Using AI to Support Teaching and Learning of Multiword Sequences

Christopher J. Hollis
Tottori University

Reference Data:

Hollis, C. J. (2026). Using AI to support teaching and learning of multiword sequences. In B. Lacy, M. Swanson, & H. Yoder (Eds.), *LanguageS: Learning, Teaching, Assessing – JALT 50 Years – Challenges and Perspectives*. JALT. https://doi.org/10.37546/JALTPCP2025-02

Multiword sequences (MWSs) – combinations of words that frequently recur together – are a key component of L2 proficiency and oral fluency (Paquot, 2019; Tavakoli & Uchihara, 2020). Recent work shows that the density and length of MWSs and lexical bundles strongly influence listener-based ratings of fluency (Brown et al., 2023; Hougham et al., 2023; Takizawa & Suzuki, 2025). Building on this research, and on proposals for principled GenAI use (Hougham & Higginbotham, 2024) and systematic vocabulary and multiword-item instruction (Nakata & Suzuki, 2022), this paper reports an exploratory use of GenAI chatbots in a university course to emphasize MWSs as an essential part of L2 vocabulary acquisition. It outlines GenAI-mediated tasks for teaching idioms, incongruent collocations, and academic multiword sequences, and discusses perceived benefits, drawbacks, and challenges. It also offers ideas for using GenAI to support learners' MWS development while maintaining learner agency and academic integrity.

多語表現（multiword sequences: MWS）は，頻繁に共起する語の組み合わせであり，第2言語（L2）の熟達度と発話の流暢さを支える重要な要素である（Paquot, 2019; Tavakoli & Uchihara, 2020）。近年の研究では，MWS や語彙束ルの密度と長さの双方が，聞き手による流暢さの評価に強く影響することが示されている（Brown et al., 2023; Hougham et al., 2023; Takizawa & Suzuki, 2025）。本稿では，これらの研究知見に加え，生成AIの原則的な活用（Hougham & Higginbotham, 2024）及び体系的な語彙・多語表現指導（Nakata & Suzuki, 2022）に関する提案を踏まえ，MWS をL2 語彙習得の中核的要素として重視した大学授業における生成AIチャットボットの試験的な活用について報告する。本稿では，イディオム，非直訳的コロケーション，及びアカデミックな多語表現を対象とする生成AI仲介タスクの概要を説明し，その利点，欠点，及び課題を議論する。さらに，学習者の主体性と学術的誠実性ィを維持しつつ，生成AIを用いて学習者の MWS 発達を支援するためのアイデアを提示する。

Multiword sequences (MWSs) such as idioms, collocations, phrasal verbs, and lexical bundles (see Appendix for working definitions and examples) are now widely viewed as central to L2 speech fluency rather than optional "extras" for advanced learners (Peltonen, 2024; Tavakoli & Uchihara, 2020). Recent studies show that more proficient L2 speakers tend to use a wider range of MWSs, draw on them more frequently, and select more sophisticated sequences than less proficient peers, and that these patterns are visible in both learner corpora and listener judgments (Takizawa & Suzuki, 2025; Vercellotti, 2021). Listener-based research indicates that temporal fluency indices such as speech rate and mean length of run are closely associated with how speakers deploy MWSs in real time (Asari, 2021; Brown et al., 2023; Hougham et al., 2023), supporting the view that formulaic language is integral to how fluent a speaker sounds, not just how fast they talk.

In many Japanese university English classes, however, instruction and assessment still focus primarily on single words and sentence-level grammar, leaving formulaic language undertaught and underpracticed (Nakata & Suzuki, 2022; Uchihara & Clenton, 2020). Corpus-informed work and classroom observations suggest that students often know many individual words but have limited control of high-utility formulaic expressions such as discourse markers, stance expressions, and idiomatic collocations, which can make their speech sound hesitant or unnatural despite reasonable grammatical accuracy (Nordlund & Norberg, 2025). However, generative artificial intelligence (GenAI) tools such as ChatGPT and Google Gemini have made it easy for learners to access large amounts of contextualized language, including MWSs, while also raising concerns about hallucinations, over-reliance, and academic integrity in higher education (Hougham & Higginbotham, 2024; Kohnke et al., 2023; Sims, 2023). These tensions make it crucial to consider how GenAI can be harnessed to promote, rather than undermine, learners' independent control of MWSs.

