

Proceedings of the 2026 Teacher Development Symposium

Nagoya University of Foreign Studies
Center for Language Education and Development

Edited by
Berke Alp
Misa Fujinuma

CONTENTS

Introduction	3
Using Values-Rich Stories to Enhance Language Learning in the EFL Classroom	5
<i>Ginny DEL ROSARIO</i>	
Teaching AI Literacy Through Business Writing: A Practice Report on AI-Assisted L2 Writing in a Japanese University	24
<i>Misa FUJINUMA</i>	
Diagnosing and Measuring Pragmatic Development through Corpus-Informed Instruction: A Classroom Study of English Apologies	44
<i>Taylor MEIZLISH and Christopher ZIFFO</i>	
The Relationship Between GenAI-Writerism and English Writing Self-Efficacy in a Japanese University Context: A Pilot Study	60
<i>Naoya SHIBATA and Vivian TURK</i>	
Teaching the Room: A Qualitative Inquiry into Seating and Grouping in the Language Classroom	83
<i>Vivian TURK</i>	

Introduction

We are pleased to announce the 2026 NUFS Teacher Development Symposium Proceedings. This edition features five insightful and timely articles derived from presentations at the fifth Teacher Development Symposium, hosted by the Centre for Language Education and Development (CLED) at Nagoya University of Foreign Studies (NUFS) on January 24th, 2026. The symposium provided a valuable platform for educators and researchers from all over the world to exchange knowledge, share experiences, and discuss activities relevant to teaching English as a foreign language, as well as to engage in discussions on theoretical and practical issues in language education.

The first article, by Ginny Del Rosario, presents a literature-based instructional framework that uses values-rich stories, such as *The Eagle and the Chicken* and Bo Sanchez's *How Many Cows Are You Worth?*, to support meaningful language use and reflection in English language teaching. Misa Fujinuma presents a study in the second article where generative AI tools were deliberately integrated into an advanced, English-medium business writing course, using a transformative learning and self-authorship framework to make students' AI use visible and open to discussion throughout the writing process. In the third article, Taylor Meizlish and Chris Ziffo use a corpus comparison to identify gaps in Japanese university students' English apology strategies and evaluate a targeted instructional intervention to address them. Naoya Shibata and Vivian Turk introduce the concept of GenAI-Writerism in the fourth article, exploring how learners position generative AI output as a source of linguistic authority and examining its correlation with L2 writing self-efficacy. Finally, Vivian Turk examines how seating and grouping arrangements in first-year Core English classes shape students' participation, belonging, and perceived learning.

The Teacher Development Symposium Committee extends sincere gratitude to the contributors who submitted their manuscripts and navigated the double peer review process, as well as to the reviewers who provided constructive and insightful feedback. Additionally, the call for papers for the sixth NUFS Teacher Development Symposium, scheduled for January 23rd, 2027, has been announced on our website. We welcome proposals for Short Research Presentations and Practice-oriented Presentations or Workshops.

Editors

Berke ALP

Misa FUJINUMA

**The 2026 Teacher Development
Symposium Committee**

Jason WALTERS (Chair)

Misa FUJINUMA

Gamze KELLE

Marissa ODA

Vivian TURK

Christopher ZIFFO

Using Values-Rich Stories to Enhance Language Learning in the EFL Classroom

Ginny DEL ROSARIO

Immaculate Conception Academy

Abstract

This practice article presents a literature-based instructional framework that uses values-rich stories to support meaningful language use and reflection in English language teaching. To illustrate the framework, the article presents an exemplar lesson based on the stories *The Eagle and the Chicken* and Bo Sanchez's *How Many Cows Are You Worth?* Drawing on Bandura's Social Cognitive Theory, literature-based language instruction, and storytelling pedagogy, the framework integrates modeling, guided discussion, and scaffolded language activities to promote learner engagement and personal reflection.

The framework follows a five-stage sequence: Motivation, Vocabulary, Storytelling, Comprehension, and Reflection. Through this sequence, literary texts are used as vehicles for vocabulary development, discussion, comprehension, and exploration of themes such as identity, self-worth, resilience, and personal growth. Sample classroom activities and teacher reflections are included to demonstrate how the framework may be adapted for use in English language classrooms. The article suggests that values-rich storytelling can provide a meaningful context for language learning while encouraging learners to connect language use with personal experience and reflective thinking.

Keywords: literature-based instruction, storytelling, values education, English language teaching, Social Cognitive Theory

Using Values-Rich Stories to Enhance Language Learning in the EFL Classroom

Language teachers continually seek instructional approaches that promote both language development and meaningful learner engagement. While grammar exercises, vocabulary practice, and textbook-based activities remain important components of language instruction, language is often learned most effectively when learners use it for authentic communication and meaningful interaction. Contemporary language teaching therefore, emphasizes communicative competence, learner participation, and opportunities for learners to connect classroom learning with their own experiences.

One approach that supports these goals is storytelling. Stories provide learners with meaningful and contextualized language input while creating opportunities for discussion, interpretation, and personal response. Because stories engage learners intellectually and emotionally, they can encourage participation, stimulate reflection, and provide meaningful reasons for language use. Literature-based approaches similarly recognize the value of authentic texts in promoting language development, critical thinking, and learner engagement.

Values-rich stories are particularly useful because they combine language learning with opportunities to explore important human experiences and personal values. Stories that address themes such as identity, perseverance, courage, and self-worth invite learners to move beyond literal comprehension and engage in discussion, reflection, and personal meaning-making. Such experiences can help create language-learning environments in which communication is purposeful and personally relevant.

This article presents a literature-based instructional framework that uses values-rich storytelling to support meaningful language use and reflection in English language teaching. To illustrate the framework, the article presents an exemplar lesson based on *The Eagle* and the Chicken and Bo Sanchez's (2007) *How Many Cows Are You Worth?* Both stories explore themes of identity, self-belief, personal potential, and self-worth, providing opportunities for language learning as well as reflective discussion.

The framework follows a five-stage sequence: Motivation, Vocabulary, Storytelling, Comprehension, and Reflection. Designed to support meaningful language use and learner engagement, the framework combines storytelling, discussion, and reflection within a structured lesson sequence. Developed for NUFs's teacher-development symposium, it is presented as an instructional exemplar rather than as a report of classroom implementation and is intended as a flexible guide that teachers may adapt to different learners, stories, and educational contexts.

The purpose of this article is to illustrate how values-rich storytelling may be integrated into language instruction in a structured and meaningful way. By combining language learning with opportunities for discussion and reflection, teachers may create classroom experiences that support communicative development while encouraging learners to engage thoughtfully with important human values and experiences.

Theoretical Background

This article draws on three complementary perspectives: Bandura's Social Cognitive Theory, literature-based language instruction, and storytelling pedagogy. Together, these perspectives help explain how stories can support meaningful language learning while also encouraging reflection on personal values and experiences.

Bandura's Social Cognitive Theory provides a psychological explanation for how learners develop beliefs, attitudes, and behaviors through observation and reflection. Storytelling provides learners with models in the form of characters, situations, and narratives that illustrate human actions and their consequences. Literature-based language instruction, in turn, provides the pedagogical framework through which learners engage with stories, discuss their meanings, and use language in authentic and purposeful ways. Taken together, these perspectives suggest that stories can serve not only as sources of language input but also as opportunities for learners to reflect, communicate, and construct meaning.

Social Cognitive Theory

Bandura's Social Cognitive Theory emphasizes the importance of observational learning, modeling, and self-efficacy in human development. According to Bandura (2001), individuals learn not only through direct experience but also by observing others and reflecting on their actions and outcomes. In educational settings, such observations can influence learners' beliefs about their own capabilities and potential for success.

Applied to language learning, the theory suggests that learners benefit from engaging with meaningful experiences and observing models that demonstrate attitudes, behaviors, and problem-solving strategies. Stories provide rich opportunities for such learning because they present characters who face challenges, make decisions, and experience consequences. Through these narratives, learners may identify with characters, reflect on their choices, and consider how similar situations relate to their own lives.

Values-rich stories are particularly relevant from a social cognitive perspective because they often portray perseverance, courage, self-belief, and personal growth. By discussing and reflecting on these experiences, learners may develop greater awareness of

their own abilities and aspirations. Such reflection aligns with Bandura's concept of self-efficacy, which plays an important role in motivation, persistence, and performance.

Literature-Based Language Instruction. Literature-based language instruction emphasizes the use of authentic literary texts as meaningful contexts for language learning. Rather than treating language as an isolated system of grammar rules and vocabulary lists, this approach places language within stories, experiences, and human interactions that encourage learners to engage actively with meaning.

Literary texts expose learners to authentic language and encourage interpretation, discussion, and personal reflection. Collie and Slater (1987) argue that literature provides valuable opportunities for learners to explore language through imagination and cultural understanding. Similarly, Lazar (1993) emphasizes that literary texts provide authentic language input while stimulating learners' imagination and encouraging deeper engagement with characters, themes, and moral lessons. Wright (1995) likewise notes that stories help learners internalize vocabulary and language structures while fostering personal involvement in the learning process.

More recent scholarship continues to support the value of literature in language education. Mardiani and Baharuddin (2023) argue that literature-based instruction contributes not only to language skills development but also to learner motivation and cultural understanding. Similarly, Hossain (2024) notes that literature-based language learning can enhance linguistic proficiency, cultural awareness, and critical thinking while engaging learners with authentic literary texts. These findings suggest that literature is a valuable resource for contemporary language classrooms.

Research has also shown that literature-based approaches can positively influence students' attitudes toward reading and learning. Tunnell and Jacobs (1989), for example, reported positive outcomes associated with literature-based instruction for a wide range of learners, including struggling readers and language learners. Darmawati et al. (2020) further observed that literature-based instruction can support language development while enhancing learners' confidence in reading, writing, listening, and speaking.

Because literary texts present language in meaningful and authentic contexts, literature-based instruction provides opportunities for vocabulary development, reading comprehension, communicative competence, and critical thinking. At the same time, it encourages learners to connect language learning with personal experiences, making classroom learning both linguistically and personally meaningful.

Storytelling and Language Learning. Storytelling has long been recognized as a powerful pedagogical tool in language education. Stories naturally capture learners' attention and provide meaningful contexts in which language can be understood and used. Unlike isolated language exercises, stories present vocabulary, grammar, and discourse structures within coherent narratives, helping learners understand how language functions in real communication.

Storytelling supports language learning in several ways. First, it promotes comprehension by providing contextualized language input. Second, it helps learners remember vocabulary and concepts through narrative structure and emotional engagement. Third, it creates opportunities for learners to discuss ideas, express opinions, and respond to characters and events using the target language.

Recent research likewise highlights the role of storytelling in fostering learner engagement and meaningful participation. Rahmawati et al. (2025) found that storytelling can enhance learners' confidence, engagement, and willingness to participate in English language activities. Their findings suggest that storytelling provides meaningful opportunities for language use while creating a supportive environment for communication and self-expression.

Storytelling also contributes to learners' personal development. Educational researchers have noted that narratives help learners interpret experiences and construct meaning about themselves and the world around them. Values-rich stories, such as fables and moral narratives, are particularly effective because they present characters who face challenges, make decisions, and experience consequences. Through these stories, learners are encouraged not only to understand language but also to reflect on human values and social realities.

From the perspective of Social Cognitive Theory, storytelling serves as a vehicle for observational learning. Characters function as models whose actions and experiences invite reflection and discussion. Through literature-based instruction, these narratives become opportunities for learners to engage in meaningful communication while exploring themes that resonate with their own lives.

Consequently, literature-based instruction that incorporates storytelling can serve as a powerful pedagogical strategy. By integrating language learning with reflection on human values and experiences, teachers can create classroom experiences that are both linguistically rich and personally meaningful. This theoretical foundation underpins the values-rich storytelling framework presented in this article.

Instructional Context. This article presents an illustrative lesson designed to demonstrate how values-rich stories may be incorporated into English language instruction. The lesson was adapted from classroom materials used by the author and was developed for presentation at NUFS's teacher-development symposium focused on English language teaching. To align with the symposium theme, the stories and activities were modified to illustrate how the framework might be applied in an EFL context. The lesson is therefore intended as an instructional exemplar rather than a report of a specific classroom implementation. Its principles may be adapted to a variety of English language teaching contexts and learner populations.

The Values-Rich Storytelling Framework. This article proposes a simple instructional framework for designing literature-based language lessons using values-rich stories. The framework follows a five-stage sequence:

Motivation → Vocabulary → Storytelling → Comprehension → Reflection

The sequence is not intended to replace existing language-teaching frameworks. Rather, it builds upon the traditional pre-reading, while-reading, and post-reading sequence commonly used in language classrooms while extending it through the explicit inclusion of a Reflection stage and the use of a secondary application story that promotes personal meaning-making and values formation. The framework is designed as a flexible guide that teachers may adapt according to their learners' needs, proficiency levels, instructional goals, and chosen texts.

Motivation

The lesson begins with a motivation activity that activates learners' prior knowledge and stimulates interest in the topic. This stage may involve quotations, photographs, short scenarios, reflective questions, or brief sharing activities. By connecting the lesson to learners' experiences, the motivation stage helps establish a meaningful context for subsequent language learning activities.

Vocabulary

Before the story is presented, selected vocabulary items are introduced to support comprehension. This stage familiarizes learners with key words and expressions that may be important for understanding the narrative. Activities may include matching exercises, visual aids, contextualized sentence practice, or pronunciation work. Pre-teaching selected vocabulary reduces cognitive load and allows learners to focus more fully on the meaning of the story.

Storytelling

The story is then presented through oral storytelling, reading, or a combination of both. During this stage, learners are exposed to contextualized language within an engaging narrative. Teachers may use expressive reading, voice modulation, gestures, visual supports, and prediction questions to sustain attention and encourage active listening. The storytelling stage serves as the primary source of linguistic input while also introducing the themes that will be explored throughout the lesson.

Comprehension

Following the storytelling stage, learners demonstrate understanding of the narrative through comprehension activities. These may include discussion questions, written responses, graphic organizers, role-plays, collaborative tasks, or story retellings. In addition to checking understanding, this stage provides opportunities for learners to use language for communication and meaning-making.

Reflection

The Reflection stage represents the distinctive feature of the framework. While traditional post-reading activities often focus primarily on comprehension and language practice, this stage encourages learners to connect the story's themes to their own experiences, beliefs, and aspirations. Reflection activities may involve guided discussion, journal writing, personal response questions, or the introduction of a secondary application text that reinforces the lesson's central message.

Among these activities, the use of a secondary application story plays a particularly important role in extending reflection and discussion. While the main story introduces the central theme, the secondary story allows learners to revisit that theme in a different context and from a different perspective. This transfer of ideas encourages learners to move beyond comprehension of a single narrative and consider how broader concepts such as identity, resilience, and self-worth may be expressed across different situations. By comparing and connecting the two stories, learners are provided with additional opportunities for discussion, interpretation, and personal reflection.

Through reflection, learners move beyond understanding what happened in the story to considering why it matters and how it relates to their own lives. This process encourages meaningful language use while supporting personal growth, self-awareness, and deeper engagement with the themes explored in the text.

Because the framework is flexible, it can be adapted to a wide range of literary texts, learner profiles, proficiency levels, and educational contexts. Teachers may modify individual

stages, expand particular activities, or select alternative stories while maintaining the overall structure of the framework.

Lesson Design and Teaching Procedure

Lesson Overview

Grade Level: Grade 6

Subject: English Language

Main Story: The Eagle and the Chicken

Application Story: How Many Cows Are You Worth?

Contextual Note

The lesson presented in this article reflects the instructional sequence and sample activities used in the teacher-development presentation from which this paper is derived. The vocabulary items, discussion questions, and classroom tasks are included as illustrative examples of how the framework may be implemented in practice. As with any instructional model, teachers are encouraged to adapt materials, vocabulary selection, and learning activities according to their learners' proficiency levels, instructional objectives, and classroom contexts.

Lesson Objectives

Students will be able to:

- understand key vocabulary related to the story;
- demonstrate comprehension of narrative events and themes;
- express opinions and personal reflections in English;
- use selected vocabulary in oral and written responses; and
- reflect on themes of identity, resilience, and self-worth.

To provide a clearer view of how the framework may be applied in practice, Table 1 presents an illustrative lesson procedure. The table outlines the sequence of instructional stages, sample teacher actions, student participation, and intended learning outcomes. This example demonstrates how the Motivation–Vocabulary–Storytelling–Comprehension–Reflection framework may be adapted for use in English language teaching contexts.

Table 1

Sample Lesson Procedure Using the Values-Rich Storytelling Framework

Lesson Stage	Teacher Activities	Student Activities	Learning Objectives
Motivation	The teacher begins by presenting the quote, “ <i>You become what you believe,</i> ” and invites students to think about its meaning. The teacher briefly explains that our beliefs about ourselves can influence what we are able to do and achieve. Students are then asked reflective questions such as: “ <i>Have you ever felt that you could do something better if you believed in yourself?</i> ” This helps connect the idea of self-belief to their own experiences and prepares them for the theme of the story.	Students respond to the questions and share their ideas about the meaning of the quote, as well as their thoughts on abilities and personal goals. They may also describe situations in which they felt unsure about their own abilities.	Activate prior knowledge, stimulate curiosity, and encourage students to express initial thoughts in English.
Vocabulary	The teacher introduces key vocabulary words from the story, including <i>eagle, chicken, naturalist, wings, stretch,</i> and <i>soar</i> . Visual aids or pictures may be used to support comprehension. The teacher models pronunciation and asks students to repeat the words.	Students repeat the vocabulary words, match them with definitions, and create simple sentences using the new vocabulary. Pair work activities allow students to practice using the words in short conversations.	Develop vocabulary knowledge and prepare learners for understanding the story.
Storytelling	The teacher narrates the story “The Eagle and the Chicken,” using expressive voice, gestures, and pauses to highlight important events. The teacher occasionally asks prediction questions such as “What do you think will happen next?” to maintain engagement.	Students listen carefully to the story and visualize the events. Some students may take brief notes about important details. They respond to prediction questions and express their expectations about the story.	Improve listening comprehension and expose students to contextualized language.
Comprehension	After the storytelling stage, the teacher asks comprehension questions to check understanding. The teacher encourages students to explain the events of the story in their own words and to discuss the main message.	Students work in small groups to answer comprehension questions and summarize the story. Group representatives share their responses with the class.	Reinforce understanding of the narrative and develop speaking skills.
Reflection	The teacher introduces the second story, “How Many Cows Are You Worth?”, and guides students in discussing how society sometimes measures people’s value based on appearance or external standards. The teacher encourages students to reflect on their own potential.	Students participate in reflective discussions and share their personal interpretations of the stories. Some students write short reflections about the lesson they learned.	

The illustrative lesson procedure shown in Table 1 demonstrates one possible application of the values-rich storytelling framework. Each stage is designed to build upon the previous one, guiding learners from initial engagement with the topic toward deeper linguistic and conceptual exploration. By moving from motivation to reflection, the framework situates language learning within a meaningful narrative context rather than relying solely on isolated language exercises. The sequence is intended to support vocabulary development, listening and speaking practice, comprehension, and reflective discussion while encouraging learners to engage with themes of identity, self-belief, and personal growth.

Vocabulary Activities

Before reading the story, key vocabulary words are introduced. The vocabulary examples presented below were selected for the original lesson demonstration and are intended to illustrate the vocabulary stage of the framework. Teachers may modify the vocabulary list, activities, and level of linguistic challenge according to their learners' age, proficiency level, and instructional goals.

Example vocabulary list:

- eagle
- chicken
- naturalist
- wings
- stretch
- soar

Matching Activity

Match the words with their meanings.

1. common, ordinary farm bird
2. body part used by birds for flying
3. a big, powerful bird
4. a person who studies plants, animals, and the natural environment
5. the act of reaching out to lengthen muscles to their full capacity
6. to rise to a higher level

Students write the correct definitions in their notebooks.

Storytelling Stage

The teacher reads the story “The Eagle and the Chicken” aloud. Teacher modeling includes expressive reading, gestures, and pauses to maintain student engagement. Students listen carefully and visualize the events.

Comprehension Activities

Activity A: Mini Role-Play (Pairs) - One student is the naturalist; the other is the eagle. They act out one or two lines.

Activity B: Guide Questions (Students may discuss answers in small groups.)

1. Why did the eagle behave like a chicken?
2. What helped him remember who he really was?
3. What does the eagle symbolize?

Possible responses may include:

- The eagle behaved like a chicken because he was raised among chickens and believed he was one of them.
- The naturalist helped the eagle recognize his true nature and potential.
- The eagle symbolizes individuals who underestimate their abilities or fail to recognize their potential.

Reflection

1. Have you ever felt like the eagle - holding yourself back?
2. What limits you to “spread your wings”?
3. Who are the people who can help? Who can be your “naturalist”?

The second story, “How Many Cows Are You Worth?” is then introduced. This story reinforces the theme of self-worth. Students discuss how people sometimes judge others based on external standards.

Example reflection question:

“Should people measure their value based on what others say?”

Classroom Implementation and Student Interaction

Facilitating Storytelling, Discussion, and Reflection

The storytelling stage represents a central component of the framework because it provides learners with meaningful and contextualized language input while fostering engagement with the narrative. During this stage, the teacher serves not only as a narrator but also as a facilitator of learning. Effective storytelling may involve the use of voice modulation, gestures, facial expressions, and strategic pauses to sustain learners' attention and highlight significant events in the story.

In an illustrative application of the framework, the teacher may present *The Eagle and the Chicken* through expressive narration, emphasizing moments that reveal the story's central themes of identity, self-belief, and personal potential. At selected points, the teacher may pause to ask prediction or reflective questions such as, “*What do you think the eagle is*

feeling at this moment?” These brief interactions encourage learners to engage actively with the narrative rather than simply listening passively.

Teachers may also encourage learners to visualize the story by imagining the settings, characters, and events being described. For example, learners may be asked to consider how the eagle's environment influenced its beliefs about itself. Visualization activities can support comprehension while encouraging deeper engagement with the narrative.

Following the storytelling stage, learners may participate in small-group discussions focused on comprehension and interpretation. Collaborative learning provides opportunities for learners to share ideas, clarify understanding, and explore different perspectives. In language-learning contexts, small-group discussions may also encourage participation by providing a supportive environment for communication.

The teacher may subsequently facilitate a whole-class discussion in which groups share their interpretations of the story. Questions such as the following can guide discussion:

- Why did the eagle behave like a chicken?
- What helped him remember who he really was?
- What does the eagle symbolize?

