers actually do, and that the genre conventions they learn are only tendencies to bear in mind; consequently, a student comes to develop a greater appreciation for the freedom that writers actually enjoy. In contrast, learners who primarily rely on a steady diet of texts custom-made to fit the specifications of a certain genre will come to believe that effective writing is all clockwork convention and rigorous regularity. And like the serpent swallowing its own tail, students fed a regular diet of GenAI output may well become teachers who see GenAI writing as the only model texts their students need.

Being a product of such a system, I know how difficult it is to resist the instinct to fall back on PEEL and the structure of the five-paragraph essay when attempting an unfamiliar topic. Yet, as Bereiter and Scarmadalia, (1987) have been telling us since 1987, writing is more than simply retelling what we know; it is a transformative process where what we think we know changes through the need to express it. Writers are their first own readers: a gap in our rhetoric – some word that doesn't feel quite right, a sentence that couldn't capture what we hoped to express, or paragraphs that just don't link up very well – sometimes pushes us to reevaluate the thought that eluded our language; going back to reconsider this thought, we might rework the string of propositions that comprise it and emerge with fresh insight, which in turn sets in motion another search for the right word, the best syntax, a better way to organize our text. And so it is through these recursive cycles of self-reading and rewriting that the text takes shape, a text organized according to the stream of ideas it hopes to address, simultaneously familiar enough to fit within the rules of the genre it belongs to and different enough to suit the writer's purpose and personality. A writer knows when he needs to be pragmatic and efficient, and when he needs to work things out more patiently and reflectively; which path he chooses determines to a large extent how completely he clings to the rules of a particular genre.

Good writing is ludic, a chess match where both reader and writer win if they work together to bend the rules. The rules are the features and conventions of each genre; a writer can choose to operate safely within the space these expectations demarcate or opt for the road less taken. In her essay on AI-generated writing, Morrison (2023, p. 158) describes how her attempts at subverting reader expectations make her writing seem "almost willfully insubordinate, self-sabotaging." GenAI output, in contrast, is "a marvel of correct, mild-mannered, balanced, objective prose," properly organized and studiously formatted. Would it be quixotic to believe that the ability to craft an imperfect text, the kind where one might find a distinctly recognizable voice inseparable from certain inflections of tone and quirks of syntax, might represent a certain advantage in a world where everyone's writing appears increasingly similar?

Could GenAI be a collaborator in the process of developing a personal voice? As we have seen with feedback and planning, doing so requires a strategic approach that draws on existing scholarship. In their case study of a young academic adept at integrating ChatGPT into every phase of her writing process, Jacob et al. (2024) describe how "Kailing" rejects ChatGPT output that does not match her personal style even when it appears eloquent; and how she begins to rely less and less on the chatbot when she detects recurrent patterns in its lexicogrammatical choices. In a word, it is possible to OUT-GROW GenAI. Thus the prototypicality that characterizes GenAI output can be harnessed two ways, depending on where a learner is as a writer: at the early stage when genre knowledge is lacking, its output demonstrates the language features that should be emulated; whereas at the later stages when

learners are ready to advertise their identity as unique voices, GenAI output represents the kind of language one should seek to avoid, modify, or subvert.

4. The intelligent use of intelligent things

As I try to stay focused on completing this essay, my son is waiting for me to pause so that he can show me his latest effort at writing, a diary entry in Mandarin chronicling this morning's exciting events: going to school with his sister, having toast for breakfast, taking the bus. He wrote it sitting next to me, stopping every now and then to ask me how to write a particular word or phrase or whether he had chosen the right punctuation mark. He tries to hide the words from me, insisting that I must not read it before he is done. As I (secretly, slyly) read his composition, certain turns of phrase catch my eye. I notice the chunks he has borrowed and acquired from the texts I've been reading to him, the personal favorites that I've (secretly, slyly) hoped he will also come to love.