Developments such as these pose a practical and ethical question for language teachers: How can GenAI be used to support principled teaching and learning of

MWSs without turning learners into passive copiers of AI-generated text? Recent work on AI-supported language education and policy suggests that, when used with clear guidelines, GenAI tools can provide detailed explanations, level-appropriate input, and rich practice opportunities (Hunte & Wheatley-Holness, 2025; Teng, 2023; UNESCO, 2023; US Department of Education, 2023). At the same time, authors emphasize the need for transparent classroom policies and explicit discussion of acceptable and unacceptable uses of GenAI so that learners remain active agents in their own language development (Heaven, 2023; Hougham & Higginbotham, 2024; Rudolph et al., 2023). In short, the challenge is to design GenAI-mediated tasks that foreground learner agency and engagement with MWSs rather than passive dependence on AI output.

Against this background, this paper reports on classroom practices at a Japanese national university that used GenAI chatbots to foreground MWSs as an essential part of L2 vocabulary learning. The paper, based on classroom observations and student feedback, has three primary aims: (1) to highlight the importance of incorporating MWSs into university English curricula, (2) to share GenAI-mediated activities designed to support effective teaching of MWSs and learners' acquisition of them, and (3) to discuss the benefits, drawbacks, and pedagogical challenges of using GenAI large language model (LLM) applications in foreign language education.

Literature Review

Background: Why Teach Multiword Sequences?

A growing body of recent research shows that multiword sequences are central to L2 spoken fluency and should be treated as core learning targets. Studies of formulaic sequences and perceived proficiency indicate that learners who use a wider range of appropriate MWSs are rated as more proficient than peers who rely mainly on word-by-word constructions, even when their vocabulary size is similar (Durrant & Schmitt, 2009; Paquot, 2019; Vercellotti, 2021). Studies on listener judgments also demonstrate that temporal fluency indices such as speech rate and mean length of run are closely tied to how speakers deploy MWSs in real time (Brown et al., 2023; Hougham et al., 2023). This stands in alignment with Tavakoli and Uchihara's (2020) finding that an increased frequency of MWS usage directly correlates with a faster articulation rate while simultaneously minimizing mid-clause breakdowns, providing structural processing advantages that ease a speaker's real-time temporal delivery. Ultimately, as investigations into perceived oral production highlight, a robust command of formulaic language heavily dictates how fluent an L2 speaker sounds to evaluators (Asari, 2021). These collective empirical insights underscore why MWSs were treated as core speaking targets in the present course design.

Recent studies have highlighted in particular the role of lexical bundles, recurrent multiword strings that serve textual and interactional functions, in L2 oral proficiency. Hougham et al. (2023) showed that the length of lexical bundles contributes differentially to holistic proficiency ratings, with both shorter and longer bundles offering unique benefits for how advanced learners sound in oral tasks. Their findings suggest that teachers should expose students to a range of bundle types, from short verb-object or adjective-noun combinations to longer discourse-organizing sequences such as "on the other hand, I think…" This conclusion is strongly reinforced by work on pragmatic markers and perceived fluency in rating scales (Brown et al., 2023). Listener-based research by Takizawa and Suzuki (2025) similarly indicates that raters' fluency judgments are sensitive to the density and appropriacy of MWSs in L2 argumentative speech, even when basic speed and pause measures are relatively similar across speakers (Tavakoli & Wright, 2020). Taken together, this body of work underlines that what learners say with MWSs can matter as much as how quickly they say it.

From a processing perspective, formulaic sequences are argued to ease the cognitive load of speaking by allowing learners to retrieve larger chunks as single units rather than constructing every utterance word by word (Pawley & Syder, 1983; Wray, 2002). Recent syntheses of fluency research emphasize that speed, breakdown, and repair measures are tightly intertwined with learners' ability to mobilize such chunks to structure discourse and signal relationships between ideas (Peltonen, 2024; Tavakoli & Wright, 2020). When learners lack commonly used MWSs for core communicative functions such as expressing opinions, agreeing and disagreeing, and signaling contrast, they frequently resort to word-by-word constructions that sound hesitant or unnatural, even when grammatical accuracy is relatively high (Peltonen, 2024; Wood, 2010, 2020). In communicative classrooms, this means that helping learners build an accessible repertoire of MWSs is as important as teaching them to produce grammatically accurate sentences.

