

ISSN 2436-1993

一般社団法人大学英語教育学会 (JACET)

2025 年度

関東支部紀要

Vol. 13

*JACET-KANTO Journal*



## はじめに

JACET 関東支部紀要『JACET-Kanto Journal Vo.13』をお届けいたします。今号には、厳正な査読を経て採択された論文 2 本および研究ノート 2 本を掲載しています。

今号に並んだ 4 本の論考を概観しますと、現在の英語教育研究が向き合っている重要なテーマが浮かび上がってきます。まずは「学習者の自律性」です。自己効力感といった心理的要因、ポートフォリオを用いたリフレクション、自己評価プロセスなど、学習者が自身の学びをいかに内省し、調整していくかという、学びの主体性に光が当てられています。もう一つのテーマは「ライティング」です。教員による添削やフィードバック、ループリックの活用といった従来の課題のみならず、機械翻訳や文章作成支援ツールといった最新のテクノロジーとの付き合い方が模索されています。

これらのテーマは、当然のように生成 AI の発展と深い関わりを持っています。学習者の自律性は、今後 AI との関係によりその在り方に変化が生じていくかもしれません。また、ライティングは AI 活用と実力養成のバランスが最も悩ましいスキルと言えるでしょう。そうした意味で、今号は今の時勢を鮮やかに反映した一冊と言えるかもしれません。

最後になりますが、今号の刊行にあたり貴重な研究成果をご投稿いただいた皆さま、そして多忙な中で真摯に査読をお引き受けいただいた査読委員の先生方、そして貴重な成果を会員の皆さまに届けるべく尽力いただいた紀要編集委員会の先生方に、心より御礼申し上げます。

2026 年 3 月 31 日

JACET 第 6 代関東支部長 鈴木彩子

# 目 次

はじめに .....	JACET 関東支部長 鈴木 彩子... 1
目次 .....	2
<b>【論文】</b>	
・ The Effect on the Development of Writing Proficiency and Self-Reflection Through Bi-Semester Self-Assessment .....	OI (SUGANUMA) Yoko... 3
・ 英語学習に対する自己効力感尺度と“英語力”の関連：日本の英語教育研究で使用される尺度の傾向と社会的比較の影響 .....	藤田 恵里子... 21
<b>【研究ノート】</b>	
・ 文章作成支援ツールの使用による英作文の書き直しに対する日本人英語学習者の認識 .....	甲斐 順... 41
・ 機械翻訳の使用が日本人大学生の英語ライティング学習に与える影響：脳機能イメージングの観点から .....	佐竹 幸信... 57
<b>JACET 関東支部・紀要投稿規定（第 14 号） .....</b>	<b>78</b>
<b>JACET 関東支部・紀要募集要項（第 14 号） .....</b>	<b>80</b>
<b>JACET 関東支部紀要第 13 号 編集委員会・査読者一覧 .....</b>	<b>84</b>
<b>編集後記.....</b>	<b>85</b>

# The Effect on the Development of Writing Proficiency and Self-Reflection Through Bi-Semester Self-Assessment

Yoko OI (SUGANUMA)

*Seisen University*

## Abstract

This study examined changes in writing proficiency and self-reflection across two semesters in an academic writing class. A total of 55 Japanese university students compiled portfolios and wrote reflective comments after each assignment, guided by five criteria: task fulfilment, organization/coherence, lexical usage, grammatical accuracy, and overview. At the end of each semester, students reviewed their portfolios and provided free-form self-assessments. Text mining was used to analyze these reflections, comparing changes between semesters. Additionally, two raters evaluated students' first, fifth, and tenth writing tests using the same criteria. A one-way repeated ANOVA showed a significant development in writing performance between the beginning and end of the year. Students increasingly emphasized organization and coherence, noting improvements in expressing arguments and opinions. While grammar and vocabulary remained ongoing challenges, students reported greater self-awareness and strategic planning for future writing. The bi-semester self-assessment process enabled learners to recognize their progress and fostered a sense of accomplishment. However, further research is recommended to explore effective integration of portfolios into academic writing instruction and teacher intervention.

**Keywords:** portfolio, writing, reflection, self-assessment, formative assessment

## 1. Introduction

As the background to this study, it has often been noted that the writing skills of learners of English as a second language (ESL) or English as a foreign language (EFL) receive less attention in both research and teaching than other skills such as speaking (Amiran & Mann, 1982; Lipstein & Renninger, 2007). One key reason for this is the limited time dedicated to English composition in Japanese middle and high schools, and even when writing is taught, it often involves simplistic tasks such as sentence rearrangement or translating Japanese sentences into English, which are still treated as “writing instruction” (Hirose, 2003; Oi, 2019). As a result, many students struggle to adapt to academic writing at university level, where they are required to produce substantial amounts of English text without any prior experience in fundamental paragraph writing (Mulvey, 2016).