These questions encourage learners to move beyond literal comprehension and consider the symbolic and thematic dimensions of the story.

The Reflection stage extends this process by encouraging learners to connect the narrative to their own experiences and beliefs. Reflection questions may include:

- Have you ever felt like the eagle—holding yourself back?
- What limits you from “spreading your wings”?
- Who are the people in your life who encourage and support you?

Through such discussions, learners are invited to examine themes of identity, self-worth, resilience, and personal growth while continuing to use language for meaningful communication.

The storytelling, discussion, and reflection activities described above illustrate how values-rich stories may be used to create meaningful opportunities for language use and personal engagement. As Murdoch (2002) observed, stories provide rich opportunities for developing language skills while encouraging discussion, interpretation, and learner involvement. By integrating storytelling with guided discussion and reflection, teachers can create learning experiences that are both linguistically meaningful and personally relevant.

Teacher Reflection Notes

Teachers are encouraged to observe how learners respond to storytelling activities and to reflect on the effectiveness of the instructional strategies used.

Reflection questions for teachers can include:

1. Did students participate actively in discussions?
2. Which vocabulary activities were most effective?
3. Did students demonstrate understanding of the story's message?

Such reflections can help teachers adapt and refine the framework for different classroom contexts while maintaining its emphasis on meaningful language use and personal reflection.

Discussion

The instructional framework presented in this article illustrates one way in which storytelling may be integrated into language instruction to support meaningful language use and personal reflection. Rather than focusing exclusively on vocabulary acquisition or grammatical accuracy, the framework encourages learners to engage with stories that explore themes such as identity, perseverance, self-worth, and personal growth. In doing so, it seeks to position language learning within meaningful and engaging contexts.

The five-stage sequence—Motivation, Vocabulary, Storytelling, Comprehension, and Reflection—draws upon established principles of language teaching while providing a structured pathway for learner engagement. The motivation stage activates prior knowledge and encourages learners to connect the lesson topic to their own experiences. The vocabulary stage prepares learners for comprehension by introducing key words and expressions. Storytelling provides contextualized language input, allowing learners to encounter language within meaningful narratives. Comprehension activities support understanding of the text, while the Reflection stage encourages learners to articulate personal responses and connect the story's themes to their own lives.

A notable feature of the framework is its emphasis on reflection as an integral part of language learning. While many language lessons focus primarily on comprehension and language practice, the Reflection stage extends learning beyond the text by inviting learners to consider the relevance of the story to their own experiences, beliefs, and aspirations. This emphasis on personal meaning-making aligns with literature-based approaches that view language learning as both a cognitive and an affective process.

Another distinctive element of the framework is the use of a secondary application story. Whereas the primary narrative introduces the central theme, the secondary story

enables learners to revisit and apply that theme in a new context. By examining similar ideas across two narratives, learners are encouraged to make connections, compare perspectives, and engage in deeper discussion and personal reflection. In this way, the secondary story serves as a bridge between story comprehension and broader exploration of values and life experiences.

Values-rich stories may also contribute to learner engagement and motivation. Stories that explore universal human experiences and challenges provide meaningful topics for discussion and encourage learners to express opinions, share experiences, and consider multiple perspectives. Such opportunities for communication support the development of language skills while fostering deeper engagement with the lesson content.

From the perspective of Social Cognitive Theory, storytelling may further support learners by presenting characters who model perseverance, courage, self-belief, and personal growth. Through discussion and reflection, learners are encouraged to examine these experiences and consider their relevance to their own lives. Although the framework presented in this article is not intended as an empirical evaluation of instructional effectiveness, it illustrates how storytelling can be used to integrate language learning with opportunities for reflection and meaningful communication.

Overall, the framework suggests that literature-based storytelling can serve as a valuable resource for English language teaching. By connecting language learning with reflection on meaningful themes and human experiences, teachers may create classroom environments that support communicative development while encouraging thoughtful engagement with important values and life lessons.

Pedagogical Implications for English Language Teachers

The use of values-rich storytelling in language classrooms has several important implications for English language teachers.

First, storytelling suggests that language learning need not be limited to mechanical exercises or isolated grammar instruction. Instead, teachers may design lessons that integrate language practice with meaningful content, discussion, and personal reflection. By situating language within engaging narratives, teachers can create opportunities for learners to use language for authentic communication rather than solely for skill practice.

Second, storytelling may support the development of communicative competence. When learners discuss characters, themes, conflicts, and lessons from a story, they practice expressing opinions, explaining ideas, justifying viewpoints, and responding to others. Such

activities provide meaningful contexts in which learners can use language for communication and interaction.

Third, literature-based instruction may contribute to learner engagement and motivation. Stories naturally capture attention, stimulate curiosity, and invite emotional involvement. When learners become invested in a narrative, they may be more willing to participate in discussions, collaborative activities, and reflective tasks associated with the story.

Another important implication is the potential of storytelling to support holistic education. In many educational contexts, language learning is expected to contribute not only to linguistic development but also to personal growth and character formation. Values-rich stories such as *The Eagle and the Chicken* and *How Many Cows Are You Worth?* provide opportunities for learners to reflect on themes such as perseverance, self-worth, resilience, and personal potential while simultaneously engaging in language-learning activities.

The framework presented in this article also offers a practical model that teachers may adapt to a variety of instructional contexts. The sequence of Motivation, Vocabulary, Storytelling, Comprehension, and Reflection provides a structured yet flexible approach to lesson design. While the lesson example presented in this article was developed for a teacher-development symposium, the framework itself may be adapted to different learner populations, proficiency levels, literary texts, and educational settings.

Ultimately, the framework illustrates how storytelling may be used not only as a source of language input but also as a vehicle for meaningful communication, reflection, and learner engagement. By integrating language learning with stories that explore significant human experiences and values, teachers can create classroom experiences that are both linguistically purposeful and personally meaningful.

Limitations and Future Directions

This article presents an instructional framework and illustrative lesson rather than an empirical investigation of classroom effectiveness. Consequently, no claims are made regarding measurable learning outcomes or the comparative effectiveness of the framework. Future research may explore how values-rich storytelling influences learner engagement, language development, reflective thinking, and communicative competence across different educational contexts and learner populations.

Conclusion

This article has presented a literature-based instructional framework that uses values-rich storytelling to support meaningful language use and reflection in English language

teaching. Drawing on the stories *The Eagle and the Chicken* and *How Many Cows Are You Worth?*, the article has illustrated how storytelling may be used to integrate language learning with exploration of themes such as identity, perseverance, self-worth, and personal growth. The proposed framework—Motivation, Vocabulary, Storytelling, Comprehension, and Reflection—provides a structured yet flexible approach to designing story-based language lessons. While the lesson example included in this article was developed for a teacher-development symposium, the framework may be adapted to a variety of learner populations, proficiency levels, literary texts, and educational contexts.

The instructional activities presented in this article illustrate how vocabulary development, storytelling, discussion, and reflection may be integrated within a coherent lesson sequence. By emphasizing both language use and personal meaning-making, the framework encourages learners to engage with stories as sources of communication, interpretation, and reflection.

For English language teachers, storytelling represents a valuable pedagogical resource that extends beyond traditional language instruction. By integrating literature with opportunities for discussion and reflection, teachers may create learning experiences that are both linguistically meaningful and personally engaging. The framework presented in this article is offered as one possible model for incorporating values-rich storytelling into language classrooms and may serve as a starting point for further adaptation, exploration, and classroom application.

References

- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Collie, J., & Slater, S. (1987). *Literature in the language classroom: A resource book of ideas and activities*. Cambridge University Press.
- Darmawati, B., Rahman, A., Halim, A., & Basri, M. (2020). Teaching English through literature-based instruction: An integrated study of language and literature. In *Proceedings of the International Conference on Community Development (ICCD 2020)* (Advances in Social Science, Education and Humanities Research, Vol. 477). Atlantis Press.
- Hossain, K. I. (2024). *Literature-based language learning: Challenges and opportunities for English learners*. *Ampersand*, 13, 100201. <https://doi.org/10.1016/j.amper.2024.100201>
- Lazar, G. (1993). *Literature and language teaching: A guide for teachers and trainers*. Cambridge University Press.
- Mardiani, M. & Baharuddin, A. F. (2023). *Literature-based EFL instruction: Benefits for language skills development, learning motivation, and cultural understanding*. *LLT Journal: A Journal on Language and Language Teaching*, 26(2), 796–811. <https://doi.org/10.24071/llt.v26i2.6366>
- Murdoch, G. (2002). Exploiting well-known short stories for language skills development. *IATEFL LCS SIG Newsletter*, 23, 9–17.
- Rahmawati, A., Suhatmady, B., Sucahyo, D., & Hermagustiana, I. (2025). Storytelling in English language learning: Boosting confidence and discourse in EFL contexts. *BORJU: Borneo Journal of Language and Education*, 7(1), 1–13. <https://doi.org/10.24903/bej.v7i1.1895>
- Sanchez, B. (2007). *8 secrets of the truly rich: How you can create material wealth and gain spiritual abundance at the same time*. Shepherd's Voice Publications.
- Tunnell, M. O., & Jacobs, J. S. (1989). Using "real" books: Research findings on literature-based reading instruction. *The Reading Teacher*, 42(7), 470–477.
- Wright, A. (1995). *Storytelling with children*. Oxford University Press.

Appendix A

The Eagle and the Chicken

A baby eagle once fell from its nest and landed in a chicken farm.

The farmer raised the eagle among the chickens, so the eagle grew up acting like a chicken, pecking at the ground and never looking at the sky.

One day, a naturalist visited and was shocked to see an eagle living like a chicken. He said, “This is an eagle—the king of the sky!” But the farmer insisted, “No, he’s a chicken now.”

The naturalist lifted the eagle onto a fence and said, “Eagle, stretch your wings and fly!”

But the eagle looked at the chickens, felt safe, and jumped back down.

The next day, the naturalist tried again from the roof—but the eagle still returned to the coop.

Finally, he brought the eagle far away to the foot of a high mountain. He held the eagle toward the bright sun and said, “Eagle, you belong to the sky. Spread your wings and fly!”

This time, the eagle looked at the sun, stretched his great wings, and—slowly at first, then powerfully—soared into the sky, finally becoming who he was born to be.

Appendix B

How Many Cows Are You Worth?

In a small village, a man had to offer cows as a dowry when he wanted to marry a woman.

The number of cows showed how beautiful she was, one to ten cows.

Brad wanted to marry a kind, intelligent young woman named Angelina. But she wasn't considered beautiful, and her family believed she was worth only one cow.

When Brad asked for her hand, the father said, "Just one cow."

Brad surprised everyone: "I will not give you one cow... I will give you ten."

People thought he was crazy. Angelina herself was shocked. But Brad insisted:

"Your daughter is worth more than the highest price a man can pay."

Years passed. And something magical happened—Angelina became the most beautiful woman in the village.

Because she believed she was a ten-cow woman, she carried herself like one... and the whole village saw her that way.

Teaching AI Literacy Through Business Writing: A Practice Report on AI-Assisted L2 Writing in a Japanese University

Misa FUJINUMA

Nagoya University of Foreign Studies

Abstract

This practice report describes an advanced, English-medium business writing course for third and fourth year English majors at a Japanese private university, where AI tools were integrated as a planned part of instruction rather than treated mainly as a disciplinary risk. The course responded to students' tendency to view AI as cheating while also facing workplaces that expect responsible, AI-supported communication. Grounded in transformative learning, self-authorship, and sociocultural scaffolding, each assignment followed a process-writing cycle with explicit AI literacy activities designed to make AI use visible and discussable. Two end-of-semester course-reflection questionnaires suggested a shift in students' self-reported practices from secret, answer-focused AI use toward more selective use for ideas, language refinement, and feedback, alongside ongoing concerns about overreliance and loss of the writer's voice. This report offers practical principles for embedding AI literacy into business writing courses in Japanese higher education.

Keywords: AI literacy, business writing, L2 writing, student agency

Teaching AI Literacy Through Business Writing: A Practice Report on AI-Assisted L2 Writing in a Japanese University

Generative AI (GenAI) is increasingly embedded in Japanese workplaces, making the ability to interpret and critically evaluate AI output an important professional competence (Access Partnership, 2023). Society 5.0 likewise emphasises human judgment, creativity, and problem-solving in digital transformation (Cabinet Office, Government of Japan, 2019). For teachers of business writing, these developments demand a shift in focus. If workplaces increasingly rely on AI for routine tasks while expecting humans to contribute creativity, judgment, and collaboration, then writing education cannot centre only on policing AI use. Instead, it must help students learn to use AI critically, ethically, and strategically as part of their communication toolkit. AI literacy, understood as the capacity to understand, use, monitor, and critically evaluate AI and its output, is therefore becoming a key professional competence, particularly for students who are about to enter AI-rich business environments (Cabinet Office, Government of Japan, 2024).

This practice report examines an AI-integrated, process-oriented business writing course in which GenAI and grammar checkers were used for brainstorming, drafting, feedback, revision, and reflection. It investigates how students described changes in their perceptions and use of AI (RQ1) and how reflection, peer and AI-generated feedback, and group discussion shaped their sense of agency and ownership over AI-assisted writing (RQ2). Here, AI literacy refers to the capacity to use and evaluate AI critically and ethically while retaining responsibility for the final text.

Literature Review / Theoretical Background

GenAI-Assisted L2 Writing: Affordances and Challenges

Generative artificial intelligence (GenAI) has become increasingly integrated into second language (L2) writing instruction, supporting learners across the writing process from ideation to proofreading (Yang et al., 2025). In their review of 73 empirical studies, Yang et al. (2025) found that GenAI was most frequently used for feedback and revision, followed by content development and vocabulary and grammar support. GenAI can therefore support L2 writers at multiple stages; however, its educational value depends on students' ability to evaluate and act on its output critically. Reported challenges include overreliance, inaccurate or inconsistent feedback, limited contextual understanding, and ethical concerns.

Existing research also remains concentrated in particular settings and genres. Yang et al. reported that most studies examined academic writing, whereas only two of the 73 studies focused on professional writing. In addition, 29 studies did not state a clear theoretical

framework. These patterns indicate a need for context-sensitive, theoretically informed accounts of how AI literacy can be developed in professional L2 writing classrooms.

Critical AI Literacy and Learner Agency

A central concern in AI-assisted writing is the possibility that learners may treat GenAI output as more reliable or linguistically authoritative than their own judgment. Shibata (2026) conceptualises this orientation as *GenAI-Writerism*: a psychological belief that GenAI can produce linguistically superior and appropriate writing and can therefore be trusted as a reliable writing authority. Importantly, Shibata distinguishes GenAI-Writerism from AI literacy. AI literacy concerns what learners can do with AI critically and effectively, whereas GenAI-Writerism concerns what they believe about AI's linguistic authority.

Shibata's study of Japanese university EFL students suggests that strong trust in GenAI may be associated with low writing self-efficacy and limited language or genre knowledge. Students who felt unable to judge vocabulary, grammar, or genre appropriateness were more likely to accept GenAI output without critical evaluation. This pattern is particularly relevant in L2 writing contexts, where fluent or apparently native-like AI output may be interpreted as automatically appropriate.

At the same time, Shibata does not argue that trust in GenAI is inherently harmful. GenAI can provide useful language models, ideas, and feedback. The risk arises when learners accept AI suggestions without considering audience, purpose, genre, or the reasons for a proposed revision. Such uncritical reliance may weaken writers' voice and ownership (Shibata, 2026).

Feedback, Reflection, and Process-Oriented Use

Recent classroom studies suggest that GenAI can be most productive when integrated into a visible writing process rather than used to generate finished text. Mullin (2026), for example, incorporated AI as a post-drafting grammar and vocabulary support tool in a Japanese university context. Students first produced their own writing without AI assistance, then used AI to check their drafts and reflected on what they learned from the feedback, including which suggestions they accepted or rejected. This sequence foregrounded learning progress and reflection rather than mechanical correction alone.

Huang et al. (2025) similarly demonstrate that GenAI feedback and peer feedback may serve different but complementary functions. Learners valued GenAI for immediate feedback, error detection, vocabulary support, and structural suggestions. At the same time, they raised concerns about inaccurate, impersonal, or difficult-to-interpret feedback and about becoming overly dependent on the tool. Peer feedback, by contrast, was associated with

learner autonomy, critical thinking, self-correction, and attention to content and expression. Although Huang et al. compared the two feedback modes rather than testing an integrated model, their findings suggest the value of asking learners to compare feedback sources rather than treating AI as a replacement for peers or teachers.

Together, these studies suggest that AI literacy involves making feedback use visible, comparing advice from different sources, and explaining revision decisions. Reflection can help students retain responsibility for the final text.

AI Literacy in Business Writing

The need for critical AI literacy is particularly significant in business communication. Iaia and Pironti (2024) argue that implementing AI in business communication should not be understood as simple tool adoption. Instead, it involves organisational learning, knowledge sharing, human evaluation, and responsibility for communication outcomes. In this view, AI can provide forecasts, suggestions, or behavioural guidance, but decision-making remains a human responsibility.

This emphasis aligns with the demands of business writing. Professional writers must make choices that reflect a specific purpose, audience, organisational context, and ethical responsibility. Even linguistically accurate AI-generated language may be ineffective or inappropriate if it does not match the intended reader, message, relationship, or communicative goal. Therefore, students preparing for AI-rich workplaces need opportunities to use AI while retaining responsibility for the content, voice, and consequences of their communication.

Theoretical Orientation and Present Study

This practice report draws on transformative learning, self-authorship, and sociocultural scaffolding to interpret students' engagement with AI-assisted writing. Transformative learning highlights how reflection and dialogue may help learners reconsider assumptions, such as the belief that AI is always cheating or always correct (Mezirow, 1997, 2009). Self-authorship focuses on learners' movement away from dependence on external authorities, including AI, toward making judgments based on their own developing values and interpretations (Baxter Magolda, 2007). Sociocultural scaffolding frames peer interaction, AI feedback, teacher guidance, and structured writing stages as forms of temporary support that can help students make increasingly independent writing decisions (Vygotsky, 1978; Wood, Bruner, & Ross, 1976).

Building on these perspectives, this study examines an AI-integrated, process-oriented business writing course for advanced Japanese university students. It explores how students

described changes in their perceptions and use of AI, and how reflection, peer and AI-generated feedback, and group discussion related to their sense of agency and ownership over AI-assisted writing.

Research Issues and Questions

Despite the growing relevance of AI in professional communication, students may regard it as both a prohibited shortcut in university writing and a necessary workplace tool. This study therefore examines how an AI-integrated, process-oriented business writing course supported students' critical and ethical AI use, agency, and ownership.

RQ1: How did students describe changes in their perceptions and use of AI in writing during an AI-integrated, process-oriented business writing course?

RQ2: How did students describe the role of reflection, peer and AI-generated feedback, and group discussion in shaping their sense of agency and ownership over AI-assisted writing?

Method

The research was conducted as classroom-based qualitative research supplemented with descriptive survey data. The design used small-scale, mixed-methods classroom surveys that included five-point Likert scale items and open-ended comments.

Participants and the Research Context

This year-long study involved two English-medium advanced business-writing classes for third- and fourth-year English majors at a Japanese private university. Students generally had upper-intermediate to advanced English proficiency (CEFR B2–C1). In Group 1, enrolment was 14 students in the first semester and 16 in the second; three students joined, and one did not continue. In Group 2, enrolment was 15 students in the first semester and 13 in the second; two students did not continue, including for study abroad. Total enrolment was therefore 29 students in each semester. I designed and taught both classes as a teacher-researcher. This may have influenced questionnaire responses, for example through social desirability or reluctance to criticise the course. To partially address this, participation was voluntary, responses were anonymous, and questionnaire responses were not used for grading or formal course evaluation.

The university had no institution-wide generative-AI policy. Teachers established course-specific rules and explained them in their syllabi. The university informed students that they must not submit AI-generated text as their own work, must verify AI outputs because they may be inaccurate, and must not enter personal or confidential information into AI tools. Course policies consequently ranged from prohibition to limited or guided use. In this course, AI was used under explicit conditions and to develop AI literacy. Students were

told in the syllabus that AI would be used regularly, detailed written guidance would be given in the first two classes, and unmodified AI-generated text was prohibited in assignments. They could access Gemini through university Google accounts, although tools and frequency of use varied.

Data Collection Instruments and Procedures

Three anonymous, voluntary questionnaires were administered: two at the end of the first semester and one at the end of the second semester. The “AI Use in Writing Class” questionnaire was completed by 21 students and asked students to reflect on their AI use and attitudes before and after the first semester. The first-semester course-reflection questionnaire was completed by 27 students, and the second-semester course-reflection questionnaire was completed by 21 students. Students also maintained AI trackers for pedagogical reflection. Although submitted with final drafts as a course requirement, they were not standardised or systematically retained and analysed as research data and were therefore excluded from the formal analysis.

Overall course. In the first class, students received an orientation that introduced different types of AI tools, including generative systems, writing enhancement tools, and translation tools, and invited discussion of perceived benefits and risks. They were asked to set intentions for AI use, for example, learning to use grammar checkers as a minimum requirement, and discussed ethical issues such as transparency, accuracy, plagiarism, and critical evaluation. To make AI use visible, students maintained an AI tracker throughout the course.

Writing cycle and AI integration. Each major assignment followed the same four-stage cycle: brainstorming and researching, drafting, review, and editing (Table 1). At each stage, students took part in both non-AI activities and AI-supported activities, followed by peer or group discussion. Table 1 summarises the four stages of the writing cycle, their purposes, and the main non-AI and AI-supported activities used in class.

Table 1***Writing cycle stages, AI integration, and reflective activities***

Stage	Purpose	Main focus in class
Brainstorm /Research	Generate and select ideas	Individual paper brainstorming, then AI brainstorming; brief notes distinguishing own and AI ideas and reasons for selection
Draft	Create initial text from own ideas	Handwritten or typed draft from students' ideas; optional AI language support; short AI tracker entry on if/how AI was used
Review	Develop feedback literacy and critical stance	Peer review using a checklist; AI feedback with prompts; small-group discussion of which suggestions to accept or reject
Edit	Refine language and structure while keeping voice	Prompt-crafting activity; AI for options (verbs, phrasing, clarity); short written comments explaining feedback choices

Brainstorming. Students first brainstormed individually on paper without technology, in order to commit to their own ideas. They then repeated the brainstorming with an AI tool, using sample prompts such as “Suggest problems or opportunities for [company or job]” or “What is a trending topic in [industry]?” In small groups, they compared their own ideas with AI-generated ideas, and discussed which ideas seemed realistic, unique, or suitable for proposals, and noted which ideas came from themselves and which came from AI. This sequence positioned AI as a source of possibilities while requiring students to evaluate and select ideas in relation to their own purposes.