In the Japanese EFL context, these issues are particularly salient. Corpus-informed work has shown that students often know many individual words but have limited experience with high-utility formulaic expressions that occur frequently in authentic spoken English, including discourse markers, stance expressions, and idiomatic collocations (Nakata, 2025; Nordlund & Norberg, 2025; Uchihara & Clenton, 2020; Wood, 2015). Classroom-based studies and reviews argue that effective pedagogy should therefore help learners identify, understand, and repeatedly use such expressions in

meaningful contexts, rather than treating them as incidental by-products of general vocabulary learning (Nakata, 2025; Paquot, 2019; Tavakoli & Uchihara, 2020). These recommendations align with broader vocabulary and fluency research, which shows that explicit attention to high-frequency formulaic expressions, combined with spaced retrieval and productive use, can substantially improve learners' fluency and confidence (Durrant & Schmitt, 2009; Garner & Crossley, 2018; Hougham et al., 2023). In this study, these insights shaped the choice of target MWSs and the design of GenAI-mediated practice cycles.

Despite this growing body of evidence, incorporating MWS instruction into already crowded curricula remains challenging. Teachers must balance textbook requirements, test preparation, and diverse learner needs, while also deciding which MWSs to prioritize and how to provide repeated, contextualized practice (Tavakoli & Wright, 2020; Wood, 2015). These practical constraints motivated the search for tools that could help generate rich examples and practice opportunities for MWSs without overwhelming teachers' preparation time, which is one reason GenAI was explored as a pedagogical resource in the present study. The following section shows how GenAI was used to relieve some of these practical pressures while keeping MWSs at the center of instruction.

Methodology

Design: Using GenAI to Support MWS Teaching and Learning

This practice-oriented report describes how GenAI chatbots were integrated into one elective English communication course at a Japanese national university to support the teaching and learning of multiword sequences. The class size was 18 2nd-, 3rd-, and 4th-year students from a range of majors. Lessons met once a week for 90 minutes over one semester. Course objectives emphasized spoken communication at approximately CEFR B1-B2 level, with regular pair and group tasks, short presentations, and written assignments.

GenAI chatbots–specifically, ChatGPT (OpenAI, 2025) and Gemini (Google, 2025)–were introduced as support tools for noticing, exploring, and practicing MWSs rather than as sources of complete texts. At the start of the course, students received a short orientation that (a) explained what GenAI is and how LLMs generate language, (b) outlined classroom rules for acceptable/unacceptable use, and (c) emphasized that AI output must always be checked, adapted, and personalized (Hougham & Higginbotham, 2024; Kohnke et al., 2023). Students were explicitly told not to paste AI-generated sentences directly into graded assignments and were reminded that, as Hougham and Higginbotham put it, "students should use generative AI to support their writing processes rather than to produce finished products that they simply submit as their own work" (2024, p. 14). Ultimately, the aim was to position GenAI as a scaffold for building students' own MWS repertoires, not as a shortcut around genuine self-expression. In short, AI was framed as a *support* for learning, not a tool for doing all the work for the students.

The overall design followed three principles derived from recent work on MWSs, vocabulary learning, and GenAI in education. First, activities targeted high-utility MWSs that are frequent in authentic discourse and relevant to the course tasks, such as idioms for everyday topics, incongruent collocations that differ from Japanese and English, and academic expressions for discussions and presentations (Durrant & Schmitt, 2009; Nordlund & Norberg, 2025; Tavakoli & Uchihara, 2020). Second, GenAI was used to provide rich, level-appropriate input and prompts that could support repeated encounters and productive use of these sequences, aligning with research on spaced retrieval and practice for vocabulary learning (Garner & Crossley, 2018; Nakata et al., 2020). Third, each activity built in reflection and verification stages in which students evaluated AI suggestions against their own intuitions, peer feedback, and teacher guidance, echoing recent recommendations to foster critical digital literacy and avoid uncritical dependence on AI (Hougham & Higginbotham, 2024; Rudolph et al., 2023; UNESCO, 2023). In practical terms, this meant that no GenAI output was treated as "correct" until students had questioned and adapted it.