On the teaching side, it has been reported that writing classes tend to become monotonous, and educators often struggle to provide effective feedback (Oi, 2019), thus emphasizing the need for innovative tools such as an academic writing portfolio, which can address these challenges for both students and teachers. Improving English writing skills, which has recently been highlighted as an essential productive skill, requires creative and novel approaches, rather than simply repeating the cycle of completing isolated assignments and conducting summative assessments, which makes it difficult to support learners' growth closely. For effective formative assessment, the ideal is to have a system that enables a comprehensive view of each learner's developmental process, provides appropriate feedback at each stage, and guides them towards the next level of learning. Formative assessment is often referred to as "assessment for learning" because it involves monitoring learners' progress, identifying areas where they lack understanding, and using feedback to help them achieve their next learning goals and consolidate their knowledge. Writing activities, in comparison to oral ones, tend to lack active peer interaction within the classroom and are often solitary and monotonous.

As one solution aimed at guiding academic writing classes toward more reflective and developmental activities, a portfolio is expected to enhance learners' metacognitive abilities, allowing them to reflect on their progress, deepen their introspection, and set their own learning goals. The use of a portfolio in academic writing classes might foster more autonomous learning, as it serves as a foundation for continuing self-directed study even after graduating from university. Therefore, the aim of this study is to investigate learners' awareness of academic writing through bi-semester self-assessment, while also focusing on the process of learning English writing, and it aims to develop a portfolio that facilitates effective formative assessment, the result of which will be beneficial for both teachers and learners.

## **2. Literature Review**

### **2.1 Portfolio**

Since the 1980s, writing instruction methods grounded in cognitivism have garnered significant interest. These methods focus on the writing process rather than the final written product in terms of cognitivism, emphasizing how writers articulate their thoughts in written form. A portfolio is particularly well suited to this approach, as it can make the writer's cognitive processes transparent and visible, and even record the pre-writing phase. Moreover, it facilitates guiding learners in revising or rewriting their work by enabling an awareness of their cognitive position, thereby motivating them to pursue educational objectives through the regulation of affect, motivation, and attitudes, which are essential to fostering self-development and personal goal setting. A portfolio helps learners to be aware of their position and encourages them to strive for learning goals by regulating and controlling affect, motivation, and behavior, which in turn lead to

the support of self-development and self-goals. In other words, a portfolio enhances self-awareness, which is considered to be a state variable, whereas self-consciousness is considered a trait variable (Klenowski, 2010; Lam, 2018; Weigle, 2002; Yin, 2014; Zimmerman & Schunk, 2001). This heightened self-awareness is intrinsically linked to self-regulation, which comprises three components: self-observation, self-judgment, and self-reaction. Reflective opportunities at the conclusion of each term are intended to cultivate students' self-regulatory capacities. In a nutshell, it is believed that learners' awareness through bi-semester self-assessment promotes self-regulation. Zimmerman and Schunk noted the importance of three self-regulation processes, namely self-observation, self-judgment, and self-reactions (Earl, 2013; Earl & Katz, 2008; Lam, 2014), and Zimmerman and Schunk (2001, p. 99) posited that these processes influence a range of cognitive and affective variables, including: concentration and attention; organizing, rehearsal of information to be remembered, and effective use of resources; beliefs about the self, learning tasks, and outcomes; and the experience of satisfaction and pride in one's work.

Learners strive for these goals in an autonomous and self-disciplined way. They shape their self-concepts through continual engagement in goal-oriented processes as proactive agents. To fulfil these aims, learners also engage in self-monitoring and self-evaluation processes to enhance their self-awareness and self-definition, as well as regulating their own developmental trajectories. As learners grow and develop, learning processes, behaviors, and strategies evolve accordingly. It is believed that a portfolio is useful for nurturing self-awareness, self-monitoring, and goal setting, with the significance of the latter lying in its demonstration of learners' metacognitive development, because they reflect on their planning strategies, learning processes, and methods of monitoring and evaluating comprehension or output. In order to acquire the ability to set goals, the capacity for self-regulation is a prerequisite for learners.

As previously stated, a portfolio is expected to develop self-awareness, which is seen as a precursor to self-evaluation and affect in the development of self-system processes. It is considered a state, whereas self-consciousness is seen as a dispositional trait (Sadler, 2010; Zimmerman & Schunk, 2001). In striving toward their goals, learners act with autonomy and self-discipline, as active agents molding and creating their self-concept through the continual engagement of processes in goal-directed activity. Learners engage in self-monitoring their progress and assess their achievements, thereby fostering greater self-awareness and self-definition in regulating and controlling their own self-development process. As learners grow and develop, so do learning processes and behaviors. In essence, the development of self-system structures and processes is assumed to be the fundamental phenomenon that explains the development of learning processes and behaviors, increasing capabilities for regulating and controlling affect, motivation, and behavior, all in support of self-development and self-goals (Paris & Paris, 2001; Zimmerman & Schunk, 2001).