Drafting. Drafts were then written by hand or on computers without copying AI-generated sentences directly. Students could consult AI for language support, but the emphasis was on producing the main text themselves so that later AI suggestions would respond to their own writing rather than replace it. After drafting, students recorded in their AI tracker whether and how they had used AI, including which tools they used, what they asked, and at which points in the drafting process. The tracker made students' AI-use decisions visible and reinforced their responsibility for the draft.

Reviewing. In the review stage, students first exchanged drafts and used a checklist to comment on clarity, specificity of objectives and benefits, persuasiveness, and any confusing sections. Next, they submitted their drafts to an AI tool with prompts such as “What are the strengths of this draft?” or “What questions might a manager still have after reading this?” and students were encouraged to have a conversation with AI and try different prompts. Groups then discussed which feedback from peers and which feedback from AI they

found useful, which advice they chose not to follow, and why. Comparing feedback sources required students to evaluate suggestions rather than treat AI feedback as automatically correct.

Editing. Before editing their drafts, students completed a prompt crafting activity using a simple sentence. They compared AI outputs produced by broad prompts such as “Fix this sentence and make it better” with outputs from more focused prompts such as “Suggest a list of strong action verbs I can use in this sentence,” and discussed which responses helped them maintain their own voice. They then applied similar prompt strategies to their own drafts, using AI to generate options for verbs, phrasing, or more concrete objectives while deciding themselves which changes to implement. Alongside their revised drafts, students discussed in small groups and later submitted short written comments explaining which peer, AI, and teacher feedback they decided to accept or reject and the reasons for these decisions. This reflection positioned students as the final decision-makers responsible for both the text and their use of AI.

Data Analysis Instruments and Procedures

Data were collected through three online questionnaires completed in English. At the end of the first semester, students completed the AI Use in Writing Class questionnaire (n=21) and a course-reflection questionnaire (n=27). The former included five-point Likert-scale items on AI-use frequency, attitudes towards AI in writing, and confidence in responsible AI use, as well as open-ended questions about students’ views and use of AI before and after the course. The reflection questionnaire asked about students’ experiences, concerns, and changing AI use during the semester. A third questionnaire at the end of the second semester (n=21) asked students to reflect on changes in their understanding of business writing and use of AI. As responses were anonymous, data could not be matched across administrations and do not indicate individual longitudinal change.

For quantitative analysis, Likert-scale responses were exported from Google Forms and summarised using descriptive statistics. Items were grouped by whether they addressed AI use, attitudes, and confidence before the course or after the first semester. Frequencies and percentages for each response option were calculated to identify class-level patterns and changes. These summaries were then used to address RQ1 by indicating how students’ reported perceptions and self-reported practices differed before the course and after the first semester.

For the qualitative analysis, I conducted an inductive thematic analysis of all open-ended responses. Comments were organised into six categories: views of AI before the course

and after the first semester; reported AI use before the course and after the first semester; and concerns or other comments after the first and second semesters. Responses were reread to develop familiarity with the data, and keywords and phrases were highlighted and assigned brief descriptive codes. Segments could receive multiple codes when they expressed more than one idea. Through repeated comparison within and across categories, related codes were developed into broader exploratory themes.

Themes included fear of or prohibitive views towards AI versus tool-based views; copying AI output versus using AI for brainstorming, organisation, and feedback; and concerns about overreliance, loss of voice, prompt difficulty, hallucination, and more careful or strategic AI use. Representative quotations were selected to illustrate recurrent themes and the range of views. Changes in themes across categories were examined in relation to the research questions: shifts in AI views and writing-stage use informed RQ1, while themes of responsibility, voice, and decision-making informed RQ2 on agency and ownership in AI-assisted writing.

Results/Findings

RQ1: Perceptions and Use of AI

How students viewed AI before the course and after the first semester. Questionnaire responses showed that many students initially saw AI as a cheating tool that was not allowed in writing classes. In an open-ended item about views before the course, comments were most often coded as AI as cheating / prohibited (10 coded comments) or Lack of knowledge / too difficult to use (4). At the same time, a closed item indicated that a total of 67% of respondents reported relying on AI “always,” “often,” or “sometimes” to get direct answers (see Table 2), suggesting that AI use for obtaining direct answers coexisted with perceptions of AI as prohibited or risky. Some comments admitted that students “used AI for almost all questions to cheat,” while others noted that AI could help if used “in the correct way.” Taken together, these responses suggest that many students viewed AI in academic writing as prohibited and risky, while also reporting its use to obtain direct answers.

Table 2***Frequency of Students' Reliance on AI for Direct Answers Before the Course***

Response option	n	%
Always	1	5
Often	7	33
Sometimes	6	29
Rarely	4	19
Never	3	14

Note. AI Use in Writing Class questionnaire; n = 21.

By the end of the first semester, responses indicated more positive views of AI. Of the 21 respondents, 11 (52%) selected “much more positive” and 9 (43%) selected “somewhat more positive”; thus, 20 respondents (95%) reported a more positive view. One respondent (5%) selected “somewhat more negative,” and none selected “much more negative” (see Table 3). The comment linked to the “somewhat more negative” response suggested a move away from blind trust rather than rejection of AI: the writer reported previously “using AI for almost all questions,” but now recognised that “sometimes AI’s information is wrong” and that “AI is useful but sometimes it’s dangerous,” and described changing from asking AI to write sentences to using it mainly to “give me some example” rather than to “seek the answer.”

Table 3***Change in Positivity Toward Using AI in Writing After the First Semester***

Response option	n	%
Much more positive	11	52
Somewhat more positive	9	43
No change	0	0
Somewhat more negative	1	5
Much more negative	0	0

Note. AI Use in Writing Class questionnaire; n = 21.

Comments about changed understanding and attitudes further illustrate this shift. Responses described AI as helping to “expand opinions and expose different ways of thinking” and as “highly effective for creative learning,” with one describing “smart utilization of AI” as “highly effective for creative learning” and another calling AI “the best partner on [the] internet,” showing a change from a secret cheating tool to a helpful learning tool.

How students reported using AI before the course and after the first semester.

Students’ descriptions of their writing indicated a similar change in their reported practices,

from having AI take over their writing to using AI to guide and support their own work. In comments about how they used AI before the course, responses were most often coded as AI gives answers (3 coded comments) and AI writes/rewrites text (2), with some comments mentioning grammar/word correction (2) or no/little AI use (1). Several students described AI as a “teacher” that “submit answers” or as a tool to “make the sentence,” showing that students often let GenAI write for them rather than using it to guide their decisions.

When asked how they now use AI differently after the first semester, comments were most frequently coded as grammar/wording help (7 coded comments), brainstorming and question generation (6), advice or feedback on drafts (6), examples or ideas (5), and more limited or careful use (2). Students described moving from “make it better” prompts to asking for “some example” and then choosing only expressions they already understood, using AI to suggest alternative wording or stronger verbs, and consulting AI when unsure whether a paragraph was “strong” rather than asking it to write the paragraph itself. Because individual responses could receive more than one code, these numbers represent coded comments rather than unique students. Students also valued AI for providing alternative ideas, supporting organisation and language choices, and making assistance available when teachers were not present.

Confidence in responsible AI use. Confidence ratings aligned with these qualitative shifts. On the item “How confident do you feel now about using AI responsibly and effectively for your writing?,” 2 responses (10%) selected “Extremely confident,” 8 (38%) chose “Very confident,” and 9 (43%) selected “Moderately confident,” whereas only 2 (10%) reported being “Slightly confident” and none chose “Not confident at all” (see Table 4). Most reported medium-to-high confidence in using AI responsibly, suggesting that they could use AI as a supportive tool rather than a shortcut, but still recognised that their skills and judgment could improve.

Table 4

Confidence in Using AI Responsibly and Effectively for Writing After the First Semester

Response option	n	%
Extremely confident	2	10
Very confident	8	38
Moderately confident	9	43
Slightly confident	2	10
Not confident at all	0	0

Note. AI Use in Writing Class questionnaire; n = 21.

RQ2: Agency, ownership, and reflection

Concerns and comments after the first semester. Open-ended responses after the first semester showed ongoing worries about how to use AI responsibly. In the coded comments on remaining concerns, Overreliance / dependence on AI was the most frequent code, appearing in more than half of the comments, with several students describing themselves as “too dependent on AI” or letting AI “do all the work.” A smaller group reflected Loss of individuality / originality, such as “Individuality is lost” or worries that copy-and-paste use of AI would replace their own voice. Students also raised concerns about Accuracy and technical limits, Incorrect or low-quality information, Prompting skills / how to ask, and the Effort / burden of responsible use, mentioning that AI is “not accurate,” may lead to misunderstanding, is “troublesome” to check, and is hard to prompt effectively.

These patterns suggest that, in their comments, students’ strategic use of AI improved and they also became more aware of the ethical and identity-related part of AI-assisted writing. Rather than rejecting AI, they reported responsible use as an ongoing challenge that required balancing usefulness with the risks of overreliance, voice loss, and misinformation, and these concerns led to the decision to increase in-class discussion of AI suggestions and prompt crafting during the second semester.

Concerns and comments after the second semester. End-of-semester reflections suggested that many students had developed more confident, self-authored positions towards AI-assisted writing. Responses to “How has your understanding, usage, or attitude toward AI use in writing changed, particularly after this semester?” most frequently highlighted AI as useful support / tool and Better prompting / specific questions, followed by Maintain own voice / think for myself, Use AI for ideas / brainstorming / structure, and a More critical / balanced attitude. One student explained, “Before, I used AI to write for me. Now, I use AI to help my thinking... The writing itself is done by me,” while another described AI as “an ally in refining my own ideas.” These comments illustrate a more balanced relationship between AI support and personal authorship.

Several students explicitly linked their changed attitudes to course activities, including peer review, teacher feedback, and explicit work with prompts. They reported that learning to ask more specific questions made AI responses “more logical and convincing,” and that AI worked best when used alongside human feedback rather than as the main authority, echoing Better prompting / specific questions and AI as useful support / tool. Comments such as “I must never stop thinking myself, instead of leaving everything to AI” and “I’m going to keep using AI to make a choice in myself” exemplify Maintain own voice /

think for myself and More critical / balanced attitude, showing students articulating an internal standard for responsible AI use.

Some students nevertheless continued to identify appropriate AI use, overreliance, and accuracy as ongoing concerns. Overall, the findings suggest that discussing AI suggestions and prompt crafting supported greater control, responsibility, and critical engagement, while students recognised that their relationship with AI would continue to evolve.

Discussion

RQ1: Perceptions and use of AI

The self-reported shift towards using AI for brainstorming, language refinement, and feedback is consistent with common L2-writing uses identified by Yang et al. (2025) and with Mullin's (2026) process-based approach. However, because the evidence is anonymous self-report data, the findings indicate changing perceptions and reported practices rather than demonstrated improvement in writing ability.

RQ2: Agency, ownership, and reflection

Students' continuing concerns about overreliance, accuracy, and voice support the view that AI literacy requires ongoing judgment rather than technical confidence alone. Consistent with Shibata (2026), reflection and discussion may have helped students question AI authority and articulate responsibility for revision decisions. The second-semester reflections suggest, but do not establish causally, that prompt work and feedback comparison supported a more agentive stance.

Implications for Teaching

For writing pedagogy, this practice report suggests five principles for integrating AI literacy into business writing in Japanese higher education. First, students should brainstorm without AI before using it to expand or refine their ideas, helping preserve ownership of the initial content. Second, prompt design should be taught as part of writing instruction because specifying context, audience, constraints, and goals develops rhetorical awareness as well as more purposeful AI use. Third, AI trackers or brief process notes should make students' AI use visible and open to reflection. Fourth, students should compare peer and AI-generated feedback and explain which suggestions they accept, modify, or reject; this reflects Huang et al.'s (2025) finding that AI and peer feedback can play complementary roles, with AI offering immediate linguistic and structural support and peer feedback supporting autonomy, critical thinking, and attention to meaning. Fifth, students should reflect on their revision decisions and risks such as overreliance and loss of voice. This responds to Shibata's (2026) call for learners to evaluate AI output, consider alternatives, and validate revision decisions.

Together, these practices support the critical evaluation and deliberate interaction with AI output emphasised in L2 writing research (Yang et al., 2025) and maintain the human responsibility required in AI-supported business communication (Iaia & Pironti, 2024).

Limitations

This practice report is based on a small sample size of advanced English majors in two classes at a single Japanese private university, so the findings are context-specific and cannot be generalised to other institutions, course types, or proficiency levels. The study did not include a pre-course baseline survey or a controlled comparison group, and the main data sources were anonymous self-report questionnaires, raising the possibility of social desirability bias and other teacher-researcher effects. Qualitative coding was conducted by one researcher using inductive procedures without inter-coder reliability checks, and AI tracker logs and writing products were not analysed systematically. As a result, the themes reported here should be interpreted as exploratory patterns in students' self-reported perceptions and practices rather than as evidence of causal impact, and any claims about agency or ownership refer to students' own accounts rather than independently verified changes in their writing. Future research could incorporate more systematic use of AI tracker logs and writing samples, examine how similar approaches work with lower-year or differently motivated students, and explore ways to make the human-AI division of labour more visible in students' texts without undermining their sense of ownership.

Adaptation to other contexts

Building on these implications, the core elements of the design can be adapted to different proficiency levels, class sizes, and institutional AI policies. Although this course was taught to intermediate-to-advanced English majors in small classes, the central ideas are that AI use is kept visible, that students regularly reflect on how and why they use AI, and that they justify choices about peer, AI, and teacher feedback.

For lower-proficiency or non-English-major students, teachers can use shorter texts, model prompts, sentence frames, or shared drafting before asking students to evaluate AI suggestions. In larger classes, teachers may limit the number of prompts or feedback points and use brief checklists or oral reflection. Where AI access is restricted, teachers can demonstrate AI feedback on sample texts and ask students to evaluate the suggestions; short process notes or revision explanations can still make students' decision-making visible.

Conclusion

This practice report suggests that making AI use visible, discussable, and reflective may support a shift from secretive or answer-seeking use towards more selective use for

brainstorming, language refinement, and feedback. Students nevertheless continued to identify overreliance, voice, and accuracy as concerns, underscoring that AI literacy involves ongoing ethical and rhetorical judgment.

Acknowledgments

I would like to thank the students in my business writing courses who generously participated in this study and shared their experiences with AI-assisted writing. Their thoughtful reflections and willingness to engage critically with AI tools made this research possible.

References

- Access Partnership. (2023). *The economic impact of generative AI: The future of work in Japan*. <https://cdn.accesspartnership.com/wp-content/uploads/2023/06/Economic-Impact-of-Generative-AI-The-future-of-work-in-Japan.pdf>
- Baxter Magolda, M. B. (2007). Self-authorship: The foundation for twenty-first-century education. *New Directions for Teaching and Learning*, 2007(109), 69–83. <https://doi.org/10.1002/tl.266>
- Cabinet Office, Government of Japan. (2019, March 29). 人間中心の AI 社会原則 [Social principles of human-centric AI]. <https://www8.cao.go.jp/cstp/ai/aigensoku.pdf>
- Cabinet Office, Government of Japan. (2024). 第 1 章 第 3 節 AI 活用に向けたリスキリングと教育 [Chapter 1, Section 3: Reskilling and education for AI utilization]. https://www5.cao.go.jp/j-j/sekai_chouryuu/sh24-01/s1_24_1_3.html
- Huang, J., Mizumoto, A., & Huang, Y. L. (2025). Comparing peer feedback and generative artificial intelligence feedback in a Japanese English-as-a-foreign-language speaking context: Impacts on motivation, engagement, and writing self-efficacy. *International Journal of Applied Linguistics*. <https://doi.org/10.1111/ijal.70020>
- Iaia, L., & Pironti, M. (2024). Supporting the implementation of AI in business communication: The role of knowledge management. *Journal of Knowledge Management*, 28, 85–95. <https://doi.org/10.1108/JKM-12-2022-0944>
- Mezirow, J. (1997). Transformative learning: Theory to practice. *New Directions for Adult and Continuing Education*, 1997(74), 5–12.
- Mezirow, J. (2009). An overview on transformative learning. In K. Illeris (Ed.), *Contemporary Theories of Learning: Learning Theorists in Their Own Words* (pp. 90–105). Routledge.
- Mullin, B. (2026). Responsible AI integration in university EFL writing classes in Japan: Grammar-checking tasks and ethical framing at Kobe University. *Journal of the School of Languages and Communication, Kobe University*, 22, 1–10. <https://doi.org/10.24546/0100502588>
- Shibata, N. (2026). Generative artificial intelligence writerism: Factors in overestimation of generative artificial intelligence output. *Journal of Humanistic English Language Education Research*, 6, 6–21. https://researchmap.jp/nshibata_708/published_papers/53858205

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100.
- Yang, C., Li, R., & Yang, L. (2025). Revisiting trends in GenAI-assisted second language writing: Retrospect and prospect. *Journal of Educational Computing Research*, 1–45.
<https://doi.org/10.1177/07356331251367309>

Appendix A

AI Use in Writing Class (First Semester)

1. How often did you rely on AI to get direct answers before this course?
 - 1 = Never
 - 2 = Rarely
 - 3 = Sometimes
 - 4 = Often
 - 5 = Always
2. Before this course, how did you view the use of AI in your writing assignments?
 - Open-ended response
3. After this course, how positive do you feel about using AI in your writing?
 - 1 = Much more negative
 - 2 = Somewhat more negative
 - 3 = No change
 - 4 = Somewhat more positive
 - 5 = Much more positive
4. How has your understanding or attitude toward AI use in writing changed as a result of this course?
 - Open-ended response
5. How do you use AI differently in your writing process now compared to before the course? Please give examples.
 - Open-ended response
6. For which parts of the writing process do you now use AI most often? (Select all that apply.)
 - Brainstorming ideas
 - Outlining
 - Drafting
 - Editing/proofreading
 - Other: _____
7. What benefits do you see in using AI as a tool for brainstorming, organizing, or improving your writing?
 - Open-ended response
8. How confident do you feel now about using AI responsibly and effectively for your writing?

- 1 = Not confident at all
- 2 = Slightly confident
- 3 = Moderately confident
- 4 = Very confident
- 5 = Extremely confident

9. What challenges or concerns do you still have about using AI for writing assignments?

- Open-ended response

10. Any other comments or questions?

- Open-ended response

Appendix B

End-of-Semester Course-Reflection Survey (First and Second Semesters)

1. Class period (G1/G2, day, time).
2. How has your understanding and approach to professional business writing changed throughout this course?
 - Open-ended response
3. Describe a significant challenge you faced in your writing process (planning, drafting, revising, or using feedback/technology).
 - Open-ended response
4. Was there anything about the class organisation or my teaching style that helped you learn well? Is there anything you think could be improved?
 - Open-ended response
5. How has your understanding, usage, or attitude toward AI use in writing changed, particularly after this semester?
 - Open-ended response
6. Any other comments or questions?
 - Open-ended response

Diagnosing and Measuring Pragmatic Development through Corpus-Informed Instruction: A Classroom Study of English Apologies

Taylor MEIZLISH and Christopher ZIFFO

Nagoya University of Foreign Studies

Abstract

Second language learners may produce grammatically accurate apologies that lack pragmatic nuance. This classroom-based study used a corpus comparison to identify gaps in Japanese university students' English apology strategies and then evaluated a targeted instructional intervention. An English L1 apology corpus ($n = 60$) served as a benchmark, and a learner corpus ($n = 60$) revealed systematic underuse of forms of address, exclamation words, and intensifiers. Forty-five first-year students were assigned to a test group or control group and completed a pretest discourse completion task (DCT). The test group received a 90-minute strategy-focused lesson, while the control group received no treatment. Posttest results showed that the test group increased use of the three target strategies across contexts varying in relative power, social distance, and imposition, with the strongest gains in intensifiers and address forms. Findings support corpus-informed target selection and suggest that even brief explicit instruction can enhance pragmatic strategy use.

Keywords: interlanguage pragmatics, apologies, explicit instruction, corpus-informed instruction, Japanese EFL

Diagnosing and Measuring Pragmatic Development through Corpus-Informed Instruction: A Classroom Study of English Apologies

Pragmatic competence—the ability to know what to say, how to say it, and when to say it—is central to maintaining rapport and managing relationships in interaction. Through speech acts, such as requests, apologies, refusals, and expressions of gratitude, speakers negotiate responsibility, express stance, and signal interpersonal alignment. For second language learners, difficulty often lies not in using appropriate grammar, but in selecting appropriate language strategies to express their intent. In foreign language contexts where exposure to naturalistic interaction is limited, such abilities are often expected to develop implicitly, leaving learners with grammatically accurate yet pragmatically less appropriate responses. Previous research has also shown that pragmatic features are often underrepresented in language teaching materials, leaving learners with limited exposure to contextualized models of speech act use (Vellenga, 2004).

Despite the recognized importance of speech acts, classroom pragmatics instruction is often evaluated impressionistically by teachers. Teachers may feel that learner language sounds more natural after instruction, but such judgments are difficult to quantify. This raises a practical question for teacher development: how can pragmatic growth be diagnosed and measured systematically in ordinary classroom settings? A corpus-informed approach offers one solution. By comparing learner production with a benchmark corpus of proficient L1 users and operationalizing strategies as countable categories, significant gaps can be identified, and improvement can be measured through changes in strategy frequency.

The present study applies this method to English apologies, a speech act that requires speakers to manage responsibility and repair relationships after an offense. The study proceeded in two phases. An L1 benchmark corpus and a learner corpus were compared to diagnose underused strategies. These corpus-based findings then informed a focused classroom intervention. Pretest and posttest responses were analyzed through strategy counting to measure improvement across situations varying in power, social distance, and imposition.