Within this framework, three clusters of GenAI-mediated activities were designed: (1) raising awareness of idioms and collocations, (2) exploring incongruent collocations that differ between Japanese and English, and (3) practicing academic multiword sequences for discussions and presentations. Each cluster followed a similar cycle: students prompted a chatbot to generate candidate expressions in context; selected and edited target MWSs; incorporated them into short dialogues, mini-scenes, and scripts; and then used these in speaking tasks where classmates listened for and tallied occurrences of the target sequences. Homework and quizzes recycled MWSs that multiple students identified as new and useful, aiming to support both recognition and productive use over time, in line with research highlighting the value of repeated, contextualized encounters for formulaic language learning (Durrant & Schmitt, 2009; Garner & Crossley, 2018; Nakata et al., 2020). The next section illustrates how these principles played out in the GenAI-mediated classroom activities.

Results

Activities: GenAI-Mediated Activities for MWSs

Activity 1: Raising Awareness of Idioms and Collocations

The first cluster of activities aimed to raise students' awareness of idioms and collocations as key types of MWSs and to link this awareness to their own vocabulary learning goals. After a short teacher-led introduction that contrasted transparent combinations (e.g., drink coffee) with less transparent idioms and collocations (e.g., hit the books, break the ice, spill the beans), students worked in pairs with a GenAI chatbot to generate expressions related to familiar themes such as travel, university life, and relationships. A typical prompt asked the chatbot to "list five frequently used idioms about travel at CEFR B2 level, with short, simple explanations and example sentences."

Pairs skimmed the output, selected several idioms that were new or unclear, and asked follow-up questions to check meaning and usage (e.g., "Can you give two more examples of 'hit the books' in easy English?"). The students then wrote short dialogues and mini-stories that incorporated their target idioms in context, first drafting with support from AI when needed and then revising to make the language more personal. For example, after prompting the chatbot for B2 travel idioms, one student pair selected "on a shoestring budget" and adapted the AI's example into their own role-play dialogue: "We traveled around Tottori on a shoestring budget, staying only at inexpensive hotels." The students then presented their findings in small groups and then to the class. In subsequent lessons, selected dialogues were used as role-plays and short listening texts, and idioms that multiple students identified as useful were recycled in quizzes and follow-up speaking tasks. This cycle allowed GenAI to supply varied, theme-linked examples while keeping human evaluation and adaptation at the center of learning, consistent with calls for explicit, contextualized practice of formulaic expressions (Durrant & Schmitt, 2009; Paquot, 2019).

Activity 2: Exploring Incongruent Collocations

The second set of tasks focused on collocations that differ between Japanese and English, an area where learners often produce non-targetlike combinations despite knowing the component words (Koya, 2003; Uchihara & Clenton, 2020). Building on a brief explanation of incongruent collocations, students first brainstormed everyday Japanese expressions such as *kopii wo toru* (make a copy, not take a copy) and *koe ga futoi* (have a deep voice, not have a thick voice) and predicted how these might be expressed in English. They then used a GenAI chatbot to check their hypotheses, prompting it to "give four frequently used English collocations that are different from Japanese for the topic of daily school life, with simple explanations of the differences and CEFR B2 example sentences."

Students compared AI-suggested collocations with their own guesses and with teacher feedback, highlighting pairs where literal translations would be misleading or sound unnatural. In particular, learners focused on Japanese-English contrastive pairs where literal translation fails, substituting non-targetlike expressions such as "take a copy" or "thick voice" with natural collocations like "make a copy" and "deep voice." Each pair chose two or three collocations that were new or surprising and wrote short dialogues and mini-scenes that contrasted non-targetlike and targetlike expressions, sometimes drawing on humorous misunderstandings similar to those described in work on intercultural pragmatics (Kecskes, 2018). After peer review and teacher commentary, these collocations were added to a shared class list, which was later used to design quiz items and speaking prompts. This approach operationalized recommendations to target high-value, problematic collocations and recycle them over time, while leveraging GenAI to generate candidate items and explanations efficiently (Durrant & Schmitt, 2009; Nordlund & Norberg, 2025).

Activity 3: Practicing academic multiword sequences

The third set of activities targeted academic MWSs and lexical bundles relevant to discussion and presentation tasks, building on research linking such bundles and pragmatic markers to listener-based fluency judgments (Brown et al., 2023; Hougham et al., 2023; Tavakoli & Wright, 2020). In other words, students focused on the kinds of expressions that help speakers clearly signal opinions, contrast ideas, and organize their turns in academic talk. In preparation for small-group discussions and short presentations, students prompted a GenAI chatbot to "provide 10 frequently used multiword expressions for expressing opinions and agreeing or disagreeing in academic discussions at the CEFR B2 level, with short explanations and example sentences." The class briefly examined the list, and the teacher highlighted items that corresponded to stance and discourse-organizing bundles identified in fluency research (e.g., I would argue that..., On the other hand,..., In other words,...).