I wonder how this relationship might change if I told him to send his texts to a chatbot instead for instant feedback. I would prefer him to come up with topics that really matter to him, that come from his own observations about the world. Because I care about what he reads, I would rather he learn good writing from texts I myself have enjoyed and seek to emulate. I believe many teachers feel the same way about their students.

More than two decades ago, Warschauer and Healey (1998, p. 67) urged teachers to explore "the intelligent USE of CALL" while awaiting the arrival of "intelligent CALL." While the explosion of literature on GenAI suggests that the wait is over, there are still plenty of questions for intelligent teachers to discuss. If teachers grade less than what they currently do, how might the quality of teaching be affected and how much do learners benefit from GenAI feedback? If learners outsource planning to a chatbot, would the attention thus freed up necessarily be channelled towards language? Would model texts that flawlessly reproduce the features of a certain genre INSPIRE good writing?

Such questions are critical but also difficult to answer alone. Teachers have the experience of classroom work to turn to, and their relationships with students to draw on. But we would fare much better if the research community supports us with the use of carefully constructed studies that seek to more fully unravel the impact of GenAI on the experience of language learning.

Competing Interests. The author declares no competing interests.

References

- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. Assessing Writing, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745
- Batstone, R. (2005). Planning as discourse activity: A sociocognitive view. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 277–295).

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Routledge.

- Block, M. K., & Strachan, S. L. (2019). The impact of external audience on second graders' writing quality. *Journal of Literacy and Language Arts*, 58(2), 68–93. https://scholarworks.wmich.edu/reading_horizons/vol58/iss2/5.
- Brannon, L., Courtney, J. P., Urbanski, C. P., Woodward, S. V., Reynolds, J. M., Iannone, A. E., Haag, K. D., Mach, K., Manship, L. A., & Kendrick, M. (2008). EJ Extra: The five-paragraph essay and the deficit model of education. *English Journal*, 98(2), 16–21. https://doi.org/10.58680/ej20086823
- Breen, M. (1987). Learner contribution to the task design. In C. N. Candlin, & D. Murphy (Eds.), *Language learning tasks* (Vol. 7, pp. 23–46). Prentice-Hall International.
- Cohen, M., & Riel, M. (1989). The effect of distant audiences on students' writing. *American Educational Research Journal*, 26(2), 143–159. https://doi.org/10.3102/00028312026002143

Cope, B., & Kalantzis, M. (Eds.). (1993). The powers of literacy : A genre approach to teaching writing. Falmer Press.

Derewianka, B. (2003). Trends and issues in genre-based approaches. *RELC Journal*, 34(2), 133–154. https://doi.org/10.1177/003368820303400202

Derewianka, B., & Jones, P. (2016). Teaching language in context. Oxford University Press.