Literature Review

Interlanguage pragmatics (ILP) conceptualizes pragmatic competence as the ability to use linguistic resources to accomplish communicative goals and express intentions in ways that are socially and contextually appropriate (Kasper & Rose, 2002). A central distinction in ILP research is between pragmalinguistics, the linguistic forms used to perform speech acts, and sociopragmatics, the social perceptions and contextual judgments that guide appropriate

language use (Leech, 1983; Kasper & Rose, 2002). Development in these domains does not necessarily parallel grammatical development; learners may acquire grammatically accurate language while still struggling to choose language that fits the social situation. Empirical studies have shown that L2 learners often diverge from L1 speakers in how they select, combine, or distribute speech act strategies, particularly across variations in power, social distance, and imposition (Bardovi-Harlig, 2013; Taguchi, 2015). Such findings underscore that pragmatic development involves not only acquiring forms but learning how different strategy choices communicate different social relationships. More recent work has continued to emphasize the importance of developing pragmatic competence that enables learners to communicate effectively across diverse linguistic and cultural settings (Sánchez-Hernández & Martínez-Flor, 2022).

Among speech acts, apologies have been widely examined as a complex speech act set composed of multiple semantic formulas that can be combined in different ways (Olshtain & Cohen, 1983; Blum-Kulka & Olshtain, 1984). Early cross-cultural research demonstrated that English apologies frequently include strategies such as explicit expressions of regret, acknowledgment of responsibility, explanations, offers of repair, and promises of forbearance. The distribution of these strategies varies systematically according to contextual variables such as power, social distance, and offense severity.

Subsequent interlanguage studies have shown that L2 learners often rely heavily on a core apology formula while underusing or differently combining supportive strategies compared to L1 speakers (Bardovi-Harlig, 2013; Prachanant, 2016). These differences are not merely stylistic; they reflect variation in how learners manage responsibility and maintain relationships with others, and they are observable through differences in strategy frequency and distribution across contexts. Cross-cultural research has also documented differences in how apologies are realized, including tendencies for Japanese apologies to emphasize expressions of remorse and responsibility, while English apologies more often incorporate explanations or offers of repair (Sugimoto, 1998; Bergman & Kasper, 1993).

Research in instructed pragmatics has consistently shown that pragmatic ability is amenable to classroom intervention. Drawing on the Noticing Hypothesis (Schmidt, 1990), scholars have argued that learners benefit when instruction explicitly highlights the relationship between language forms and their communicative functions and sociocultural constraints that might otherwise go unnoticed (Schmidt, 1993; Takahashi, 2010). Empirical syntheses indicate that explicit instruction tends to yield stronger effects than implicit exposure alone, particularly when learners engage in metapragmatic reflection and guided

practice (Jeon & Kaya, 2006; Taguchi, 2015). A recent meta-analysis focusing specifically on requests, apologies, and refusals likewise concluded that explicit instruction has a positive effect on learners' pragmatic development across a range of instructional contexts (Derakhshan & Shakki, 2023). However, while many studies demonstrate positive instructional effects, classroom practice often lacks systematic procedures for selecting instructional targets or for evaluating gains beyond holistic judgments of appropriateness. As a result, pragmatic improvement is frequently inferred rather than operationalized through clearly defined and measurable linguistic features. Corpus-informed approaches offer a principled means of addressing both target selection and evaluation in pragmatics instruction. Research on corpus-based learning and data-driven learning has shown that corpus evidence can support language learning by drawing attention to recurring linguistic patterns (Boulton & Cobb, 2017). By comparing learner production with benchmark data from experienced L1 users, researchers and teachers can identify systematic differences in strategy frequency and distribution across contexts. Because speech act strategies can be operationalized as identifiable and countable categories, corpus comparison allows pragmatic gaps to be diagnosed systematically rather than impressionistically. Moreover, when these categories are applied consistently to pretest and posttest data, development can be measured through observable changes in strategy use, providing a practical indicator of pragmatic behavior. In this way, corpus-based strategy analysis provides a replicable framework that connects diagnosis, instruction, and quantifiable evaluation within classroom-based research.

Research Issues and Questions

This project was undertaken to address the gap between learners' pragmatic strategy use and that of proficient L1 English users. Although many learners develop adequate control of linguistic forms through classroom instruction, pragmatic features such as strategy selection, degree of intensification, and interpersonal alignment are often underrepresented in formal curricula and are expected to develop implicitly. By systematically comparing learner apology responses with those of L1 users, the study aimed to identify specific areas of pragmatic divergence and to use these findings to inform targeted instructional design. Guided by this aim, the study addressed the following research question:

To what extent does corpus-informed, strategy-focused instruction increase first-year Japanese university students' use of English apology strategies?

Method

The study consisted of two phases. Phase 1 established instructional targets through a corpus-based comparison of apology strategy use in L1 and L2 data. Phase 2 evaluated a classroom intervention using a pretest–posttest design with a test group and a control group.

All data were elicited using a written discourse completion task (DCT), in which participants read short situation descriptions and wrote what they would say in each context, administered via Google Forms. Although DCTs have been criticized for not fully capturing the dynamic and interactive nature of authentic conversation, they remain one of the most widely used methods in interlanguage pragmatics because they allow researchers to systematically control situational variables and elicit comparable responses across large groups of participants (Golato, 2003; Ogiermann, 2018). This level of control is particularly valuable when the aim is to compare the frequency and distribution of speech act strategies across groups and contexts rather than to examine conversational interaction itself. Recent reviews have also concluded that, despite their limitations, DCTs remain an appropriate and effective method for investigating speech act realization when interpreted alongside an understanding of their methodological constraints (Timpe-Laughlin & Youn, 2021).

Participants and the Research Context

Three participant groups contributed data. First, the L1 benchmark corpus was compiled in 2024 and consisted of written apology responses from 60 English L1 users. Participants were native speakers of English from the United Kingdom, the United States, and Australia. All were university English-language lecturers working in Japan with more than five years of teaching experience. This corpus was treated as a pragmatic benchmark representing experienced, contextually appropriate English apology behavior. Second, the L2 learner corpus was collected in 2025 and consisted of written apology responses from 60 first-year Japanese university students majoring in English, who had approximately six years of formal English study. To reduce the influence of near-native English ability, students who had lived in an English-speaking country for more than one year or who had a fluent English-speaking parent were excluded from the study. Third, the instructional phase involved a separate cohort of 45 first-year university students who did not participate in the Phase 1 corpus collection but were drawn from the same institutional context. Participants were divided into a test group ($n = 24$) and a control group ($n = 21$). Both groups were enrolled in first-year English courses within English-related programs and had comparable educational backgrounds. All participants had studied English for approximately six years in secondary school prior to entering university. Because the study was conducted in an intact program

setting, group assignment followed class-based administrative constraints rather than strict randomization. As a result, the design is best characterized as quasi-experimental and produced slightly unequal group sizes (test group $n = 24$; control group $n = 21$).

Data Collection Instruments and Procedures

The DCT consisted of three apology situations designed to vary in relative power, social distance, and imposition between the interlocutors (Brown & Levinson, 1987). Each prompt provided a short description of the setting and the relationship between the interlocutors, followed by a brief exchange in which participants wrote what they would say. The full situation prompts are provided in Appendix A.

Situation 1 (low power, distance, and imposition) involved forgetting to reply to a close friend's movie invitation and apologizing the next day. Situation 2 (moderate power, distance, and imposition) involved arriving late to a study meeting with an unfamiliar classmate shortly before a deadline. Situation 3 (high power, distance, and imposition) involved giving incorrect change at a part-time job and responding to a manager's correction.

For Phase 2, both the test group and the control group completed the DCT as a pretest. The test group then received a 90-minute online lesson via Zoom, while the control group received no treatment. Immediately afterward, both groups completed the same DCT as a posttest. Using an immediate posttest allowed the study to measure short-term changes in strategy use following instruction but did not assess long-term retention.

Data Analysis Instruments and Procedures

Responses were analyzed using 14 apology strategy categories based on Prachanant's (2016) taxonomy and expanded inductively through iterative coding of the data by the authors. The categories represent pragmatic strategies commonly found in English apologies: address, alerter, apology, denial of intent, empathy, exclamation, forbearance, gratitude, humor, intensifier, offer to repair, reason, self-blame, and solidarity. Coding was conducted at the strategy level. Each response was coded for the presence or absence of each strategy category. Because strategy annotation is time-consuming, a hybrid coding process was adopted. A large language model (GPT-5) was used as a first-pass annotator to flag potential instances of each strategy category within responses, which were then verified through full human review. This process followed a human-in-the-loop workflow described by Meizlish and Ziffo (2026) and Yu et al. (2024). Two human researchers then reviewed all annotations, resolved disagreements through discussion, and finalized the coding. This workflow was used to reduce processing time while maintaining human oversight for interpretive decisions.

For Phase 1 (diagnostic benchmarking), strategy frequencies were calculated as the percentage of responses that contained at least one instance of each strategy across the three situations. Percentages allowed comparison across groups of different sizes and allowed pragmatic differences to be operationalized through observable differences in strategy frequency.

For Phase 2 (instructional evaluation), analyses focused on the three target strategies: forms of address, exclamation words, and intensifiers. For each situation, pretest and posttest percentages were compared within each group, and movement toward the L1 benchmark was interpreted as positive development. Because the goal was pedagogical diagnosis and evaluation of a short intervention, results are reported descriptively.

Instructional Treatment

The instructional treatment consisted of a 90-minute online lesson delivered via Zoom. The lesson was designed to address pragmatic strategies identified as underused in the corpus comparison with the L1 benchmark corpus. Instruction focused on three target strategies: forms of address, exclamation words, and intensifiers.

The lesson began with awareness-raising activities in which students examined sample apology responses and discussed how small linguistic choices shaped interpersonal stance in apologies. Students then completed guided brainstorming tasks to expand their repertoire of intensifiers (e.g., really, so, extremely, terribly), exclamation words (e.g., Oh no, Oh, Ah), and address forms (e.g., first name, title plus family name). To link form to context, the class discussed when each strategy is common in English and how its appropriateness may change with power, distance, and imposition.

Next, students practiced with controlled role-play activities based on the three DCT situations. The role plays provided opportunities to incorporate the target strategies while allowing learners to focus on appropriate use in context with a manageable cognitive load. The lesson concluded with a short review activity designed to consolidate learning and reinforce the target strategies.

Results

Phase 1: Benchmark Comparison of L1 and L2 Strategy Use

Table 1 summarizes the preliminary corpus comparison. Overall, first-year students produced substantially fewer forms of address, exclamation words, intensifiers, and reasons than L1 users. These gaps were used to select instructional targets for Phase 2. Because reasons are highly context-specific, they were excluded from the instructional focus, which instead targeted address forms, exclamation words, and intensifiers.

Table 1*Differences Between L1 and L2 Users in Apology Strategy Use*

Strategy	L1 benchmark (%)	First-year students (%)
Address	66	24
Alerter	12	3
Apology	91	86
Denial of intent	9	11
Empathy	13	15
Exclamation	17	4
Forbearance	33	31
Gratitude	4	3
Humor	4	2
Intensifier	67	39
Offer to repair	55	50
Reason	51	25
Self-blame	65	60
Solidarity	2	0

Note. Percentages are used to allow comparison between groups of different sizes and across the three situations.

Phase 2: Instructional Effects by Situation

The following sections report pretest and posttest results for the three target strategies. The test group completed the strategy-focused lesson between the pretest and posttest, whereas the control group did not.

Situation 1: not replying to a good friend's message. Situation 1 represented a low power, distance, and imposition context. The test group showed large gains in intensifier use and meaningful gains in address forms and exclamation words. The control group showed only small changes and decreased use of exclamation words.

Table 2*Target Strategy Use in Situation 1*

Strategy	Test pre	Test post	Control pre	Control post	L1 benchmark
Intensifier	38	75	38	43	72
Exclamation	12	17	12	5	38
Address	37	54	37	48	65

Note. Values represent the percentage of responses containing the strategy.

Situation 2: arriving late for a classwork meeting with a classmate. Situation 2 represented a moderate power, distance, and imposition context. The test group again increased intensifier

use substantially and showed a small increase in exclamation words. Both groups remained far below the L1 benchmark for address forms.

Table 3

Target Strategy Use in Situation 2

Strategy	Test pre	Test post	Control pre	Control post	L1 benchmark
Intensifier	52	79	52	48	83
Exclamation	0	4	0	0	13
Address	27	29	27	29	77

Note. Values represent the percentage of responses containing the strategy.

Situation 3: making a cash-handling mistake at work. Situation 3 represented a high power, distance, and imposition context. The L1 benchmark showed no exclamation words in this situation. The test group increased intensifier use and address forms, while the control group showed only small gains.

Table 4

Target Strategy Use in Situation 3

Strategy	Test pre	Test post	Control pre	Control post	L1 benchmark
Intensifier	27	54	27	29	47
Exclamation	0	0	0	0	0
Address	10	29	10	14	57

Note. Values represent the percentage of responses containing the strategy.

Summary of Results

Across all three situations, the test group increased use of the target strategies on the posttest and consistently moved closer to the L1 benchmark. Average intensifier use increased from 39% to 69% (+30 percentage points) in the test group, compared with 39% to 40% (+1 point) in the control group. Average address use increased from 25% to 37% (+12 points) in the test group, compared with 25% to 30% (+5 points) in the control group. Exclamation words showed smaller gains in the test group (4% to 7%, +3 points) and decreased in the control group (4% to 2%, -2 points).

Table 5

Average Target Strategy Use Across the Three Situations

Strategy	Test pre	Test post	Control pre	Control post	L1 benchmark
Intensifier	39	69	39	40	67
Exclamation	4	7	4	2	17
Address	25	37	25	30	66

Note. Values are arithmetic means of the three situation-specific percentages reported in Tables 2–4.

Discussion

This study set out to (a) identify pragmatic gaps in first-year Japanese university students' apology production through a corpus comparison and (b) evaluate whether explicit, strategy-focused instruction could reduce those gaps. Using a corpus-informed strategy-counting framework, the study first diagnosed systematic differences between learner production and an L1 benchmark and then examined whether targeted instruction could produce measurable changes in strategy use. Overall, the pattern of results suggests that targeted instruction can increase learners' use of specific apology strategies. Following a single instructional session, the test group increased the frequency of the three target strategies across all contexts, whereas the control group showed little change. These results suggest that corpus-informed strategy analysis can be used both to guide instructional targeting and to track changes in pragmatic strategy use through observable shifts in strategy frequency.

The strongest and most consistent gains were found for intensifiers. This result is compatible with noticing-based accounts of instructed pragmatics. Intensifiers are short, high-frequency items that can be readily integrated into familiar apology frames (e.g., I'm really sorry), making them relatively accessible targets for explicit instruction (Taguchi, 2015; Biber et al., 1999). Because these items can be inserted easily into existing apology structures, they may represent relatively accessible targets for short instructional interventions. The intervention explicitly drew attention to these items, expanded learners' lexical options through brainstorming, and required immediate use during role plays. Together, these activities likely increased the salience and retrievability of intensifiers during the posttest (Schmidt, 1990, 1993). This pattern suggests that corpus-identified strategy gaps involving highly frequent and structurally simple items may be particularly responsive to brief explicit instruction. At the same time, the test group exceeded the benchmark in two situations, suggesting a risk of overgeneralization. Future instruction may therefore benefit from contrasting strong and mild offenses and from highlighting that intensification is optional and context-sensitive, rather than a universal marker of politeness.

Forms of address also increased in the test group, particularly in the low power, low distance, and low imposition situation and the high power, high distance, and high imposition situation. However, address use remained substantially below the L1 benchmark, especially in Situation 2 and Situation 3. One explanation is sociopragmatic transfer. In Japanese, address practices and honorific systems differ from English norms (Braun, 1988; Sugimoto, 1998), and learners may avoid direct naming even when it is pragmatically routine in English.

As a result, learners may hesitate to include names or titles in situations where English speakers would normally use them to signal interpersonal alignment. Another possible explanation is a task effect. Because the DCT was written, some learners may have prioritized the core apology move and omitted optional openers such as names and titles. Written responses may therefore underrepresent address strategies that would be more likely to appear in spoken interaction. These points suggest that address forms may require more extensive input, modeling, and feedback than intensifiers, including attention to how address choices index power, distance, and imposition in different interactional contexts (Brown & Levinson, 1987).

Exclamation words showed the smallest instructional gains. In Situation 1, exclamations were common in the L1 benchmark, but learners remained well below this level even after instruction. Exclamations may be less transparent as a pedagogical target because they are brief, may appear informal, and often function through prosody and interactional stance rather than through clear lexical meaning (Ameka, 1992). As a result, learners may not readily recognize them as part of the pragmatic structure of an apology. In addition, the benchmark indicated that exclamations were not used in the workplace situation, which may have reinforced an interpretation that exclamations are limited to more informal contexts. For teacher practice, this pattern suggests that exclamations may benefit from audio-visual input and prosody-focused modeling, rather than relying solely on written examples.

From a teacher development perspective, the study illustrates a workflow for moving from diagnosis to instruction. Corpus comparison first identified systematic gaps in learner strategy use, which then informed the selection of instructional targets. Rather than attempting to teach all apology strategies, the intervention selected a small set of high-frequency, pedagogically accessible features based on corpus evidence. This aligns with recommendations that classroom pragmatics instruction should be selective, contextualized, and supported by metapragmatic reflection (Ishihara & Cohen, 2010; Taguchi, 2015). In addition to its pedagogical implications, the study also demonstrates a practical methodological approach for classroom-based research. The use of a hybrid AI and human coding process enabled strategy-level analysis at a scale that would be difficult for an individual teacher-researcher to complete alone. While such tools should not replace expert judgment, they can extend the analytic capacity of teacher-researchers when used with careful human review.

Limitations and Future Directions

Several limitations should be acknowledged. First, the instructional study involved a relatively small number of participants (24 students in the test group and 21 in the control group), which limits generalizability. Second, the intervention consisted of a single 90-minute lesson and the posttest was administered immediately afterward. The results therefore reflect short-term instructional responsiveness rather than longer-term pragmatic development, and do not provide evidence of retention over time.

A further limitation concerns the analytic focus of the study. The analysis examined whether strategies were present, rather than evaluating their appropriateness, naturalness, or interactional timing. Learners may increase strategy frequency without fully mastering how to deploy a strategy in a contextually appropriate and interactionally nuanced way. In addition, the DCT format provides an efficient elicitation method but primarily taps participants' reported responses in hypothetical situations rather than real-time interactional competence (Golato, 2003; Kasper & Dahl, 1991).

Future research should examine longer instructional sequences, include delayed posttests, and incorporate interactional data such as recorded role plays or naturally occurring interaction. Qualitative analysis of learner language could also help clarify whether increased strategy use reflects improved sociopragmatic judgment or simply the addition of lexical elements identified during instruction. Future studies could also apply the corpus-informed diagnostic approach used in this study to other speech acts, which would further test its value for systematic classroom-based pragmatics research.

Conclusion

This study used a corpus-based comparison to identify pragmatic gaps in first-year Japanese university students' English apologies and then evaluated a targeted instructional response. Benchmarking against L1 data indicated that learners underused several strategies, most notably forms of address, exclamation words, and intensifiers. Using these corpus-identified gaps to guide instruction, a single lesson designed around awareness raising and controlled practice produced clear increases in the three target strategies for the test group, with the largest gains observed for intensifiers and forms of address.

These findings suggest that a corpus-informed strategy-counting framework can be used both to guide instructional target selection and to track changes in learners' use of specific pragmatic strategies. For teacher development, the study highlights the practicality of combining corpus-informed diagnosis, targeted instruction, and strategy-level analysis in classroom research. This approach offers a replicable model for teachers who wish to

investigate pragmatic development in their own instructional contexts using systematic and measurable methods.

References

- Ameka, F. (1992). Interjections: The universal yet neglected part of speech. *Journal of Pragmatics*, 18(2-3), 101-118. [https://doi.org/10.1016/0378-2166\(92\)90048-G](https://doi.org/10.1016/0378-2166(92)90048-G)
- Bardovi-Harlig, K. (2013). Developing L2 pragmatics. *Language Learning*, 63(Suppl. 1), 68-86.
- Bergman, M. L., & Kasper, G. (1993). Perception and performance in native and nonnative apology. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 82-107). Oxford University Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Longman.
- Blum-Kulka, S., & Olshtain, E. (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, 5(3), 196-213. <https://doi.org/10.1093/applin/5.3.196>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348-393. <https://doi.org/10.1111/lang.12224>
- Braun, F. (1988). *Terms of address: Problems of patterns and usage in various languages and cultures*. Mouton de Gruyter.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Derakhshan, A., & Shakki, F. (2023). *Instructed second language pragmatics for the speech acts of request, apology, and refusal: A meta-analysis*. Springer. <https://doi.org/10.1007/978-3-031-37093-9>
- Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics*, 24(1), 90-121. <https://doi.org/10.1093/applin/24.1.90>
- Ishihara, N., & Cohen, A. D. (2010). *Teaching and learning pragmatics: Where language and culture meet*. Pearson Longman.
- Jeon, E. H., & Kaya, T. (2006). Effects of L2 instruction on interlanguage pragmatic development: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 165-211). John Benjamins.
- Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, 13(2), 215-247.
- Kasper, G., & Rose, K. R. (2002). *Pragmatic development in a second language*. Blackwell.
- Leech, G. N. (1983). *Principles of pragmatics*. Longman.