Students then selected several target expressions they wanted to practice and wrote short discussion scripts in pairs, with the instruction to keep content simple but to use each target MWS multiple times in slightly different ways. In their scripts, students embedded stance and discourse-organizing bundles generated by the LLM, producing exchanges such as:

Student A: *I would argue that* eating at the campus cafeteria is the best option for lunch because it is cheap and convenient.

Student B: *That's true*, but *on the other hand*, it gets extremely crowded during the noon break. *From another perspective*, bringing a *bento* from home saves even more money.

Student A: *I see your point*, but *I would also argue that* eating at the cafeteria lets us socialize with friends. *On the other hand*, if we want a quiet lunch, eating near the station is much better.

After rehearsing with partners, they participated in live discussions without reading their scripts (although note cards listing the new target expressions were allowed), while classmates listened for and tallied occurrences of the target expressions. This task design drew on fluency-oriented research emphasizing the role of structured repetition in automatizing target chunks (Tavakoli & Uchihara, 2020), while leveraging the fluency benefits inherent in interactive, dialogic task performance (Tavakoli, 2016). Following the discussions, students returned to the chatbot to request alternative expressions and more formal/informal variants and compared AI suggestions with teacher judgments about naturalness and appropriacy. This cycle aimed to expand learners' repertoires of academic MWSs, strengthen links between formulaic language and perceived fluency, and promote a more critical, reflective stance toward GenAI feedback, as well as giving students a chance to *enjoy* using MWSs in real communication. Together, these activities positioned GenAI as a source of examples and prompts, with students and the teacher responsible for selecting, adapting, and using them.

Discussion

Benefits, Drawbacks, and Challenges of Using GenAI for MWS Instruction

Across these activities, several perceived benefits, drawbacks, and challenges of using GenAI to support MWS instruction emerged. However, it is important to mention up front that, since this is an exploratory practice-oriented pedagogical classroom report, the benefits and risks discussed here reflect teacher observations and students' comments, not results from any formal pre-/post-tests or controlled comparisons.

As for the benefits, many students reported that working with chatbots made idioms, collocations, and academic expressions feel more accessible and interesting than when they encountered them only in textbooks or word lists. This echoes claims that enriched, contextualized input can facilitate formulaic language learning (Paquot, 2019). GenAI also allowed rapid generation of level-appropriate examples and variations, which supported repeated encounters, a condition associated with durable gains in vocabulary (Durrant & Schmitt, 2009; Garner & Crossley, 2018). In the academic-discussion activities, the ability to request alternative phrasings and more formal versions appeared to help students expand their repertoires of lexical bundles and pragmatic markers in ways that align with recent work on their role in perceived fluency (Brown et al., 2023). For these learners, GenAI functioned less as a phrase generator and more as a catalyst for noticing, recycling, and personalizing language in ways that supported more confident, fluent-sounding speech.

At the same time, several drawbacks and risks became apparent, many of which resonate with current critiques of GenAI in education. Some students were tempted to copy AI-generated sentences directly into their dialogues and scripts without sufficient adaptation, raising concerns about over-reliance on AI phrasing and reduced opportunities for genuine language production (Heaven, 2023; Rudolph et al., 2023). The chatbots also occasionally produced infrequent, awkward, or culturally odd MWSs, especially when prompts were vague or under-specified, confirming warnings about hallucinations and the need for careful human evaluation of AI output (Hougham & Higginbotham, 2024; Kohnke et al., 2023). Questions about academic integrity surfaced when students realized how easily GenAI could generate complete essays and scripts, underscoring the importance of transparent classroom policies and explicit discussion of acceptable uses (Sims, 2023; UNESCO, 2023; US Department of Education, 2023). These issues suggest that GenAI-mediated MWS work *must* be coupled with explicit conversations about ethics and authorship, not only about language.