- Ellis, R. (2021). The effects of pre-task planning on second language writing: A systematic review of experimental studies. *Chinese Journal of Applied Linguistics*, 44(2), 131–165. https://doi.org/10.1515/CJAL-2021-0009
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language teaching research*, 3(3), 215–247. https://doi.org/10.1177/136216889900300303
- Gibbons, S. (2019). "Death by PEEL?" The teaching of writing in the secondary English classroom in England. *English in Education*, 53(1), 36–45. https://doi.org/10.1080/04250494.2019.1568832
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5–24. http://doi.org/10.1177/10125/73474
- Hong, W. C. H. (2023). The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, 5(1), 37–45. https://jeti.thewsu.org/index.php/cieti/article/ view/103
- Ingley, S. J., & Pack, A. (2023). Leveraging AI tools to develop the writer rather than the writing. *Trends in ecology & evolution*, 38(9), 785–787. https://doi.org/10.1016/j.tree.2023.05.007
- Jacob, S., Tate, T., & Warschauer, M. (2024). Emergent AI-assisted discourse: A case study of a second language writer authoring with ChatGPT. Journal of China Computer-Assisted Language Learning. https://doi.org/10.1515/jccall-2024-0011
- Johnson, M. D., & Abdi Tabari, M. (2023). Planning in L2 writing: A research synthesis and meta-analysis. *System*, *118*, 103152. https://doi.org/10.1016/j.system.2023.103152
- Kellogg, R. T. (1994). The psychology of writing. Oxford University Press.
- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy, & S. Ransdell (Eds.), The science of writing: Theories, methods, individual differences and applications (pp. 57–72). Erlbaum.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. https://doi.org/10.1177/00336882231162868
- Lee, I. (2013). Research into practice: Written corrective feedback. *Language Teaching*, 46(1), 108–119. https://doi.org/10.1017/ S0261444812000390
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49, 100734. https://doi.org/10.1016/j.jslw.2020.100734
- Magnifico, A. M. (2010). Writing for whom? Cognition, motivation, and a writer's audience. *Educational Psychologist*, 45(3), 167–184. https://doi.org/10.1080/00461520.2010.493470
- Marton, F., & Säljö, R. (1976a). On qualitative differences in learning: I Outcome and process. *British Journal of Educational Psychology*, 46(1), pp. 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x
- Marton, F., & Säljö, R. (1976b). On qualitative differences in learning: II Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology*, 46(2), 115–127. https://doi.org/10.1111/j.2044-8279.1976.tb02304.x
- McKnight, L. (2021). Since feeling is first: The art of teaching to write paragraphs. *English in Education*, 55(1), 37–52. https://doi.org/10.1080/04250494.2020.1768069
- Miller, J. (1986). Marking papers and marking time: How women and men perceive as teachers. *Journal of Natural Inquiry and Reflective Practice*, 1(1), 26–37. https://commons.und.edu/tl-nirp-journal/vol1/iss1/5
- Miller, J. (2010). Persistence of the five paragraph essay. English Journal, 99(3), 99-100. https://doi.org/10.58680/ej20109513
- Morrison, A. (2023). Meta-writing: AI and writing. *Composition Studies*, 51(1), 155–219. https://compstudiesjournal.com/ current-issue-spring-2023-51-1/
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in second language acquisition*, *21*(1), 109–148. https://doi.org/10.1017/S0272263199001047
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Christopher Brumfit Essay Prize Winner* (pp. 287–318). Cambridge University Press.
- Schön, D. A. (1983). The reflective practitioner: How professionals think in action. Basic Books.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17(1), 38–62. https://doi.org/10.1093/applin/17.1.38
- Skehan, P. (1998). A cognitive approach to language learning. Oxford University Press.
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. Assessing Writing, 57, 100752. https://doi.org/10.1016/j.asw.2023.100752
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidlhofer (Eds.), Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson (pp. 125–144). Oxford University Press.
- Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31(2), pp. 57–71. https://doi.org/10.1017/S0261444800012970
- Wiggins, G. (2009). Real-world writing: Making purpose and audience matter. English Journal, 98(5), 29–37. https://doi.org/ 10.58680/ej20097118

Joo Jin Sim is a teacher in Singapore. As a Lead Curriculum Resource Development Specialist in the Ministry of Education, he guides the translation of research into multimedia resources, digital tools, and curricular materials that

8 Joo Jin Sim

support the teaching of English Language and Literature in English. His research interests lie in the field of second language acquisition, particularly the psychology of language learning and instructional approaches like task-based language teaching and genre-based pedagogy. Joo Jin is currently pursuing a Ph.D. in the National Institute of Education, Nanyang Technological University. For his doctoral study, he is developing a novel task sequencing model based on the Limited Attentional Capacity approach and examining its impact on the written task performance of Singaporean adolescent learners.

Cite this article: Sim, J. J. (2025). The GALL of it all: Grading and teaching in the age of GenAI-assisted language learning. *Language Teaching*, 1–8. https://doi.org/10.1017/S0261444825100785