- Meizlish, T., & Ziffo, C. (2026). *Generative AI and pragmatics: Implications for research, teaching, and assessment*. Cambridge University Press.
- Ogiermann, E. (2018). Discourse completion tasks. In A. H. Jucker, K. P. Schneider, & W. Bublitz (Eds.), *Methods in pragmatics* (pp. 229-255). De Gruyter Mouton.
- Olshtain, E., & Cohen, A. D. (1983). Apology: A speech act set. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 18-35). Newbury House.
- Prachanant, N. (2016). A cross-cultural study of apology speech act realizations. *International Journal of Languages, Literature and Linguistics*, 2(3), 146-151. <https://doi.org/10.18178/IJLL.2016.2.3.85>
- Sánchez-Hernández, A., & Martínez-Flor, A. (2022). Teaching the pragmatics of English as an international language: A focus on pragmatic markers. *Language Teaching Research*, 26(2), 256-278. <https://doi.org/10.1177/13621688211064933>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158. <https://doi.org/10.1093/applin/11.2.129>
- Schmidt, R. (1993). Consciousness, learning, and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21-42). Oxford University Press.
- Sugimoto, N. (1998). Norms of apology depicted in U.S. American and Japanese literature on manners and etiquette. *International Journal of Intercultural Relations*, 22(3), 251-276. [https://doi.org/10.1016/S0147-1767\(98\)00011-0](https://doi.org/10.1016/S0147-1767(98)00011-0)
- Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48(1), 1-50.
- Takahashi, S. (2010). Assessing learnability in second language pragmatics. In A. Trosborg (Ed.), *Pragmatics across languages and cultures* (pp. 391-421). De Gruyter Mouton.
- Timpe-Laughlin, V., & Youn, S. J. (2021). Measuring L2 pragmatics. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 254-264). Routledge.
- Vellenga, H. (2004). Learning pragmatics from ESL & EFL textbooks: How likely? *TESL-EJ*, 8(2), 1-18.
- Yu, D., Li, L., Su, H., & Fuoli, M. (2024). Assessing the potential of LLM-assisted annotation for corpus-based pragmatics and discourse analysis: The case of apology. *International Journal of Corpus Linguistics*, 29(4), 534-561. <https://doi.org/10.1075/ijcl.23087.yu>

Appendix A

Discourse Completion Task Situations

Situation #	Situation	Power	Distance	Imposition
1	You did not reply to a close friend's movie invitation last night. You see your friend the next day and say:	Low	Low	Low
2	You arrive one hour late to a study meeting with an unfamiliar classmate shortly before a deadline. You see the classmate and say:	Moderate	Moderate	Moderate
3	You give incorrect change to a customer at work and are corrected by your manager. You respond to your manager and say:	High	High	High

The Relationship Between GenAI-Writerism and English Writing Self-Efficacy in a Japanese University Context: A Pilot Study

Naoya SHIBATA¹ and Vivian TURK²

¹ *Osaka Kyoiku University*

² *Nagoya City University*

Abstract

Recent development in generative artificial intelligence (GenAI) has facilitated research on GenAI use, AI literacy, and second-language (L2) writing self-efficacy. However, existing constructs, such as AI literacy, trust in technology, and perceived usefulness, do not fully explain how L2 writers position GenAI output as a source of linguistic authority in writing. To address this gap, this study explores generative artificial intelligence writerism (GenAI-Writerism), conceptualised as a writing-specific belief orientation in which learners regard GenAI output and feedback as appropriate, reliable, and efficient resources for producing linguistically accurate extended writing (Shibata, 2026). Based on quantitative data collected through six-point Likert-scale questionnaire items and analysed using Kendall's tau correlation coefficient, this exploratory pilot study examines preliminary correlations between proposed GenAI-Writerism dimensions and L2 writing self-efficacy in a Japanese English-as-a-foreign-language tertiary context. The study also considers pedagogical implications for preventing overreliance on GenAI use and for promoting balanced GenAI-Writerism and L2 writing self-efficacy.

Keywords: generative artificial intelligence, L2 writing self-efficacy, correlation

The Relationship Between GenAI-Writerism and English Writing Self-Efficacy in a Japanese University Context: A Pilot Study

Recent development in generative artificial intelligence (GenAI) has substantially altered how second language (L2) writing is produced, evaluated, and learned in tertiary educational contexts. While earlier GenAI tools were often perceived as limited or unreliable for L2 writing, recent GenAI systems can now generate linguistically fluent and structurally standard text that resembles Inner Circle academic norms (Lee et al., 2025). As a result, concerns about overreliance have increased, and AI literacy has been widely discussed as one response to these challenges. However, beyond skills and guidelines, learners' beliefs about the reliability and appropriateness of GenAI output play a crucial role in shaping how GenAI is actually used during writing. Trust in GenAI as a source of linguistically accurate and efficient writing support may influence whether learners critically evaluate GenAI suggestions or accept them unreflectively. Shibata (2025, 2026) conceptualised this belief orientation as generative artificial intelligence writerism (GenAI-Writerism), defined as the belief that GenAI output is appropriate and reliable for efficiently producing extended writing without linguistic errors.

GenAI-Writerism is related to, but distinct from, several existing constructs. AI literacy concerns learners' knowledge, skills, and critical capacity to understand and use AI tools appropriately. By contrast, GenAI-Writerism concerns learners' beliefs about the reliability, appropriateness, and authority of GenAI output in writing. Trust in GenAI also overlaps with this construct, but trust is often discussed as a broader confidence in the system or its output. GenAI-Writerism is more writing-specific because it focuses on whether L2 writers treat GenAI-generated language, feedback, and revisions as more appropriate or reliable than their own linguistic judgement. Technology acceptance models, including perceived usefulness and perceived ease of use, help explain why learners may intend to use GenAI, but they do not fully explain how learners position GenAI as a writerly authority during text production and revision. Therefore, GenAI-Writerism offers a writing-specific lens for examining how learners evaluate GenAI-supported writing, especially in relation to linguistic accuracy, appropriateness, feedback, and authorial agency.

L2 writing self-efficacy may further interact with this belief system, shaping how learners judge the usefulness and authority of GenAI output. While prior studies have examined GenAI use, writing self-efficacy, and AI literacy separately, the belief system conceptualised as GenAI-Writerism and its relationship with writing self-efficacy remain underexplored. In particular, little empirical research has examined whether strong trust in

GenAI output aligns with, or operates independently of, learners' perceived writing competence. Drawing on quantitative data collected from Japanese EFL university students through a six-point Likert-scale questionnaire, this exploratory pilot study examines preliminary correlations between proposed GenAI-Writerism dimensions and L2 writing self-efficacy, as well as relationships among those proposed dimensions. The study does not aim to validate GenAI-Writerism as a fully established construct; rather, it provides an empirical step for identifying possible relationships that can guide future questionnaire refinement, qualitative investigation, and classroom-based research.

Literature Review

This literature review examines how GenAI shapes learners' L2 writing self-efficacy through trust in writing support and reliance on GenAI feedback.

L2 Writing Self-Efficacy and GenAI

L2 writing self-efficacy refers to learners' beliefs in their capacity to perform writing tasks successfully and to manage the linguistic, cognitive, and affective demands of writing. Recent research indicates that GenAI tools can considerably reshape these beliefs, though effects are complex and uneven. A systematic review of 112 empirical studies on GenAI-integrated L2 writing found that GenAI use often increases learners' awareness of grammatical accuracy, lexical choice, and textual organization, which can contribute positively to writing confidence (Feng et al., 2025). However, the same review reported that increased exposure to GenAI texts and feedback may also intensify learners' self-criticism, particularly when GenAI output is perceived as a benchmark for ideal writing, resulting in reduced writing self-efficacy for some students (Feng et al., 2025). In other words, learners may not only use GenAI as a support tool; they may also compare their own writing with GenAI text. When GenAI output appears more fluent, accurate, or academically appropriate than their own writing, some learners may begin to evaluate their own ability less positively. These findings suggest that self-efficacy is shaped not simply by access to GenAI, but by how learners interpret its output.

Qualitative research further highlights the affective dimensions of GenAI-assisted writing. In a study of Chinese university students using a ChatGPT-embedded writing tool, many learners reported reduced anxiety and increased confidence because GenAI provided immediate, non-judgmental support throughout the writing process (Kim et al., 2024). At the same time, students struggled to evaluate GenAI suggestions critically, especially regarding tone, argumentation, and cultural appropriateness (Kim et al., 2024). This suggests that emotional gains in self-efficacy may coexist with heightened risk of overreliance. Similar

tensions are evident in other EFL contexts. Nelson et al. (2025), for example, found that Ecuadorian EFL undergraduates perceived GenAI as enabling efficient task completion while simultaneously expressing concern that heavy reliance on GenAI could undermine long-term writing development and create a sense of “pseudo-success” (p. 15). Students were more concerned about losing writing ability than about institutional sanctions, indicating awareness that confidence built through GenAI dependence may be unstable.

From a broader technology adoption perspective, Davis’s (1989) Technology Acceptance Model (TAM) identifies perceived usefulness and perceived ease of use as key factors shaping users’ attitudes towards a technology and their willingness to engage with it. In GenAI-supported writing, these perceptions are relevant because learners may rely on GenAI when they see it as useful, efficient, and easy to access. However, TAM mainly explains acceptance and intention to use technology. It does not fully explain learners’ epistemic beliefs about writing authority, such as whether GenAI feedback should be trusted more than peer feedback, teacher feedback, or their own linguistic judgment. This is where GenAI-Writerism becomes useful; it extends the discussion from technology acceptance to learners’ beliefs about GenAI as a source of writing support and linguistic authority. Together, the studies reviewed here suggest that GenAI can both enhance and undermine L2 writing self-efficacy. Confidence gains are often linked to immediate support, surface-level correction, and emotional reassurance, whereas declines in self-efficacy may occur when learners compare their own writing unfavourably with GenAI output.

GenAI in L2 Writing: Affordances, Risks, and Trust

Beyond effects on performance and confidence, recent research highlights trust in GenAI as a key factor shaping writing decisions and learning outcomes. Research consistently shows that students use GenAI across the entire L2 writing process, including idea generation, planning, drafting, revising, and proofreading. High levels of reliance are reported during revision stages, where GenAI feedback is perceived as particularly useful (Kim et al., 2024). Students commonly associate GenAI use with improved fluency, vocabulary range, grammatical accuracy, and overall text quality (Feng et al., 2025; Kim et al., 2026).

Quantitative studies applying technology acceptance models further demonstrate that students’ intention to use GenAI writing assistants is shaped primarily by performance expectancy, social influence, affect-based motivation, learning value, and trust, while ease of use and price value do not significantly predict acceptance (Kim et al., 2026). This can suggest that belief-based and affective factors outweigh technical considerations. Despite

these perceived benefits, multiple studies warn against uncritical reliance on GenAI. Large language models generate text through probabilistic prediction rather than grounded understanding, resulting in hallucinated information, fabricated references, and confident misinformation (Sterling et al., 2025). Trust, therefore, becomes a central concern. Conceptual analyses argue that GenAI challenges traditional definitions of trust by obscuring the identity of the agent, the source of knowledge, and the nature of risk, and thus trust in GenAI is necessarily indirect, extending to developers and training data that embed cultural and linguistic biases (Wolf et al., 2024).

In L2 writing contexts, trust directly affects whether learners revise, accept, or question GenAI output, shaping both writing processes and perceived authorship. Studies of international students show that GenAI may produce more formal, assertive, and standardised prose than learners intend, sometimes exaggerating claims or altering meaning through inappropriate lexical choices (Sterling et al., 2025). When learners lack the linguistic awareness to identify these shifts, they may accept GenAI output as more reliable than their own, increasing dependence and weakening their authorial voice (Sterling et al., 2025). Feedback-focused research further illustrates this risk. Hyland (2025), for example, demonstrates that while GenAI feedback is effective for surface-level corrections, it is limited in addressing rhetorical stance, reader engagement, and interpersonal meaning. Therefore, trust in GenAI feedback does not necessarily lead to writing development unless learners engage critically with the feedback.

GenAI-Writerism, Overreliance, and Writing Agency

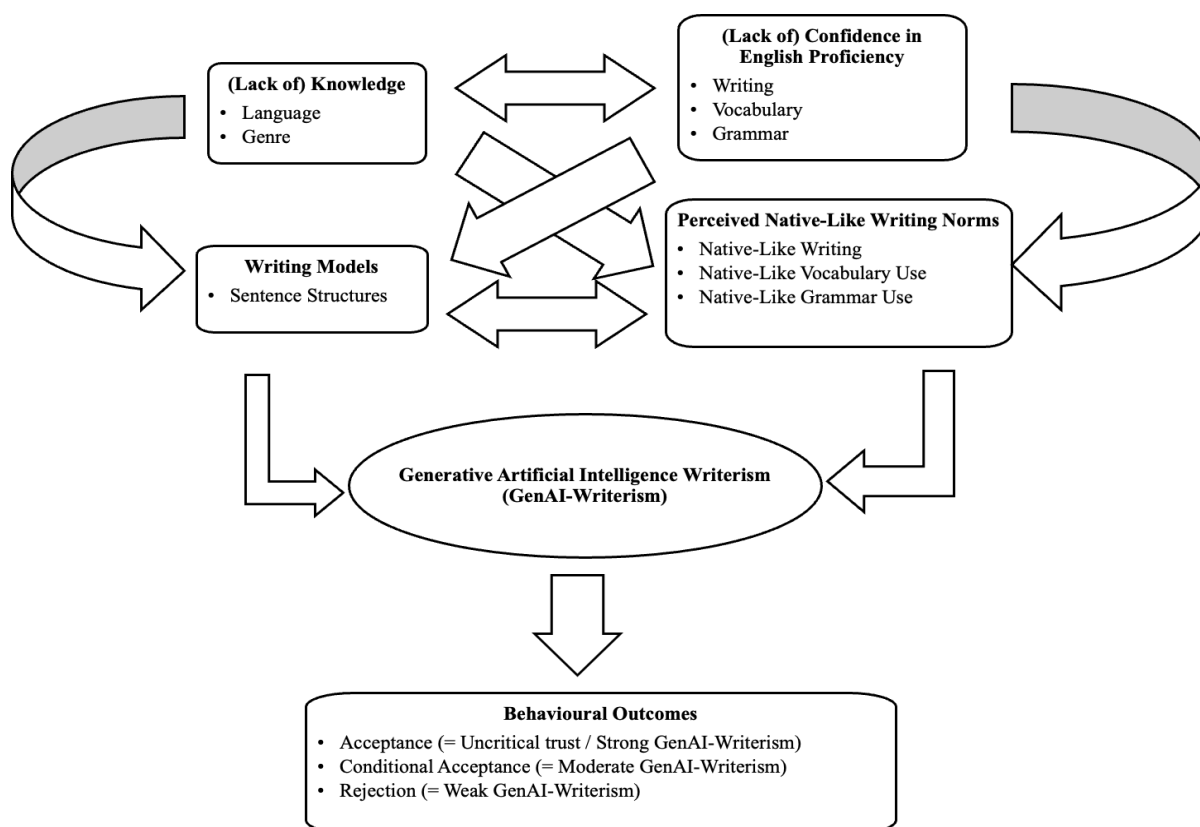
Although trust in GenAI technology can occur in many learning situations, Shibata (2025, 2026) proposed the term generative artificial intelligence writerism (GenAI-Writerism) to describe a writing-specific belief orientation in L2 writing. In this study, GenAI-Writerism is defined as a multidimensional belief orientation through which L2 writers evaluate GenAI as a source of writing support, linguistic authority, feedback, and learning assistance. It includes learners' beliefs about the linguistic quality of GenAI output, their trust in GenAI feedback, their perceived usefulness of GenAI for completing writing tasks, their tendency to rely on or accept GenAI suggestions, and their perceived capacity to use GenAI reflectively for writing development. In simple terms, GenAI-Writerism reflects how much learners trust GenAI output as a better or more reliable source of writing than their own judgment. Shibata (2026) developed the term from findings showing that some Japanese EFL university students followed GenAI feedback on language use with limited critical

evaluation because they perceived GenAI feedback as more concise and correct than peer feedback, and GenAI output as superior to their own writing.

Shibata's (2026) qualitative study also revealed that students with low L2 writing self-efficacy tended to believe that GenAI output provides linguistically appropriate use of language to help them reach their perceived native-like writing norms. Based on these findings, L2 writing self-efficacy and perceived native-like writing norms were considered a significant influence on students' trust in GenAI output, and a described interconnection between possible factors in GenAI-Writerism (Figure 1). However, because GenAI-Writerism is a newly proposed construct, it has not yet been fully theorised or empirically validated. Further exploratory studies are therefore needed to identify its possible dimensions, related factors, and effects on GenAI use.

Figure 1

Shibata's (2026) Conceptual Model of the Relationship Between GenAI-Writerism and its Possible Factors



Existing literature suggests that trust in GenAI is embedded within broader motivational, affective, and identity-related processes. For example, Kim et al. (2024) report that students often perceive GenAI not merely as a tool, but as a tutor, assistant, or peer that

offers continuous support and emotional reassurance. Such perceptions may strengthen trust and encourage habitual use, particularly under time pressure or heavy academic workloads. However, studies also warn that lower-proficiency learners may depend on GenAI output without effective evaluation, using GenAI to replace rather than extend cognitive effort (Feng et al., 2025). Similarly, research on learning outcomes shows that simple unreflective use of GenAI correlates with weaker learning gains, whereas constructive and critical engagement supports deeper understanding (Pallant et al., 2026).

From an identity perspective, GenAI-Writerism intersects with issues of voice and symbolic capital. GenAI output often reflects dominant Western academic norms, privileging direct thesis statements, linear argumentation, and assertive stance (Sterling et al., 2025). When students trust these conventions uncritically, they may internalise narrow definitions of *good writing*, reinforcing unequal hierarchies of linguistic legitimacy (Darvin, 2025). At the same time, critical frameworks emphasise the need for reflexive engagement to counter these tendencies. Lin and Chen's (2025) PAA model and 4T lenses highlight translanguaging, transculturing, and agentic AI use as ways to challenge epistemic dominance and reposition learners as critical co-constructors of meaning rather than passive recipients of GenAI output. These perspectives suggest that trust in GenAI can be conditional and negotiated rather than absolute, supporting writer development when paired with critical awareness.

Despite growing interest in AI literacy (Chiu et al., 2024), few studies directly examine how GenAI-Writerism relates to L2 writing self-efficacy. While prior research documents confidence gains, anxiety reduction, and motivation, the relationship between strong beliefs in GenAI's trustworthiness and learners' perceived writing competence remains underexplored, particularly in multilingual university contexts. Therefore, the present study examines whether GenAI-Writerism is related to learners' perceived writing abilities.

Research Questions

The literature demonstrates that GenAI plays a complex role in L2 writing. It can support confidence, efficiency, and revision while it may also introduce risks of overreliance, voice loss, and reduced agency. Trust emerges as a central but multifaceted construct, involving writerly, epistemic, and sociocultural dimensions. While previous research has examined GenAI use, writing self-efficacy, and AI literacy separately, the belief system conceptualised as GenAI-Writerism and its relationship with L2 writing self-efficacy remain insufficiently explored. Addressing this gap, the present pilot study examines preliminary correlations between GenAI-Writerism and writing self-efficacy among Japanese university

students, as well as correlations between GenAI-Writerism elements. The following research question was devised:

RQ1: To what extent does GenAI-Writerism correlate with Japanese EFL university students' L2 writing self-efficacy?

RQ2: To what extent do GenAI-Writerism elements correlate with each other?

Based on the findings, this paper also aims to consider pedagogical implications for promoting balanced, critically informed engagement with GenAI technology.

Method

Participants and the Research Context

This study was conducted at a private university in central Japan. The authors distributed an online questionnaire to their colleagues and students, who were asked to introduce the study to potential participants. Therefore, a snowball sampling strategy was utilised to collect data. Information about the study was provided in Japanese at the beginning of the online questionnaire, and students were asked to indicate whether they consented to participate before answering the survey. Students did not need to write their names and email addresses, and their teachers were not allowed to access the results; thus, their anonymity was fully ensured. As a result, 96 students consented to take the questionnaire. Among 96 students, 59 majored in British and American Cultures, 16 in English Communication, six in English Education, seven in Chinese, and eight in World Liberal Arts. In English-related majors, sixteen students had short-term study abroad experiences (two weeks to two months). All students reported that they often experience English writing activities through classes and essay/report assignments, and sometimes send messages and emails to friends and teachers in English; thus, students seem to engage in English writing activities for academic and general purposes. As this study does not aim to consider academic majors or studying or living abroad experience as important factors in GenAI-Writerism, all respondents were considered participants to explore correlations between GenAI-Writerism and their L2 writing self-efficacy.

Data Collection and Analysis Methods and Procedures

In this study, a six-point Likert scale questionnaire was utilised to examine students' English writing self-efficacy and GenAI-Writerism. The first section on English writing self-efficacy was designed based on the Common European Framework of Reference for Languages-Japan (CEFR-J) version 1.0 (Tono, 2013). CEFR-J categorises language proficiency and skills into seven levels: pre-A1, A1, A2, B1, B2, C1, and C2 (Common European Framework of Reference for Languages-Japan, 2025; Tono, 2013, 2017, 2019).

General writing activities considered appropriate in each level are described in the framework (see Tono, 2013). As the pre-A1 level of writing proficiency is considered preparatory to be able to write the alphabet, and all respondents were university students, this proficiency level was excluded from the questionnaire. Furthermore, only one statement is provided for C1 and C2 writing proficiency levels. In order to estimate item reliability, two or more question items are necessary, and thus, the item reliability of C1 and C2 was not examined in this study. As Cronbach alpha is criticised and McDonald's omega is recommended to utilise instead (Hayes & Coutts, 2020; McNeish, 2018), the item reliability for each level was estimated with McDonald's Omega (ω) through JASP (A1: $\omega = 0.90$, A2: $\omega = 0.92$, B1: $\omega = 0.94$, B2: $\omega = 0.94$). Based on Cohen et al.'s (2018, p. 774) benchmark (Table 1), omega scores for all levels were sufficient to judge the instrument as very highly reliable. All scales met accepted reliability thresholds. As this is a pilot study, the questionnaire and its dimensions are treated as exploratory rather than fully validated.

Table 1

Cohen et al.'s (2018, p. 774) Benchmark for Internal Consistency

> 0.90	very highly reliable
0.80–0.90	highly reliable
0.70–0.79	reliable
0.60–0.69	marginally/minimally reliable
< 0.60	unacceptably low reliability

The second section attempts to examine GenAI-Writerism. Authors collaboratively designed this section through online meetings and discussed which questions to include in the instrument (see Appendix). As a result, the first version of the questionnaire included seven perspectives: (1) complexity, (2) accuracy, (3) fluency, (4) appropriateness, (5) feedback, (6) trust and credibility (TR), and (7) reflective use (RU). Example items included “I believe Generative AI makes complex sentences well” for complexity, “I think Generative AI makes fewer grammar mistakes than I do” for accuracy, “generative AI helps me express my thoughts smoothly” for fluency, “I trust Generative AI to adapt its language to the writing task” for appropriateness, “I follow Generative AI feedback even when I am unsure about it” for feedback, “I feel that Generative AI is a trustworthy writing assistant tool” for trust and credibility, and “I use Generative AI as a tool to support my learning, not to replace my ideas” for reflective use. These seven perspectives were designed to capture GenAI-Writerism as a provisional multidimensional construct. Complexity, accuracy, fluency, and appropriateness represent learners' beliefs about the linguistic and rhetorical quality of GenAI output. Feedback and trust/credibility represent learners' confidence in GenAI as a source of revision

support. Reflective use was included to examine whether learners perceived GenAI as a resource for learning rather than merely as a replacement for their own writing decisions. As the scale was newly developed, these dimensions should be interpreted as tentative and exploratory. Three or four questions were assigned to each element, and item reliability was estimated to judge whether the questionnaire was usable: (complexity: $\omega = 0.76$, accuracy: $\omega = 0.71$, fluency: $\omega = 0.74$, appropriateness: $\omega = 0.77$, feedback: $\omega = 0.78$, TR: $\omega = 0.77$, RU: $\omega = 0.72$). As the 0.70 and larger is considered acceptable item reliability (Cohen et al., 2018; Dörnyei & Dewaele, 2023; Plonsky & Derrick, 2016), statements on CEFR-J and GenAI-Writerism were judged as usable in this study. As this was a pilot study, item reliability was examined, but exploratory and confirmatory factor analyses were not conducted. Therefore, the seven dimensions were treated as provisional categories rather than validated factors.