Implementing GenAI-mediated activities also involved practical challenges. Not all students had equal access to devices with similar processing speeds, and some were initially unfamiliar with English-medium interfaces, which required additional support and sometimes slowed the pace of lessons. Designing prompts that consistently elicited frequent, CEFR-appropriate MWSs took time and iterative experimentation, reflecting broader concerns that effective AI-supported instruction demands new forms of teacher expertise (Hougham & Higginbotham, 2024; Kohnke et al., 2023). Balancing time spent on AI-based exploration with other course components required conscious scheduling decisions to prevent technology from dominating class time, a concern mirrored in recent discussions of GenAI's place within broader curricula (Hunte & Wheatley-Holness, 2025; Teng, 2023). Nevertheless, the overall experience suggested that, when framed by clear guidelines and grounded in MWS and fluency research, GenAI can be integrated successfully into existing courses in ways that *support* rather than replace human interaction and teacher-led explanation.

Conclusion

Final Thoughts, Pedagogical Implications, and Further Research

This paper has explored how GenAI chatbots can be practically leveraged to bridge the gap between theoretical multiword sequence research and classroom practice. While literature consistently underscores the fact that formulaic language is central to developing oral fluency, high-utility expressions remain underrepresented in traditional EFL classrooms, leaving many Japanese university students with strong individual vocabulary knowledge but limited control over cohesive multiword chunks. The classroom activities reported here attempted to address this pedagogical imbalance. By using GenAI applications not as an unthinking shortcut, but as responsive tools for generating localized, contextualized examples and prompting metacognitive reflection, this design successfully kept *human* judgment and collaborative *interaction* at the center of the learning process.

Based on qualitative classroom observations during the course, GenAI-mediated tasks targeting idioms, incongruent collocations, and academic MWSs helped heighten students' awareness of formulaic language, broaden their range of practice, and bridge research-based fluency principles with daily instruction. At the same time, the experience highlighted the need for explicit guidelines, critical evaluation of AI output, and assessment practices that reward *students' own productive and active use* of MWSs rather than their ability to simply copy AI-generated text in an attempt to pass it off as their own original work (Heaven, 2023; UNESCO, 2023). In practical terms, this means valuing how learners use MWSs in speaking and writing, not how well they can reproduce what a chatbot has written for them.

For teachers interested in experimenting with similar approaches, three tentative implications emerge. First, treat MWSs as core instructional targets by selecting a manageable set of highvalue expressions and recycling them across tasks and assessments. Second, design GenAI prompts that foreground frequent, levelappropriate MWSs in context and build in followup activities that require personalization, comparison, and adaptation of AI suggestions (see Appendix for sample prompt templates). Third, incorporate reflection and transparency about GenAI use – for example, by asking students to document how they use chatbots and to justify the expressions they adopt – to help them develop both formulaic language knowledge and critical AI literacy (Hunte & WheatleyHolness, 2025; Teng, 2023; see Appendix for sample reflection protocols). Grounding the practice in this type of self-evaluation aligns with how Peltonen (2024) advocates for active learner self-reflection, where students are guided to analyze their own unique speaking behaviors to establish individual communicative goals. Taken together, these steps position GenAI as a structured partner in MWS learning rather than an unchecked shortcut or addon outside the instructor's oversight.

Future work could build on this exploratory classroom report by systematically tracking changes in learners' MWS use and perceived fluency over time within GenAI-mediated frameworks, drawing on analytic tools and rating approaches developed in recent vocabulary and fluency research (Garner & Crossley, 2018; Peltonen, 2024; Tavakoli & Wright, 2020). Ultimately, the aim is to keep experimenting, sharing, and refining classroom activities so that work on teaching multiword sequences keeps pace with emerging technologies and leads to more communicative confidence for learners. By adapting such practices, teachers can help ensure that vocabulary instruction remains both research-informed and responsive to the evolving realities of AI-assisted language learning. In other words, rather than treating MWS structures as an optional curricular extra, integrating structured GenAI tasks allows practitioners to turn these formulaic sequences into accessible, interactive building blocks for everyday learner communication. If teachers, learners, and researchers continue to work together in this spirit, the next generation of classrooms can treat multiword sequences as a shared resource for building richer, more fluent L2 *human* voices.