Descriptive statistics and correlational statistical testing were performed through JASP 0.95.4, a free statistical analysis software. Descriptive statistics were conducted to understand the general distribution of the datasets, means, skewness, and kurtosis. In addition, Shapiro-Wilk tests were used to judge whether parametric or nonparametric inferential statistics would be appropriate in this study. After that, correlation testing was conducted to examine the correlation between GenAI-Writerism and L2 writing self-efficacy and between GenAI-Writerism elements. In this study, Kendall's tau rank correlation coefficient (τ) was adopted because of Shapiro-Wilk test results, which are reported in the Results and Discussion section, and small sample sizes. Based on Cohen et al.'s (2018) benchmark of correlation coefficient, the strength of correlation was judged in this study: $0 < |\tau| < 0.1$ (weak), $0.1 \leq |\tau| < 0.3$ (modest), $0.3 \leq |\tau| < 0.5$ (moderate), $0.5 \leq |\tau| < 0.8$ (strong), $0.8 \leq |\tau| \leq 1.0$ (very strong). Through these procedures, the general degree of correlation between GenAI-Writerism and L2 writing self-efficacy and between GenAI-Writerism elements was estimated to consider factors in GenAI-Writerism and pedagogical implications.

Results and Discussion

Descriptive Statistics Results

Descriptive statistics indicated that, in all CEFR-J levels, except for the B1 level, no normal distribution was estimated based on noticeable skewness and kurtosis, as well as p-values of Shapiro-Wilk test results (Table 2). As a six-point Likert scale questionnaire was used in this study, three or lower can be perceived as low L2 writing self-efficacy, whereas four and higher can be considered as high L2 writing self-efficacy. Based on the results,

respondents had slightly higher confidence in A1 and A2 level writing activities (A1: Mean = 4.61, A2: Mean = 4.19), but not in the other higher level writing activities (B1: Mean = 3.54, B2: Mean = 2.97, C1: Mean = 2.81, and C2: Mean = 2.58). These patterns broadly align with Koizumi et al.'s (2022) findings on Japanese EFL learners' writing proficiency profiles.

Table 2

Descriptive Statistics Results: English Writing Self-Efficacy Based on CEFR-J

	A1	A2	B1	B2	C1	C2
Means	4.61	4.19	3.54	2.97	2.81	2.58
(SD)	(1.07)	(1.17)	(1.20)	(1.23)	(1.36)	(1.34)
Skewness	-0.96	-0.42	-0.09	0.27	0.27	0.54
Kurtosis	1.33	-0.19	-0.25	-0.52	-0.81	-0.51
p-values of Shapiro-Wilk	< 0.001	0.010	0.16	0.024	< 0.001	< 0.001

Note. SD: Standard Deviation

On the other hand, assuming from the Shapiro-Wilk test results, approximately normal distributions were identified in all GenAI-Writerism elements, except for fluency. As shown in Table 3, mean scores ranged from 3.25 to 4.12, which might indicate that respondents had a somewhat neutral degree of overall GenAI-Writerism in this study. Assuming from statistical results, students had slightly higher GenAI-Writerism regarding complexity (Mean = 4.02), accuracy (Mean = 4.07), and RU (Mean = 4.12) than the other elements (fluency: Mean = 3.96, appropriateness: Mean = 3.53, feedback: Mean = 3.32, T&C: Mean = 3.25). As Shibata (2026) reports that students, particularly those who had low perceived L2 writing self-efficacy, commented that GenAI output would help them develop linguistic complexity and accuracy in their writing products, and they sometimes overused GenAI technology; these statistical results might indicate that respondents tended to value the linguistic complexity and accuracy of GenAI output and the accessibility of GenAI functions for L2 writing activities. Using these descriptive statistical results (Tables 2 & 3), correlations between L2 writing self-efficacy and GenAI-Writerism elements are estimated in order to address research questions and theorise GenAI-Writerism and its factors.

Table 3*Descriptive Statistics Results: GenAI-Writerism Elements*

	Complexity	Accuracy	Fluency	Appropriateness	Feedback	T&C	RU
Means	4.02	4.07	3.96	3.53	3.32	3.25	4.12
(SD)	(0.91)	(0.96)	(0.91)	(0.98)	(1.06)	(1.00)	(0.90)
Skewness	-0.12	-0.25	-0.01	0.31	0.28	0.24	-0.18
Kurtosis	-0.09	-0.33	0.69	-0.09	0.12	0.35	-0.18
p-values of							
Shapiro-Wilk	0.37	0.12	0.049	0.12	0.17	0.33	0.12

Note. T&C: Trust and Credibility, RU: Reflective Use, SD: Standard Deviation

Correlations Between L2 Writing Self-Efficacy and GenAI-Writerism

The first research question in this pilot study aimed to examine correlations between Japanese EFL university students' English writing self-efficacy and GenAI-Writerism. Correlation testing revealed modest correlations only between self-efficacy in B2 level writing activities and T&C, and between self-efficacy in C2 level writing activities and T&C, whereas no correlations were detected between the other L2 writing self-efficacy elements and the other GenAI-Writerism elements (Table 4). These findings can indicate that L2 writing self-efficacy does not necessarily influence GenAI-Writerism, or vice versa.

The results do not show a consistent pattern in which higher or lower L2 writing self-efficacy is associated with stronger GenAI-Writerism. As qualitative data, such as student comments, actual GenAI use, and chat logs, were not collected, the reasons for this weak relationship cannot be identified from the present data. However, the absence of consistent correlations suggests that GenAI-Writerism may not simply reflect learners' confidence or lack of confidence in English writing. Instead, it may also be shaped by other factors, such as perceived usefulness and perceived ease of use (Davis, 1989), trust in GenAI feedback, task demands, or individual differences. For example, Du et al. (2025) report that personality traits can affect perceptions and behaviours related to GenAI output. Therefore, these findings point to the need for future research that examines GenAI-Writerism together with technology acceptance, personality traits, writing type, and actual GenAI-use behaviour.

Table 4

Correlation Between L2 Writing Self-Efficacy and GenAI-Writerism

Variables		L2 Writing Self-Efficacy						
		A1	A2	B1	B2	C1	C2	
GenAI-Writerism	Complexity	0.13	0.11		0.10	0.04	0.02	0.04
	Accuracy	0.11	0.001		-0.03	-0.002	0.02	0.01
	Fluency	0.06	0.03	-7.108×10^{-4}		-0.007	-0.02	-0.03
	Appropriateness	0.06	0.02		0.03	0.008	0.04	0.08
	Feedback	0.02	-0.008		0.01	0.11	0.14	0.15
	T&C	0.10	0.11		0.13	0.17*	0.15	0.17*
	RU	0.10	0.09		0.04	0.07	0.02	0.05

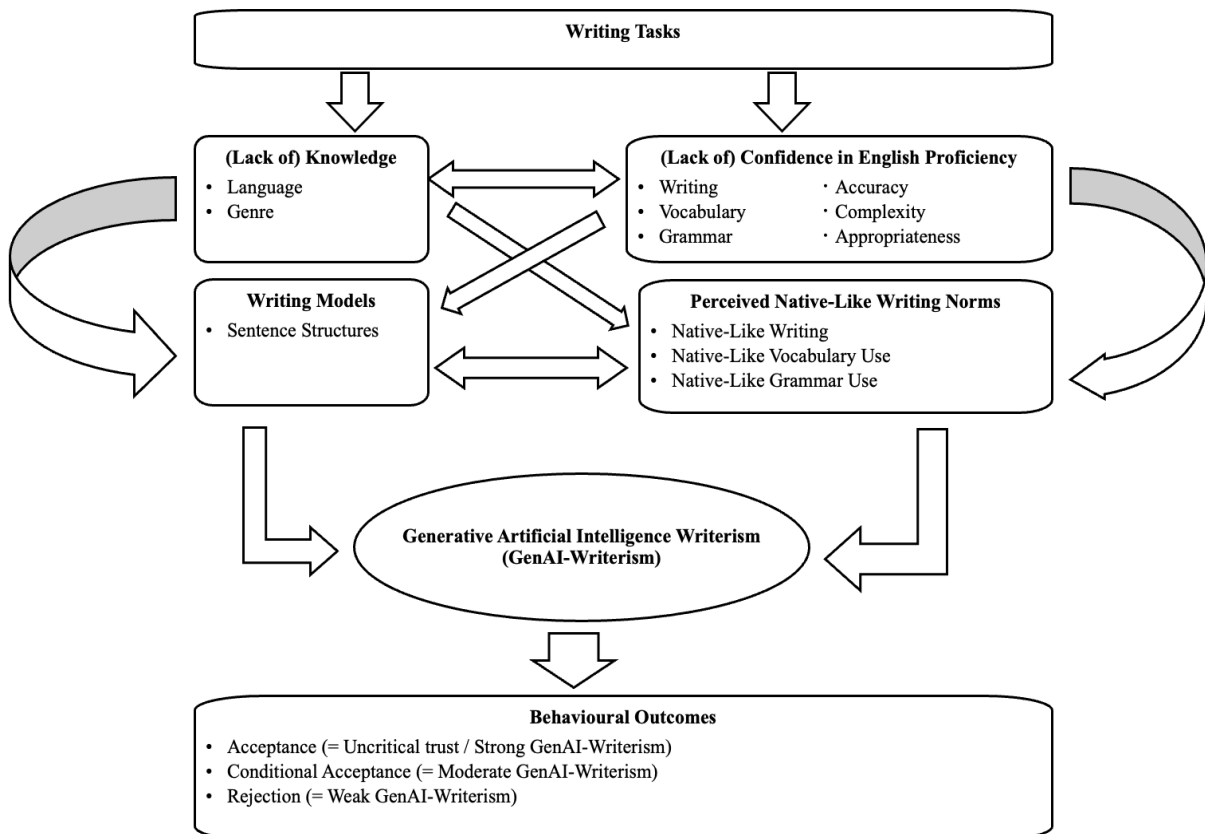
Note. T&C: Trust and Credibility, RU: Reflective Use

Note. *p < .05, **p < .01, ***p < .001

Note. $0 < |\tau| < 0.1$ (weak), $0.1 \leq |\tau| < 0.3$ (modest), $0.3 \leq |\tau| < 0.5$ (moderate), $0.5 \leq |\tau| < 0.8$ (strong), $0.8 \leq |\tau| \leq 1.0$ (very strong) (Cohen et al., 2018)

Figure 2

A Provisional Task-Sensitive Model of GenAI-Writerism and Its Possible Factors



The findings prompted a tentative revision of Shibata's (2026) conceptual model, as presented in Figure 2. The absence of consistent correlations between CEFR-J-based L2 writing self-efficacy and most GenAI-Writerism elements suggests that writing self-efficacy

should not be treated as a direct or general antecedent of GenAI-Writerism. However, the modest positive correlations between B2- and C2-level writing self-efficacy and trust and credibility indicate that the relationship may become more salient in connection with relatively advanced writing demands. Figure 2, therefore, introduces the writing task as a contextual factor that may shape learners' perceived knowledge needs, confidence in their English proficiency, and beliefs about the value and reliability of GenAI-supported writing.

One possible explanation for the modest correlations between B2- and C2-level writing self-efficacy and the trust and credibility (T&C) element is that B2- and C2-level descriptors are more closely related to advanced academic writing activities, such as essays, reports, and extended written arguments. In such tasks, students may perceive GenAI as a useful and accessible resource for managing linguistic complexity, reducing workload, and improving the quality of academic assignments. These results might also illuminate that writing tasks and genres influence learners' L2 writing self-efficacy (An & Li, 2024). Since university writing tasks can affect students' grades and may be completed under time pressure, some students may view GenAI output as a trustworthy way to support task completion (Abbas et al., 2024). However, because writing task type was not directly measured in this study, this interpretation should be treated as a hypothesis rather than as a confirmed finding. Future studies should therefore examine how task type, genre, deadline pressure, and assessment stakes influence GenAI-Writerism.

Correlations Between GenAI-Writerism Elements

The second research question examined correlations between GenAI-Writerism elements. Correlation testing revealed modest-to-moderate correlations among all proposed elements (Table 5). This suggests that students who value one aspect of GenAI-supported writing, such as complexity, accuracy, or fluency, may also tend to value other aspects, such as feedback, trust, and reflective use. However, because factor analysis was not conducted, these correlations should not be interpreted as evidence that the proposed elements form a validated scale. Rather, they provide preliminary support for treating GenAI-Writerism as a multidimensional belief orientation that requires further validation. These findings are consistent with previous research showing that students often value GenAI for surface-level linguistic improvements, including complexity, accuracy, and fluency (Feng et al., 2025; Kim et al., 2026), whilst also raising concerns about overreliance when GenAI output is accepted without careful evaluation (Hyland, 2025). Future studies should examine actual GenAI use through classroom observation, chat logs, revision histories, and interviews to understand how these beliefs shape writing behaviour.

Table 5*Correlation Between GenAI-Writerism Elements*

Variables	Complexity	Accuracy	Fluency	Appropriateness	Feedback	T&C	RU
Complexity	—						
Accuracy	0.53***	—					
Fluency	0.57***	0.43***	—				
Appropriateness	0.55***	0.49***	0.52***	—			
Feedback	0.48***	0.43***	0.46***	0.51***	—		
T&C	0.54***	0.44***	0.49***	0.48***	0.58***	—	
RU	0.51***	0.40***	0.40***	0.33***	0.36***	0.47***	—

Note. T&C: Trust and Credibility, RU: Reflective Use

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Note. $0 < |\tau| < 0.1$ (weak), $0.1 \leq |\tau| < 0.3$ (modest), $0.3 \leq |\tau| < 0.5$ (moderate), $0.5 \leq |\tau| < 0.8$ (strong), $0.8 \leq |\tau| \leq 1.0$ (very strong) (Cohen et al., 2018)

Conclusion

This exploratory pilot study examined correlations between L2 writing self-efficacy and GenAI-Writerism in a Japanese EFL tertiary educational context. The results showed little evidence of a consistent relationship between L2 writing self-efficacy and GenAI-Writerism across most patterns. Only modest correlations were found between B2/C2 writing self-efficacy and the T&C element of GenAI-Writerism. These findings suggest that GenAI-Writerism may not simply reflect learners' confidence or lack of confidence in English writing. Rather, it may be shaped by other factors, including perceived usefulness, trust, task demands, and students' broader experiences with GenAI-supported writing. The study also found modest-to-moderate correlations among the proposed GenAI-Writerism elements, which indicate that learners' beliefs about GenAI output, feedback, trust, and reflective use may be interconnected. However, because the questionnaire was newly developed and factor analysis was not conducted, the proposed dimensions should still be treated as tentative rather than validated. The findings should, therefore, be interpreted as hypothesis-generating evidence rather than as confirmation of a fully established construct.

Despite these limitations, this study contributes to current discussions on GenAI and L2 writing by positioning GenAI-Writerism as a writing-specific belief orientation that is related to, but distinct from, AI literacy, general trust in technology, and technology acceptance. Further research should validate the questionnaire, examine its relationship with AI literacy and technology acceptance, and include qualitative data such as interviews,

classroom observations, chat logs, and revision histories. Such research would help clarify how GenAI-Writerism shapes learners' actual writing and revision behaviours.

Limitations

The first limitation concerns the GenAI-Writerism questionnaire. Although item reliability was sufficiently high in this study, neither exploratory nor confirmatory factor analysis was conducted. Therefore, the seven proposed elements were based on theoretical discussion and researcher judgment rather than statistical validation. This means that unnecessary elements may have been included, whilst other important elements may have been excluded. Future studies should conduct factor analysis with a larger sample to refine the questionnaire and examine whether the proposed dimensions are empirically supported.

The second limitation is that this study relied only on questionnaire data. Students' actual GenAI use for writing activities was not observed, and data such as chat logs, revision histories, writing products, and interviews were not collected. Because of that, the study cannot fully explain how students' beliefs about GenAI translate into real writing behaviour. Since affective factors, including self-efficacy, can influence learning engagement and behaviours (Barcelos & Kalaja, 2011; Pajares, 2003; Shibata, 2019), future studies should examine students' actual GenAI-supported writing processes in addition to questionnaire responses. This could include collecting chat histories, comparing original and revised drafts, conducting interviews, videotaping writing activities, or observing classroom writing tasks to better understand the relationship between GenAI-Writerism, writing self-efficacy, and actual writing practices.

Key Takeaways

Unlike AI literacy, GenAI-Writerism should be understood not as usage behaviour, but as a belief system regarding the usefulness, reliability, and appropriateness of GenAI in writing (Shibata, 2025, 2026). Similar to motivation and other affective factors, GenAI-Writerism cannot be observed directly, but it can be explored through questionnaires, interviews, chat histories, and revision behaviour. More importantly, GenAI-Writerism is not necessarily harmful. When students use GenAI critically, it may support writing confidence, feedback engagement, and language development. However, when students accept GenAI output without evaluation, it may increase overreliance and reduce opportunities for independent writing development.

For teachers, the implication is not simply to prohibit or punish GenAI use. Instead, teachers may need to help students understand when GenAI feedback is useful, when it should be questioned, and how it can support learning without replacing their own ideas. In

writing classes, individual consultations can provide opportunities for students to show their GenAI chat logs, explain which suggestions they accepted or rejected, and discuss whether the feedback was appropriate for their writing purpose. Such practices may help students develop more reflective and responsible GenAI use while allowing teachers to better understand students' beliefs and decision-making processes.

Acknowledgments

We sincerely thank the reviewers and editors of the NUFS Teacher Development Symposium Proceedings for their constructive comments and helpful suggestions on an earlier version of this manuscript. We also thank the students who kindly participated in the questionnaire.

Notes on Contributors

Naoya SHIBATA is a specially appointed associate professor at Osaka Kyoiku University. His research interests include second language writing, theme-based instruction, materials development, curriculum design, and teacher training.

Vivian TURK is a lecturer at Nagoya City University. Her research interests include second language writing, learner identity, student autonomy, classroom-based research, and the pedagogical use of generative AI in English language education.

References

- Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, 21, Article Number 10. <https://doi.org/10.1186/s41239-024-00444-7>
- An, H., & Li, S. (2024). Task-specific writing anxiety and self-efficacy are separate from general L2 writing anxiety and self-efficacy and they have differential associations with the effects of written corrective feedback in pre-task and within-task planning. *System*, 126, Article Number 103480. <https://doi.org/10.1016/j.system.2024.103480>
- Barcelos, A. M. F., & Kalaja, P. (2011). Introduction to beliefs about SLA revisited. *System*, 39(3), 281–289. <https://doi.org/10.1016/j.system.2011.07.001>
- Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). AI literacy in education: Conceptualization, assessment, and implications. *Computers and Education Open*, 6, Article Number 100171. <https://doi.org/10.1016/j.caeo.2024.100171>
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge. <https://doi.org/10.4324/9781315456539>
- Common European Framework of Reference for Languages-Japan. (2025). <https://www.cefrij.org/index.html>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Darvin, R. (2025). Identity and investment in the age of generative AI. *Annual Review of Applied Linguistics*, 45, 10–27. <https://doi.org/10.1017/S0267190525100135>
- Dörnyei, Z., & Dewaele, JM. (2023). *Questionnaires in second language research: Construction, administration, and processing* (3rd ed.). Routledge. <https://doi.org/10.4324/9781003331926>
- Du, C., Tang, M., Wang, C., Zou, B., Xia, Y., & Du, Y. (2025). Who is most likely to accept AI chatbots? A sequential explanatory mixed-methods study of personality and ChatGPT acceptance for language learning. *Innovation in Language Learning and Teaching*, 1–22. <https://doi.org/10.1080/17501229.2025.2555515>
- Feng, H., Li, K., & Zhang, L. J. (2025). What does AI bring to second language writing? A systematic review of AI-integrated L2 writing research. *Language Learning & Technology*, 29(1), 1–27. <https://hdl.handle.net/10125/73629>

- Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But.... *Communication Methods and Measures*, 14(1), 1–24.
<https://doi.org/10.1080/19312458.2020.1718629>
- Hyland, K. (2025). EFL teachers and feedback fatigue: AI to the rescue? *Language Teaching*, 1–17. <https://doi.org/10.1017/S0261444825101018>
- Kim, J., Lee, S. S., Detrick, R., Wang, J., & Li, N. (2026). Students-generative AI interaction patterns and its impact on academic writing. *Journal of Computing in Higher Education*, 38, 504–525. <https://doi.org/10.1007/s12528-025-09444-6>
- Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies*, 30, 1265–1300. <https://doi.org/10.1007/s10639-024-12878-7>
- Koizumi, R., Agawa, T., Asano, K., & In'nami, Y. (2022). Skill profiles of Japanese English learners and reasons for uneven patterns. *Language Testing in Asia*, 12, Article Number 53. <https://doi.org/10.1186/s40468-022-00203-3>
- Lee, S., Jeon, J., McKinley, J., & Rose, H. (2025). Generative AI and English language teaching: A global Englishes perspective. *Annual Review of Applied Linguistics*, 45, 85–108. <https://doi.org/10.1017/S0267190525100184>
- Lin, A. M. Y., & Chen, Q. (2025). Towards ethical and responsible generative AI use in English medium education classrooms across Asia: A posthumanist perspective. In F. V. Lim & J. Pun (Eds.), *Designing learning with multimodality in English medium of education (EME) classrooms across Asia* (pp. 235–258). Bloomsbury.
<https://doi.org/10.5040/9781350358294.ch-010>
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3), 412–433. <https://doi.org/10.1037/met0000144>
- Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' perceptions of generative AI and academic integrity in EFL writing. *Education Sciences*, 15(5), Article Number 611. <https://doi.org/10.3390/educsci15050611>
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*, 19(2), 139–158.
<https://doi.org/10.1080/10573560308222>
- Pallant, J. L., Blijlevens, J., Campbell, A., & Jopp, R. (2026). Mastering knowledge: the impact of generative AI on student learning outcomes. *Studies in Higher Education*, 51(4), 714–735. <https://doi.org/10.1080/03075079.2025.2487570>

- Plonsky, L., & Derrick, D. J. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal*, 100(2), 538–553.
<https://doi.org/10.1111/modl.12335>
- Sterling, L., Ye, C., Ying, H., & Chen, Z. (2025). Finding your voice: Generative AI, international students, and academic writing. *Information*, 16, Article Number 289.
<https://doi.org/10.3390/info16040289>
- Shibata, N. (2019). The impact of students' beliefs about English language learning on out-of-class learning. *Relay Journal*, 2(1), 122–136. <https://doi.org/10.37237/relay/020117>
- Shibata, N. (2025, August 28). *The usefulness of ChatGPT as a feedback tool for argumentative essay writing* [conference presentation]. The 64th JACET International Convention. Tokyo, JAPAN.
- Shibata, N. (2026). Generative artificial intelligence writerism: Factors in overestimation of generative artificial intelligence output. *Hyumanistic eigokyoiku kenkyukai kiyo dai 6 go* [*Journal of Humanistic English Education in Japan*, 6], 6–21.
- Tono, Y. (Ed.). (2013). *The CEFR-J handbook*. Taishukan Shoten.
- Tono, Y. (2017). The CEFR-J and its impact on English language teaching in Japan. *JACET International Convention Selected Papers*, 4, 31–52.
- Tono, Y. (2019). Coming full circle: From CEFR to CEFR-J and back. *CEFR Journal - Research and Practice*, 1, 5–17. <https://doi.org/10.37546/jaltsig.cefr1-1>
- Wolf, M. J., Grodzinsky, F., & Miller, K. W. (2024). Generative AI and its implications for definitions of trust. *Information*, 15, Article Number 542.
<https://doi.org/10.3390/info15090542>

Appendix

第二言語ライティング活動における生成 AI の信用性に関して

(The Credibility of Generative AI in Second Language Writing Activities)

Instruction: それぞれの文章に対して以下の基準でご回答ください。(1:全くそう思わない / 2:そう思わない / 3:どちらかといえばどう思わない / 4:どちらかといえばそう思う / 5:そう思う / 6:強くそう思う) (Please answer each statement according to the following criteria. (1: Completely disagree / 2: Disagree / 3: Somewhat disagree / 4: Somewhat agree / 5: Agree / 6: Strongly agree))

Q1 (Complexity): 生成 AI は複雑な文章もうまく作成できると信じている。(I believe Generative AI makes complex sentences well.)

Q2 (Accuracy): 生成 AI 機能 (例 : ChatGPT や Grammarly) によって生成されたテキストは、常に言語的に正しいと考える。(I believe text generated by AI functions (e.g., ChatGPT and Grammarly) is always linguistically correct.)

Q3 (Fluency): 生成 AI は私の考えをよりスムーズに表現するのに役立つ。(Generative AI helps me express my thoughts more smoothly.)

Q4 (Appropriateness): 生成 AI は、私たちが指示しない限り、どのジャンルの文章でも不適切な文章を作成することはないと思う。(I do not think Generative AI makes inappropriate sentences in any writing genre unless we tell them to make them.)

Q5 (Feedback): 生成 AI は、あらゆるジャンルの文章における言語の使用を修正するための適切なフィードバックを提供してくれると信じている。(I believe Generative AI can provide me with appropriate feedback to revise my use of language in any writing genre.)

Q6 (Trust and Credibility): 生成 AI からのフィードバックは教師からのフィードバックよりも信頼できると思う。(I think Generative AI feedback is more reliable than teacher feedback.)

Q7 (Reflective Use): 生成 AI を使って自分の文章を修正すると、何かを学べる気がする。(I feel I learn something when I revise my writing using Generative AI.)

Q8 (Complexity): 生成 AI は文法的に複雑な文章を効果的に作ることができると思う。(I think Generative AI can make grammatically complex sentences effectively.)

Q9 (Fluency): 生成 AI は、連続し流れるようなテキストを生成するのが得意だと思う。(I think Generative AI is good at generating continuous, flowing text.)

Q10 (Appropriateness): 生成 AI は、さまざまな執筆目的に必要なトーンを一致させることができると信じている。(I believe Generative AI can match the tone needed for different writing purposes.)

Q11 (Trust and Credibility): 私は同級生からの提案よりも生成 AI による提案を信頼する。(I trust the suggestions made by Generative AI more than those from my peers.)

Q12 (Reflective Use): 生成 AI は私のライティング スキルの向上に役立つ。(Generative AI helps me improve my writing skills.)

Q13 (Complexity): 生成 AI は文章作成において幅広い文構造を活用できると考えている。(I believe Generative AI can use a wide range of sentence structure in writing.)

Q14 (Accuracy): 生成 AI は、文章中の文法上の間違いを避けるのに役立つ。(Generative AI helps me avoid grammatical mistakes in my writing.)

Q15 (Fluency): 生成 AIのおかげで、私は行き詰まることなく簡単に文章を書くことができる。(Generative AI makes it easier for me to write without getting stuck.)

Q16 (Appropriateness): 生成 AI は、あらゆる文章ジャンルに適した語彙を使用できると思う。(I think Generative AI can use vocabulary that is suitable for any writing genre.)

Q17 (Feedback): 生成 AI から提供された提案を直ちに受け入れる。(I accept suggestions provided by Generative AI immediately.)

Q18 (Trust and Credibility): 生成 AI は信頼できるライティングアシスタントツールであると感じる。(I feel that Generative AI is a trustworthy writing assistant tool.)

Q19 (Reflective Use): 生成 AI から新しい単語や表現を学べると信じている。(I believe I can learn new words and expressions from Generative AI.)

Q20 (Complexity): 生成 AI は、私が普段書く文章よりも長くて複雑な文章を使用するので便利だと思う。(I find Generative AI useful because it uses longer and more complex sentences than I usually write.)

Q21 (Accuracy): 生成 AI は私よりも文法の間違いが少ないと思う。(I think Generative AI makes fewer grammar mistakes than I do.)

Q22 (Fluency): 私は文章を書く際、途中で手が止まったり途切れたりしないように、生成 AI を活用する。(I use Generative AI to help me avoid pauses or breaks in my writing.)

Q23 (Appropriateness): 生成 AI が書く課題に合わせて言語表現を調整してくれると信頼しています。(I trust Generative AI to adapt its language to the writing task.)

Q24 (Trust and Credibility): 自分の文章に自信がないときは、生成 AI をより信頼する。(I trust Generative AI more when I feel unsure about my writing.)

Q25 (Reflective Use): 私は、自分のアイデアを置き換えるためではなく、学習をサポートするツールとして 生成 AI を使用する。(I use Generative AI as a tool to support my learning, not to replace my ideas.)

Q26 (Accuracy): 私はよく生成 AI を利用して、小さな文法上の問題をチェックする。(I often rely on Generative AI to check for small grammar issues.)

Q27 (Fluency): 生成 AI の文章は時々あまりにスラスラ進みすぎて、自分の考えをうまくコントロールできなくなることがあります。(Sometimes Generative AI writing flows too much, and I lose control of my ideas.)

Q28 (Feedback): 確信が持てないときでも、生成 AI のフィードバックに従う。(I follow Generative AI feedback even when I am unsure about it.)

ご協力ありがとうございました (Thank you very much for your cooperation.)

Teaching the Room: A Qualitative Inquiry into Seating and Grouping in the Language Classroom

Vivian TURK

Nagoya City University

Abstract

Seating and grouping arrangements are often treated as practical decisions, yet they shape interaction in important ways. This qualitative study examines how these arrangements affect students' participation, belonging, and perceived learning in first-year Core English classes in Japan. Drawing on classroom observations, student reflections, focus groups, and teacher journal, the study explores rows, clusters, circles, and rotational formats through narrative vignettes. Findings show that seating arrangements influence participation and perceived learning through the feelings they create, as in the model arrangement → emotion → participation → perceived learning. The study highlights the emotional and relational dimensions of classroom space and offers practical ideas for responsive classroom design.

Keywords: seating arrangements, classroom interaction, student engagement, qualitative research, EFL classroom

Teaching the Room: A Qualitative Inquiry into Seating and Grouping in the Language Classroom

Classroom seating and grouping are often treated as simple teaching logistics: rows of desks, clusters of tables, circles for discussion. Yet these everyday choices can shape how students interact, how they relate to classmates and teachers, and how they see their chances to participate and learn (Harvey & Kenyon, 2013). Research in educational psychology and classroom management has long studied seating in terms of measurable outcomes such as test scores, attention, or behaviour (Xi et al., 2017; Syaifullah et al., 2022). While this work shows that seating matters, it often leaves less room for the more human side of classroom life: feelings of inclusion or exclusion, changing peer relationships, and the ways identity and agency are shaped by physical space.

This study takes a different path. Instead of asking whether one arrangement is better than another, it explores how seating and grouping are experienced in practice, and how those experiences shape classroom interaction and learning (Nehyba et al., 2023; Norazman et al., 2019). Using a qualitative design that combines classroom ethnography, narrative inquiry, and teacher action research, the study treats seating not as a variable to measure but as a lived and relational part of classroom life. Drawing on classroom observations, student reflections, focus group discussions, and teacher reflective journal, the study examines how spatial arrangements influence students' participation, sense of belonging, and views of their own learning in the language classroom. This issue is especially important in EFL classrooms, where participation often depends on students' willingness to speak, take risks, ask for help, and negotiate meaning with peers.

This study builds on previous research by focusing more directly on lived experience. While much seating research has examined observable outcomes, this study centres student voice, thick description, and narrative representation. It treats the classroom not only as a place of instruction but also as a social and cultural space where interaction, emotion, and identity are constantly being shaped. By bringing together student perspectives and teacher reflection, the study aims to offer insights that are both theoretically useful and practical for teachers.

Literature Review

Seating and grouping are often treated as routine classroom decisions, but a growing body of research shows that they play an important role in shaping interaction, engagement, and learning. Classroom space is not just a physical background. It actively shapes students' behaviour, communication, and chances to participate (Norazman et al., 2019). In other

words, where students sit in relation to each other and to the teacher affects how classroom life develops.

Many studies have examined rows, clusters, and circular or U-shaped seating. Traditional rows are often linked to teacher-centred instruction and less peer interaction, with students positioned more as listeners than active contributors (Haghighi & Jusan, 2012). By contrast, clusters and circles are often linked to collaborative learning, more communication, and greater participation (Norazman et al., 2019). Research also suggests that seat position matters. Students who sit closer to the teacher or in more central zones often show higher engagement and performance (Xi et al., 2017). Flexible and mobile seating has also been linked to higher levels of perceived engagement and interaction in student-centred settings (Harvey & Kenyon, 2013).

At the same time, findings are not always consistent. Some studies suggest that collaborative seating improves participation and attention (Syaifullah et al., 2022), but others challenge the idea that more interactive layouts always lead to better results. Ali and Jabar (2025), for example, found that traditional rows led to higher speaking participation in an EFL setting. Nehyba et al. (2023) likewise found that students in rows showed higher speech rates than those in circular arrangements during reflective discussions. These findings suggest that the effects of seating depend not only on layout but also on task type, classroom culture, and how the class is facilitated. Because of this, recent work increasingly argues that there is no single best arrangement. What works depends on the goals of the lesson and the teaching context (Norazman et al., 2019).

This study also draws on socio-cultural and identity-oriented views of language learning. From a socio-cultural perspective, participation is not only an individual behaviour but a process shaped by social relationships, classroom practices, and access to meaningful interaction (Lave & Wenger, 1991; Wenger, 1998). Similarly, identity-focused work in language learning shows that learners' participation is connected to how they see themselves and how they are positioned by others in classroom spaces (Norton, 2013; Toohey, 2000). These perspectives are useful for this study because seating and grouping affect not only where students sit, but also how visible, confident, included, or peripheral they feel.

Pedagogically, the study is situated within communicative, task-oriented language teaching, where interaction and meaningful language use are central learning goals (Richards, 2006). In Core English classes, seating choices are closely connected to lesson goals such as pair practice, peer explanation, small-group discussion, and short speaking tasks. For this

reason, classroom layout is not treated as separate from pedagogy. It is understood as one of the conditions that can support or limit opportunities for communication.

A major weakness in the existing research is methodological. Much of it relies on quantitative measures or structured observation and focuses on outcomes such as participation frequency, attention, or grades. These approaches show that seating matters, but they often reduce classroom life to visible behaviours. They tell us little about how students experience different arrangements. There is still limited research on the emotional and relational side of classroom space: how seating shapes comfort, anxiety, inclusion, exclusion, and willingness to participate. Even when context is mentioned, student voice and cultural setting are still not studied closely enough (Haghighi & Jusan, 2012).

This methodological gap is important because qualitative classroom research can show how students make meaning from everyday classroom practices. Narrative inquiry, in particular, is useful for language education because it keeps attention on learners' lived experience, identity, and classroom stories (Barkhuizen et al., 2024). Classroom ecology perspectives also help frame the classroom as a dynamic environment where physical space, interaction, and learner agency are closely connected (van Lier, 2004).

What is still missing, then, is a view of seating and grouping as lived experience rather than as something to be measured from the outside. Existing studies can tell us whether students speak, but not how they feel about speaking, how they position themselves within different arrangements, or how classroom space shapes their sense of belonging and agency.

This study addresses that gap by using a fully qualitative, narrative approach that centres student and teacher perspectives. It examines seating and grouping not as fixed structures but as relational and emotional realities of classroom life. Through thick description and narrative vignettes, it aims to show more clearly how spatial arrangements shape participation, belonging, and perceived learning in the language classroom.

Methodology

Research Design

This study uses a qualitative classroom inquiry that brings together three related traditions: ethnography, narrative inquiry, and action research. The ethnographic side focuses on everyday classroom life through field notes, observations, and close accounts of interaction as it happens. The narrative side focuses on lived experience, using student reflections, classroom interaction, and group discussions to understand how students and the teacher make sense of seating and grouping. The action research side appears in repeated

cycles of trying different seating arrangements, observing what happens, and adjusting practice in response.

This combined design is intentional. It moves the study away from questions of simple effectiveness and toward a deeper understanding of how students experience, interpret, and negotiate classroom space. The study is guided by two research questions:

1. How do different seating and grouping arrangements shape students' experiences of participation, belonging, and perceived learning in the language classroom?
2. What do students' and teachers' narratives reveal about inclusion, exclusion, and agency within these arrangements, and what do these accounts suggest for more responsive classroom design?

Participants and Context

The study was conducted in first year Core English classes at a private foreign language university in central Japan. The study involved four Core English classes with approximately 110 first-year undergraduate students in total. Each class had about 20 to 30 students and aimed to develop students' communicative competence in English. Data were collected across 2024 and 2025. Students had mixed levels of English proficiency, which were understood through course placement, classroom performance, and regular Core English tasks rather than through a separate proficiency test.

As the teacher-researcher, I held a dual role. I taught the classes while also documenting and reflecting on classroom processes. This gave me close and sustained access to the setting, but it also required careful reflexivity when interpreting classroom events.

Students took part voluntarily in classroom-based data collection, including reflections and focus groups. Ethical care was taken throughout. Students were informed about the purpose of the study, participation was not connected to assessment, and all data were anonymised. Students could also withdraw their contributions at any time without penalty. Because the study took place in a familiar classroom setting over time, it was possible to build trust and collect rich, situated data.

Focus group participants were invited from students who had taken part in the classroom activities and had experienced several seating arrangements during the semester. Participation was voluntary, and students were not selected based on grades or performance. A total of four focus groups were conducted, with approximately five to six students in each group.

Implementation of Seating Arrangements

The seating arrangements were introduced as part of normal lesson design rather than as separate research activities. Across the data collection period, students experienced four main arrangements: rows, small clusters of three to four students, a whole-class circle, and rotational pair work. These arrangements were selected according to the lesson goal. Rows were mainly used for focused input, individual work, and written preparation. Small clusters were used for peer explanation, discussion, and collaborative tasks. The whole-class circle was used for discussion activities where students were expected to listen and respond to one another. Rotational pair work was used for short speaking practice and repeated dialogue tasks.

When the seating arrangement changed, students were briefly told the purpose of the change in simple classroom language, such as “Today we will use groups so you can help each other” or “Today we will rotate partners to practise the same question several times.” The speed-dating format involved pairs speaking for approximately six minutes before one row of students moved to a new partner. This continued for up to ten rounds, depending on the task and available time.

Each arrangement was used multiple times during each semester, so that the analysis was not based on a single lesson. The exact timing and frequency of each arrangement depended on the course schedule, classroom conditions, and lesson aims.

Data Collection

To capture seating and grouping as lived experience, the study used several qualitative methods built into ordinary classroom practice:

Classroom observations and field notes. Each lesson was observed closely, with attention to student positioning, interaction, and the way the physical layout shaped participation. After each class, detailed field notes were written. These separated descriptive notes, which recorded what was seen and heard, from reflective notes, which recorded interpretations, questions, and emotional responses.

Teacher reflective journal. Alongside the field notes, I kept a teacher’s reflective journal to record teaching intentions, in-the-moment decisions, and ongoing dilemmas. This journal helped capture how seating arrangements were experienced from the teacher’s point of view and how instructional choices changed over time.

Student reflective writing. After selected lessons, students were invited to write short reflections of two or three sentences, or to leave a short voice note, on how the seating

arrangement affected their participation, comfort, or learning. These short reflections created repeated snapshots of student experience across the semester.

Focus groups. Near the end of the term, students took part in small focus groups of five to six students. These discussions explored how they experienced different seating arrangements. Focus groups were useful because they showed both individual views and shared patterns, tensions, and differences in how students understood classroom interaction.

Visual and material artefacts. Seating maps and classroom photographs were collected for each lesson to document the physical setup. Samples of student work produced during group activities were also kept as artefacts, helping to show how collaboration developed within different arrangements.

Data Analysis

The analysis took place in two connected stages:

Thematic analysis. All textual data, including field notes, teacher's reflective journal, student reflections, and focus group transcripts, were coded inductively. Initial codes were created through close reading and then grouped into broader themes linked to participation, belonging, and perceived learning. These themes were then compared across different seating arrangements to show how specific layouts supported or limited different classroom experiences (Xi et al., 2017; Nehyba et al., 2023).

Narrative vignettes. Alongside thematic coding, rich narrative vignettes were written to present thick descriptions of key classroom moments. The vignettes were selected as contrasting cases that best illustrated the main themes from the analysis: safety and silence in rows, peer support and uneven participation in clusters, confidence and exposure in circles, and energy and fatigue in rotational pair work. Each vignette drew together observational detail, student reflections, focus group comments, and teacher's reflective journal entries. The vignettes were not included simply as examples; they worked as analytic representations of how seating arrangements were experienced in practice.

Quality Checks. Several steps were taken to strengthen the quality and transparency of the study. First, data were triangulated across observations, student reflections, teacher's reflective journal, and artefacts. Second, emerging interpretations were shared with students during focus groups so that meanings could be confirmed or refined. Third, thick description was used to show clearly how interpretations were grounded in classroom events. Finally, reflexivity was maintained throughout, with the teacher-researcher acknowledging how her dual role shaped both the collection and interpretation of data.

Findings

The findings are presented through a series of narrative vignettes that show how students experienced seating and grouping in the language classroom. Each vignette highlights a different perspective and shows how space, interaction, and identity were connected. Together, these accounts show that seating is not a neutral background to learning but an active part of classroom experience.

The vignettes are composite accounts based on recurring patterns across field notes, student reflections, focus group comments, and teacher's reflective journal entries. Names are pseudonyms. Short student quotations in the main text have been lightly edited for readability where necessary, while preserving the original meaning. Appendix C includes sample reflections closer to students' original language use to show the range of expression in the data.

Vignette 1: The Safe Corner (Rows Arrangement)

In a lesson organised in traditional rows, the classroom felt orderly and quiet. From the front of the room, attention looked even: eyes forward, notebooks open. Yet a closer look showed different levels of engagement. In the back corner, Yuta leaned slightly behind the student in front of him, partly hidden from my line of sight. He rarely raised his hand, but he took careful notes and sometimes whispered questions to his neighbour.

In his reflection, Yuta wrote: "When I sit in rows, I don't feel pressure. I can just focus. If I don't understand something, I wait until later and look it up." For him, the anonymity of rows created safety, even though it also limited participation. The teacher's reflective journal from that lesson recorded something similar: "The rows gave me control of the class, but I struggled to know what students at the back were really thinking."

Rows, therefore, created a kind of protective space. They supported quiet focus, but they also concealed silence and made some learning hard to see.

Vignette 2: The Cluster That Cares (Small Groups of Four)

On another day, desks were arranged in clusters of four. The room quickly filled with low conversation as students leaned towards each other over a reading task. In one group, Aya, who was often silent during whole-class discussion, spoke animatedly while pointing to the page. Her partner nodded, added examples, and the others built on the exchange.

Aya later wrote: "I asked my classmate to explain one word, and then I could also explain another. I felt less shy because only three people listened, not everyone." In the focus group, one student expressed a similar view: "In small groups, you don't worry about being

wrong. Friends help you.” The teacher’s reflective journal also reflected this atmosphere: “Clusters made the room feel alive. The quietest students suddenly had a voice.”

At the same time, not everyone was included equally. In one cluster, a student stayed quiet and doodled in her notebook while the others spoke. The arrangement supported peer help, but it did not automatically guarantee equal participation.

Clusters, then, created conditions for peer support, but they also required teacher attention to ensure that quieter students were not left at the edge of interaction.

Vignette 3: The Circle of Exposure (Whole-Class Circle)

During a discussion lesson, the chairs were pulled into a large circle. The aim was to create equality: everyone facing each other and every voice given value. For some students, this worked well. Haruka, who rarely spoke in groups, raised her hand with confidence and contributed an idea that drew nods around the room. In her reflection, she wrote: “I could see everyone. It felt serious, like a seminar. I wanted to say something important.”