Bio Data

Christopher J. Hollis is a junior associate professor of English at Tottori University, and a PhD student in applied linguistics at Hiroshima University. He holds an M.A. in TESL from the University of Illinois (USA). Chris has taught in Japan for over 17 years. His research interests include second language (L2) vocabulary acquisition, productive vocabulary development, assessment of vocabulary knowledge and fluency, computer-assisted language learning (CALL), artificial intelligence (A.I.), educational technology, diversity and inclusion, study abroad, international exchange, and cross-cultural communication. <Hollis@tottori-u.ac.jp>

References

Asari, Y. (2021). *The use of formulaic sequences and perceived fluency*. In Selected Papers from the 25th PAAL International Conference (pp. 1-23). Pan-Pacific Association of Applied Linguistics.

Brown, D., Fernández, J., & Huensch, A. (2023). The role of pragmatic markers in perceptions of L2 fluency in dialogue. *System, 119*, 103157. https://doi.org/10.1016/j.system.2023.103157

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching, 47*(2), 157-177. https://doi.org/10.1515/iral.2009.007

Garner, J., & Crossley, S. A. (2018). A latent curve model approach to studying L2 n-gram development. *The Modern Language Journal, 102*(3), 494-511. https://doi.org/10.1111/modl.12494

Google. (2025). *Gemini* (version 2.5 Flash) [Large language model]. https://gemini.google.com/app

Heaven, W. D. (2023, April 6). ChatGPT is going to change education, not destroy it. *MIT Technology Review*. https://www.technologyreview.com/2023/04/06/1071059/chatgpt-change-not-destroy-education-openai/

Hougham, D., Clenton, J., Uchihara, T., & Higginbotham, G. (2023). The impact of lexical bundle length on L2 oral proficiency. *Languages, 8*(4), 232. https://doi.org/10.3390/languages9070232

Hougham, D., & Higginbotham, G. (2024). Generative AI in writing classes: Seven golden rules. *The Language Teacher, 48*(2), 14-17. https://doi.org/10.37546/JALTTLT48.2

Hunte, A. A., & Wheatley-Holness, J. (2025). Enhancing quality instruction and learning experiences in higher education through ChatGPT: An exploratory case study. *Caribbean Journal of Multidisciplinary Studies*, *4*(1), 57-86.

Kecskes, I. (2018). Intercultural pragmatics. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1-7). Wiley.

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

Koya, T. (2003). A study of collocation in English and Japanese noun-verb combinations. *Intercultural Communication Studies, 12*(1), 125-145. https://media.sciltp.com/articles/sciltp/ics/2003/08-Taeko-Koya.pdf

Nakata, T. (2025). Teaching and learning multi-word items in English. *Annual Report on General Education, Tottori University, 30*, 36-50.

Nakata, T., & Suzuki, Y. (2022). *Eigo gakushuu no kagaku* [The science of English learning]. Kenkyusha.

Nakata, T., Tada, S., McLean, S., & Kim, Y. A. (2020). Effects of distributed retrieval practice over a semester: Cumulative tests as a way to facilitate second language vocabulary learning. *TESOL Quarterly, 54*(1), 29-55. https://doi.org/10.1002/tesq.596

Nordlund, M., & Norberg, C. (2025). Formulaic language in EFL teaching materials. *Education Inquiry*. Advance online publication. https://doi.org/10.1080/20004508.2025.2593701

OpenAI. (2025). *ChatGPT* (GPT-4o version) [Large language model]. https://chatgpt.com

Paquot, M. (2019). Formulaic language and learner language. *Second Language Research, 35*(2), 121-145. https://doi.org/10.1017/S0267190512000098

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). Longman.

Peltonen, P. (2024). Fluency revisited. *ELT Journal, 78*(4), 489-492. https://doi.org/10.1093/elt/ccad047

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342-362. https://scispace.com/pdf/chatgpt-bullshit-spewer-or-the-end-of-traditional-3e0b1z03.pdf

Sims, A. (2023, March 9). ChatGPT and the future of university assessment. *Times Higher Education*. https://www.timeshighereducation.com/campus/chatgpt-and-future-university-assessment

Takizawa, K., & Suzuki, Y. (2025). The role of multiword sequences in fluent speech: The case of listener-based judgment in L2 argumentative speech. *Languages*, *10*(1), 23. https://doi.org/10.1017/S0272263125000051

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance. *Language Learning, 66*(4), 735-766. https://doi.org/10.1515/iral-2016-9994

Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning, 70*(2), 506-547. https://doi.org/10.1111/lang.12384

Tavakoli, P., & Wright, C. (2020). *Second language speech fluency: From research to practice*. Cambridge University Press. https://doi.org/10.1017/9781108589109

Teng, M. F. (2023). Scientific writing, reviewing, and editing for open-access TESOL journals: The role of ChatGPT. *International Journal of TESOL Studies, 5*(4), 87-91. https://doi.org/10.58304/ijts.20230107

Uchihara, T., & Clenton, J. (2020). Dimensions of vocabulary knowledge and their relationships to second language speaking proficiency. *Journal of English for Academic Purposes, 48*, 100925.