For others, however, the circle increased anxiety. Jun sat stiffly, kept his eyes down, and spoke only when directly invited. In a focus group, he said: “I felt like everyone was looking at me at the same time. Even before I talked, I was nervous.” The teacher’s reflective journal from that lesson noted: “The circle gave me an unusual sense of openness with the class, but I also noticed several students avoiding eye contact and retreating into themselves.”

The circle, therefore, made visibility central to participation. For some students, this visibility supported voice; for others, it increased pressure.

Vignette 4: The Moving Target (Speed-Dating Pairs)

In a later lesson, I used a speed-dating format where students changed partners every few minutes to practise short dialogues. The room changed immediately: chairs scraped, laughter spread, and voices overlapped as students moved and spoke. Rina, who was usually hesitant in group work, seemed to enjoy this format. She said: “I liked talking to many people. If I made a mistake, it didn’t matter because we changed partners quickly.”

Other students reacted differently. Takashi, who preferred stable groups, found the movement tiring. He wrote: “Changing so many times was tiring. I just wanted to stay with one partner and go deeper.” The teacher’s reflective journal noted: “The classroom felt electric, but the noise level was overwhelming. I worried about those who need more processing time.”

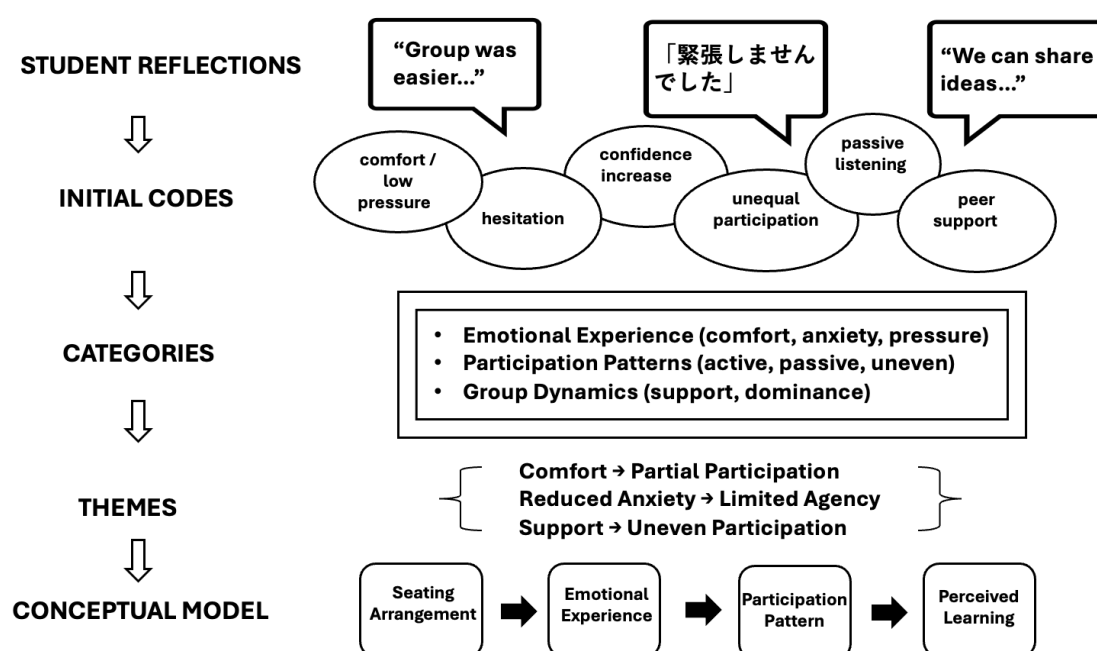
This arrangement brought energy and novelty, but it also made deeper interaction harder for some students. The activity showed that movement and repetition can support fluency practice but may also limit depth when students need more time to process ideas.

Across these vignettes, seating and grouping did not produce the same effects for everyone. Rows offered safety but also reduced visibility. Clusters supported closeness and peer help, yet some students still stayed on the edge. The circle created both empowerment and exposure. Rotational work increased energy but could reduce depth. What ties these accounts together is that space shaped feeling, and feeling then shaped participation (Ali & Jabar, 2025).

The findings support the view that no arrangement works best for all students in all situations (Norazman et al., 2019; Nehyba et al., 2023). Instead, the classroom appears as a shifting set of possibilities, where the same arrangement is experienced in very different ways by different learners. The model that came from this analysis - arrangement → emotion → participation → perceived learning - captures this relationship clearly. It shows that physical organisation shapes learning not directly, but through the feelings and interaction patterns it creates. Figure 1 illustrates the analytic movement from student reflections to initial codes, categories, themes, and the final conceptual model.

Figure 1

From student reflections to the conceptual model



This study shows that seating and grouping are not neutral classroom decisions (Norazman et al., 2019). They shape how students feel, how they take part, and how they understand their own learning (Syaifullah et al., 2022). The model developed here -

arrangement → emotion → participation → perceived learning - offers a simple way to understand this process.

Theoretical Contributions

The study adds to theory in several connected ways. First, it shows that seating and grouping should be understood not only as structural arrangements but also as emotional spaces. Across the vignettes, the same arrangement could create very different feelings. Clusters felt supportive and low-pressure for some students, while the circle created both openness and anxiety. Seating therefore carries emotional possibilities that can support or limit participation (Ali & Jabar, 2025). This extends socio-cultural theories of learning (Lave & Wenger, 1991; Wenger, 1998) by showing that participation in a community of practice depends not only on structure and interaction, but also on how learners feel within that space.

Second, the study shows how identity is shaped within classroom space. Students did not simply react to seating arrangements. They used them to position themselves in particular ways. Yuta understood rows as a place where he could be quiet and independent. Haruka saw the circle as a chance to become more visible and more academically confident. These examples suggest that classroom space is a place where identity is performed and negotiated, with students positioning themselves as visible or hidden, confident or hesitant, central or peripheral (Xi et al., 2017). This supports identity-based work in language learning (Norton, 2013; Toohey, 2000), while adding a stronger spatial dimension.

Third, the study highlights the importance of teacher reflexivity. Seating was not only experienced by students; it was also interpreted and managed by the teacher. The teacher's reflective journal showed how different arrangements created different feelings of control, uncertainty, energy, and concern. This makes it clear that classroom organisation is not a neutral technique. It is part of the ongoing emotional and interpretive work of teaching. In that sense, the study also adds to action research and reflective practice by showing how classroom space becomes part of pedagogical decision-making.

Finally, the study offers a process model of participation. Earlier studies often treated seating as a fixed factor linked to behaviour or performance (Haghighi & Jusan, 2012). By contrast, this study shows that the relationship between space and learning is mediated by emotional experience. Excitement in rotational work led some students to speak more confidently, while fatigue in the same arrangement led others to withdraw. The model therefore connects classroom management research with affective and socio-cultural views of learning, showing more clearly how space, emotion, interaction, and learning work together.

Practical Implications

The study also offers practical lessons for teachers. The vignettes make clear that seating and grouping should not be reduced to a search for the single best arrangement. Each setup offers benefits and limitations, and each works differently depending on the students and the task.

Clusters can help quieter students speak by creating small communities where peer explanation replaces some teacher intervention (Norazman et al., 2019). They can support shared learning and make participation feel safer. But clusters do not automatically produce equal participation. Some students may still stay quiet while others dominate. Teachers therefore need to watch group dynamics closely, circulate actively, and change group composition when needed.

Rows can provide psychological safety for students who prefer less social pressure (Ali & Jabar, 2025; Nehyba et al., 2023). They may work well for tasks that need concentration and individual processing. At the same time, they can hide disengagement and make it harder for teachers to notice students who are struggling in silence. Rows therefore work best when combined with low-pressure ways for students to show understanding, such as pair work, written responses, or short check-ins.

Circles can support equality and serious discussion. For some students, they make the classroom feel more open and more meaningful. But they can also increase anxiety because every student is fully visible to everyone else (Syaifullah et al., 2022; Nehyba et al., 2023). Teachers who use circles may need to reduce pressure through techniques such as think-pair-share, structured turn-taking, or preparation time before students speak.

Rotational formats such as speed-dating can increase energy and lower the emotional cost of making mistakes. For some students, the short exchanges and constant movement make speaking easier (Harvey & Kenyon, 2013). But the same format can feel mentally and emotionally tiring for others, especially when a task requires more time and depth. These activities are therefore useful for short, low-stakes communication practice, but not always for sustained discussion.

Across all arrangements, one point stands out: students experience the same layout in different ways (Xi et al., 2017; Haghighi & Jusan, 2012). No arrangement guarantees full participation or comfort. This means that teachers need flexibility. Seating should vary depending on the lesson goal, and student feedback should be taken seriously. Even short reflections can help teachers understand how students are experiencing classroom space and make better choices over time.

Two simple classroom techniques emerged as especially useful from this study. First, teachers can use a seating check-in at the end of selected lessons: one short prompt asking students how today's seating affected their speaking, comfort, or learning. This can be done in one or two minutes on paper or through a form. Second, when using clusters, teachers can assign rotating roles such as facilitator, supporter, reporter, and language checker. These roles give quieter students a clearer entry point into the task and reduce the risk that one confident student controls the group.

The approach can also be adapted to different classroom conditions. In larger classes, teachers may not be able to observe every group closely, so short written check-ins or quick reports can help identify how students are experiencing the arrangement. In lower-proficiency classes, seating changes should be paired with simple instructions, model phrases, and clear roles so that students are not overwhelmed by both language demands and movement. In fixed-furniture rooms, teachers can still adapt interaction patterns by changing whom students speak to, using front-back pair work, turning to nearby partners, or creating discussion zones within the existing layout. In non-Japanese contexts, the same reflective process can be used, but teachers should attend carefully to local norms around silence, eye contact, peer support, and public speaking.

In simple terms, effective classroom design depends on both alignment and responsiveness. Seating should fit the purpose of the lesson: clusters for collaboration, rows for focused input, circles for dialogue, and rotations for short speaking practice. At the same time, teachers need to respond to how students feel and behave in those spaces. When seating is treated as part of the lived classroom experience, it becomes a practical way to support participation, engagement, and belonging.

Limitations

This study has several limitations. First, it was conducted in a single institutional context, in first-year Core English classes at one private foreign language university in central Japan, so the findings should not be treated as generalisable to all EFL classrooms. Second, the participants who contributed reflections and focus group comments were self-selected, which may mean that students with stronger views or greater confidence were more likely to share their experiences. Third, my dual role as teacher and researcher shaped both the classroom practice and the interpretation of data, even though the teacher's reflective journal and student feedback were used to reduce this risk. Finally, the study focuses on participation, belonging, and perceived learning rather than demonstrated learning gains. For this reason,

the findings should be understood as evidence of students' classroom experiences, not as proof that one seating arrangement improves learning outcomes.

Conclusion

This study has examined how seating and grouping in the language classroom shape students' experiences of participation, belonging, and learning. Drawing on classroom observations, student reflections, focus groups, and the teacher's reflective journal, it has shown that everyday spatial decisions carry emotional and identity-related consequences.

The findings suggest that seating arrangements are not neutral parts of classroom organisation. Rows created safety but could hide disengagement. Clusters supported peer interaction but did not guarantee equal participation. Circles encouraged dialogue while increasing visibility and anxiety for some students. Rotational formats brought energy but sometimes reduced depth. The model proposed here - arrangement → emotion → participation → perceived learning - shows how classroom space may shape students' perceived learning through the feelings and interaction patterns it creates.

The study contributes to theory by treating seating as an emotional and relational part of pedagogy rather than just a structural variable. It also shows how students negotiate identity through their place in the room, and how teachers make sense of classroom space through ongoing reflection. For practice, the study suggests that seating should be matched to lesson goals while also remaining responsive to the different emotional and participation needs of students.

Rather than searching for a single ideal arrangement (Norazman et al., 2019; Nehyba et al., 2023), teachers may be better served by a flexible and reflective approach. When seating is understood as part of classroom ecology, it can be used more intentionally to support participation, engagement, and belonging.

References

- Ali, A. F., & Jabar, Z. A. (2025). The influence of seating arrangement on EFL students' speaking performance in the classroom. *Journal for Research Scholars and Professionals of English Language Teaching*, 9(50).
<https://doi.org/10.54850/jrspelt.9.50.003>
- Barkhuizen, G., Benson, P., & Chik, A. (2024). Narrative inquiry in language teaching and learning research (2nd ed.). Routledge. <https://doi.org/10.4324/9781003356394>
- Haghighi, M. M., & Jusan, M. M. (2012). Exploring students' behavior on seating arrangements in learning environment: A review. *Procedia – Social and Behavioral Sciences*, 36, 287–294. <https://doi.org/10.1016/j.sbspro.2012.03.032>
- Harvey, E. J., & Kenyon, M. C. (2013). Classroom seating considerations for 21st century students and faculty. *Journal of Learning Spaces*, 2(1).
- Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511815355>
- Nehyba, J., Juhaňák, L., & Cigán, J. (2023). Effects of seating arrangement on students' interaction in group reflective practice. *The Journal of Experimental Education*, 91(2), 249–277. <https://doi.org/10.1080/00220973.2021.1954865>
- Norazman, N., Ismail, A. H., Ja'afar, N. H., Khoiry, M. A., & Che Ani, A. I. (2019). A review of seating arrangements towards the 21st century classroom approach in schools. *Malaysian Journal of Sustainable Environment*, 7(2), 21–46.
- Norton, B. (2013). *Identity and Language Learning: Extending the Conversation* (2nd ed.). Bristol: Multilingual Matters.
- Richards, J. C. (2006). Communicative language teaching today. Cambridge University Press.
- Syaifullah, A., Munir, M., & Ariyani, A. (2022). An analysis of seating arrangement on students' learning attention. *Journal of Excellence in English Language Education*, 1(1).
- Toohey, K. (2000). Learning English at school: Identity, social relations and classroom practice. Multilingual Matters.
- van Lier, L. (2004). The ecology and semiotics of language learning: A sociocultural perspective. Kluwer Academic Publishers. <https://doi.org/10.1007/1-4020-7912-5>
- Wenger, E. (1998). Communities of practice: Learning, meaning, and identity. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803932>

Xi, L., Yuan, Z., YunQui, B., & Chiang, F.-K. (2017). An investigation of university students' classroom seating choices. *Journal of Learning Spaces*, 6(3).

Appendix A

Sample Field Note

Section	Description
Date / Class	Week 6, Core English 2
Seating Arrangement	Small clusters (groups of 3-4 students)
Descriptive Notes	The classroom was arranged into small clusters of three to four students, with desks pushed together to form compact group spaces. Most students entered the room and sat immediately with their usual peers, though one group required prompting to reorganise after initially sitting in a row. Once the activity began, the room quickly filled with low-level conversation. Students leaned toward each other, often pointing at worksheets or using gestures to support meaning. In one group near the centre, three students were actively discussing a vocabulary task. One student appeared to take a leading role, explaining meanings while the others nodded and occasionally asked short questions. In another group near the window, one student remained quiet, looking down at her paper while the other two students spoke. She occasionally smiled but did not contribute verbally during the observed period. Several students used L1 briefly to clarify instructions before returning to English. Laughter was heard intermittently, particularly in groups where students appeared more comfortable with each other. The teacher moved between groups, briefly stopping to listen or prompt discussion, but did not intervene extensively. Overall, most groups remained on task, with visible engagement across the majority of the class.

Reflective Notes	The cluster arrangement appeared to create a noticeably different classroom atmosphere compared to previous row-based lessons. The room felt more active and socially connected, with students orienting their bodies toward one another rather than toward the front. This physical orientation seemed to reduce the sense of formality and shift the focus from teacher-led instruction to peer interaction. I was struck by how quickly students began speaking once the activity started, suggesting that the arrangement itself lowered the threshold for participation. At the same time, participation within groups was uneven. In several clusters, one student naturally assumed a more dominant role, guiding the discussion and explaining answers. While this sometimes supported group progress, it also appeared to limit opportunities for quieter students to contribute. The quieter student near the window stood out to me; although she seemed attentive and not disengaged, her silence might easily have gone unnoticed if I had not been observing closely. This raises an important tension: while clusters can create a sense of safety and reduce the pressure of whole-class speaking, they may also allow some students to remain peripheral without being directly encouraged to participate. Another observation relates to language use. The brief use of L1 seemed to function as a supportive tool rather than a barrier, helping students negotiate meaning before returning to English. In this sense, the clusters appeared to facilitate collaborative problem-solving rather than strictly controlled language practice. This aligns with my broader goal of encouraging communicative confidence, but it also prompts me to consider how to balance fluency and target language use more effectively. From a teaching perspective, I experienced less direct control over the classroom compared to row seating. Instead of addressing the whole class at once, I relied on circulating and listening in on small groups. While this allowed for more personalised interaction, it also meant that I could not fully monitor all groups simultaneously. I became aware that my presence influenced group dynamics, students tended to shift into more “performative” English when I approached, which may not fully reflect their typical interaction. Overall, this session reinforced the idea that clusters support interaction and peer learning but require careful facilitation. In future lessons, I may experiment with assigning rotating roles within groups or providing more structured turn-taking prompts to ensure more balanced participation.
-----------------------------	---

Appendix B

Teacher Reflective Journal Prompts/Sample Entry

1. What seating arrangement did I choose today, and why?
2. How did students respond to this arrangement?
3. Which students seemed engaged or disengaged?
4. What surprised me about today's interaction?
5. What would I change next time?

Date / Class: Week 6, Core English 2

Seating Arrangement: Small clusters (groups of 3–4 students)

1. What seating arrangement did I choose today, and why?

I arranged the classroom into small clusters of three to four students to encourage peer interaction and reduce the pressure associated with whole-class participation. The lesson focused on vocabulary discussion and meaning negotiation, so I wanted students to rely on each other rather than on me as the primary source of input. I also hoped that smaller group sizes would make participation feel more manageable, particularly for quieter students.

2. How did students respond to this arrangement?

Overall, students responded positively. The room became active very quickly, and most groups engaged in discussion without needing much prompting. Compared to row-based lessons, there was a noticeable increase in student talk and a more relaxed atmosphere. Several students appeared more comfortable speaking, particularly those who are usually silent during whole-class interaction. However, engagement varied within groups, while some students contributed actively, others remained more passive.

3. Which students seemed engaged or disengaged?

Many students appeared engaged, especially those in groups where interaction was balanced and collaborative. I noticed that some quieter students spoke more than usual, which suggests that the smaller group setting reduced their anxiety. At the same time, at least one student in a group near the window remained largely silent throughout the activity. She appeared attentive but did not actively contribute, which raises concerns about how easily silence can be overlooked in group work.

4. What surprised me about today's interaction?

I was surprised by how quickly students began interacting once the activity started. There was very little hesitation, which suggests that the seating arrangement itself played a role in lowering the barrier to participation. I was also struck by how naturally some students took on leadership roles within their groups. While this helped sustain the conversation, it may have unintentionally limited opportunities for others to speak.

5. What would I change next time?

In future lessons, I would like to introduce more structure within the clusters to support more balanced participation. For example, assigning roles (e.g., facilitator, note-taker, speaker) or using turn-taking prompts could help ensure that quieter students have space to contribute. I would also like to be more intentional in monitoring quieter students and gently prompting their involvement. Additionally, I may experiment with occasionally reshuffling groups to prevent fixed interaction patterns from forming.

Appendix C

Student Reflection Prompt

“How did today’s seating or grouping affect your participation, comfort, or learning? Please write 2–3 sentences.”

Sample Student Reflections

Student Reflection 1

Today group was easier for me. I don’t like speak in front of class, but only 3 people is okay. My friend help me when I don’t know word. I feel little more confident. But sometimes I just listen and not speak.

Student Reflection 2

今日のグループは話しやすかったです。少人数なので、あまり緊張しませんでした。でも、自分から話すのはまだ難しくて、他の人が話しているのを聞くことが多かったです。もう少し自分から発言できるようになりたいと思いました。

Student Reflection 3

I think small groups are good because we can share ideas more freely. I feel pressure in big class discussion, but here it’s more relaxed. Also if I don’t understand something, I can ask my group first. But sometimes one person talks too much and others don’t join.

Note: Appendix samples are presented close to their original form to preserve authenticity of expression. Main-text quotations were lightly edited for readability where necessary, while preserving meaning.

Appendix D: Sample Coding of Student Reflections

The following table illustrates how selected student reflections were coded and interpreted during the thematic analysis process. Initial codes were generated inductively and later grouped into broader themes related to emotional experience, participation patterns, and perceived learning.

Table D1. Example of Coding Process

Student Reflection (Excerpt)	Initial Codes	Category	Theme
“Today group was easier for me... only 3 people is okay... my friend help me... I feel little more confident... sometimes I just listen.”	small group comfort; peer support; increased confidence; passive listening	Emotional Experience; Participation Pattern	Comfort → Partial Participation
「少人数なので、あまり緊張しませんでした... でも、自分から話すのはまだ難しくて...」	reduced anxiety; low pressure; hesitation to initiate; reliance on others	Emotional Experience; Participation Pattern	Reduced Anxiety → Limited Agency
“We can share ideas more freely... more relaxed... I can ask my group first... but sometimes one person talks too much.”	relaxed atmosphere; freedom to speak; peer scaffolding; unequal participation	Emotional Experience; Participation Pattern; Group Dynamics	Supportive Space → Uneven Participation

Analytic Notes

Across these reflections, several recurring patterns can be observed. First, small-group arrangements consistently reduced anxiety, creating a more comfortable environment for communication. Students described feeling “more relaxed” and less pressured, particularly compared to whole-class interaction. This aligns with the broader theme of emotional safety emerging from cluster arrangements.

Second, peer support played a central role in facilitating participation. Students frequently referred to receiving help from classmates or asking questions within the group before addressing the teacher. This suggests that clusters function as micro-communities of learning, where understanding is co-constructed.

However, the data also reveal important tensions. Despite increased comfort, participation remained uneven. Some students described themselves as listening rather than speaking, while others noted that dominant peers could limit opportunities for broader involvement. This indicates that while clusters create conditions for participation, they do not guarantee equal engagement.

Taken together, these coded examples illustrate how emotional experience (comfort, reduced anxiety) interacts with participation patterns (active contribution, passive listening, uneven interaction). These patterns directly informed the development of the study’s conceptual model, in which seating arrangements shape emotional responses, which in turn influence participation and perceived learning.

Nagoya University of Foreign Studies
Center for Language Education and Development

2026