UNESCO. (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. UNESCO Publishing. https://unesdoc.unesco.org/ark:/48223/pf0000385146

US Department of Education. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. Office of Educational Technology. https://www2.ed.gov/documents/ai-report/ai-report.pdf

Vercellotti, M. L. (2021). Multiword sequences in L2 English language learners' speech: The relationship between trigrams and lexical variety across development. *System, 99*, 102525. https://doi.org/10.1016/j.system.2021.102494

Wood, D. (2010). *Formulaic language and second language speech fluency*. Continuum.

Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. Bloomsbury Academic.

Wood, D. (2020). *Idioms in contemporary English (2nd ed.).* Bloomsbury Academic.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519772

Appendix

Glossary of Key Terms, Instructional Materials, and Reflective Protocols

This appendix provides definitions of key formulaic language terminology, sample prompt templates, and student reflection protocols used in the GenAI-mediated activities described in this paper.

Section A: Glossary of Key Terms

To assist readers unfamiliar with phraseological terminology, Table A1 summarizes the primary categories of multiword sequences targeted in this course.

Table A1

Definitions and Examples of Target Multiword Sequence Categories

Term	Definition	Classroom / Corpus Example
Multiword Sequences (MWSs)	Overarching term for combinations of two or more words that frequently co-occur in natural language.	*at the end of the day*, *make a decision*
Idioms	Expressions whose figurative meaning cannot be determined from the literal meanings of individual component words.	*hit the books*, *break the ice*, *spill the beans*
Collocations	Pairs or groups of words that naturally co-occur with high statistical frequency.	*make a copy* (not *take a copy*), *deep voice*
Incongruent Collocations	Collocations whose English structure differs directly from literal word-for-word translation in the learner's L1 (e.g., Japanese).	Japanese:「コピーを取る」 English: *make a copy*
Phrasal Verbs	Multi-word verbs composed of a main verb plus one or more particles (prepositions or adverbs).	*carry out*, *turn down*, *look forward to*
Lexical Bundles	Recurrent, multi-word frequency strings that fulfill specific discourse or interactional functions in academic speech/writing.	*on the other hand*, *in order to*, *I would argue that*

Note: Definitions adapted from Paquot (2019) and Wood (2015).

Section B: Sample Prompt Templates for GenAI Activities

1. Idiom Awareness Activity

"Write 2 frequently occurring idioms about [insert theme, e.g., travel/university life] with simple meaning explanations and example sentences at the [insert target CEFR level, e.g., CEFR B2 level or lower]. Please italicize the idioms."

2. Incongruent Collocations Activity

"Please write 2 frequently occurring collocations that differ between English and Japanese about [insert theme, e.g., daily school life] with simple explanations of the differences and example sentences at the [insert target CEFR level, e.g., CEFR B2 level or lower]. Please italicize the collocations."

3. Academic Multiword Sequences Activity

"Please provide 10 frequently used multiword expressions for [insert function, e.g., expressing opinions/agreeing/disagreeing] in academic discussions at the [insert target CEFR level, e.g., CEFR B2 level or lower], with short explanations and example sentences. Please italicize the academic multiword sequences"

Section C: Student Reflective Protocol for GenAI-Mediated Writing

To foster critical AI literacy, students documented their interaction process using the following log.

GenAI Reflection Log

- Target Idiom/Collocation/MWS: _____
- Prompt used: _____
- AI Output received: _____

Analysis & Justification:

- Verification: Did the AI output sound natural? Why or why not? (Check against your intuition or peer feedback).
- Adaptation: What changes did you make to the AI's suggestion? (e.g., changed the verb, adjusted the formality, personalized the context).
- Justification: Why did you decide to keep or discard the AI's version in your final script?
  - *I kept it because...* _____
  - *I changed/discarded it because...* _____