## **2.2 Written Portfolio and ePortfolio**

The adoption of ePortfolio technologies in higher education has increased significantly in recent years, accompanied by the advancement of their functional capabilities. Historically, academic attention has primarily centered on their application for assessment purposes, particularly in relation to predefined criteria or standards (Pelliccione & Dixon, 2008). However, there is growing recognition of their potential as pedagogical tools within teaching environments (Barrett, 2005; Chesney & Marcangelo, 2010). To fully realize this potential, the importance of a systematic approach to implementation and ongoing evaluation has been emphasized, with a particular focus on pedagogical integration (Roberts et al., 2016; Shepherd & Hannafin, 2011). Nevertheless, as Housego and Parker (2009, p. 409) caution, ePortfolios should not be viewed as a universal solution for promoting student reflection. It has thus been recommended that online teaching and learning environments be strategically developed, tested, and critically reviewed to ensure that such technologies are meaningfully embedded within the learning process (Phillips et al., 2012). Consequently, the use of ePortfolios for written self-reflection may give rise to certain concerns (Roberts, 2018). When students are required to submit their reflections as part of formal university assessments, the reflective writing process may shift from a personal exploration to a task driven by external expectations. When students are required to submit their reflections as part of formal university assessments, they may feel compelled to conform to standardized formats and content in order to achieve favorable grades (Hobbs, 2007; Roberts et al., 2014). As a result, the reflective writing process may shift from a personal exploration to a task driven by external expectations and, as such, the intrinsic value of the reflective process can be compromised.

While portfolios have been introduced into educational contexts to foster self-reflection, the specific effectiveness of ePortfolios in promoting reflective practice remains underexplored. Therefore, the present study employed handwritten self-reflection entries to analyze changes in students' reflective abilities across two semesters.

While the ability to listen and speak a language tends to develop naturally, writing skills—even in one's first language—are generally acquired through explicit instruction and socialization within the cultural context to which one belongs. Writers must learn to adhere to the rhetorical conventions specific to their culture and develop the specialized skills required for processes such as planning, drafting, and revising. These challenges become even more pronounced when writing in a second or foreign language. In this context, portfolio-based approaches may support language learners in enhancing their writing skills by fostering greater awareness of how to improve their overall proficiency.

### **2.3 Reflection**

The concept of reflection has been defined in various ways, particularly in descriptions of its cognitive and developmental processes (Rogers, 2001). Strampel and Oliver (2010, p. 973) describe reflection as “a complex process that strongly influences learning by increasing understanding, inducing conceptual change, and promoting critical evaluation and knowledge transfer.” Reflection and self-reflection are deeply intertwined with learning, serving as catalysts for the development of one’s capacity for autonomous action. Moreover, they are essential for both the evaluation and regulation of learning, as they involve assuming personal responsibility for one’s thinking and behavior. This study adopts Korthagen’s definition of reflection (1999, p. 193), which frames it as “the mental process of structuring or restructuring an experience, a problem, or preexisting knowledge or recognition.” In this context, self-reflection is understood as the capacity to critically assess one’s own strengths and weaknesses, and to identify learning challenges or opportunities. It involves both the ability and willingness to engage in personal introspection—for example, by questioning the impact of a specific event, thought, or experience on oneself.

### **3. Research Questions**

The following two research questions are proposed:

1. How did students’ writing proficiency develop through bi-semester self-assessments by compiling a portfolio?
2. How did students’ self-reflection evolve through bi-semester self-assessments by compiling a portfolio?

### **4. Method**

#### **4.1 Participants and Materials**

The study was conducted over two semesters in the 2024 academic year. The participants comprised 55 university students (18 male, 37 female), aged from 18 to 20 years, who majored in education and psychology in academic writing classes. The students were expected to compile all of their English writing for academic writing classes in a portfolio. A total of 12 types of writing task were given to students, who also received lessons on topics and categories such as narrative and argumentative writing before completing their written texts. Table 1 presents a list of the writing tasks that were to be completed in class. For the final submission of each writing task, students were asked to write between 200 and 300 words in 20 minutes without using dictionaries; these tasks were set as tests in class. Two raters evaluated students’ writing based on five analytical components, each of which was worth four points, amounting to a maximum of 20 points: 1) task fulfillment; 2) organization/coherence; 3) lexical use; 4) grammatical accuracy; and 5) overview. At the end of each semester, students were asked to review and evaluate their whole semester’s writing based on the same five assessment criteria

as the above. They also had to write reflective comments on the portfolio, including what they found and thought based on these five criteria.

**Table 1**

*Writing Tasks*

The first semester	The second semester
1. Self-introduction	7. Problem solution about health
2. Describe others*	8. Classification (music and sports)
3. My hometown	9. Definition of education
4. Last weekend	10. Write your opinion
5. Write an email to your friend	11. Reasons for poverty
6. Similarities between two countries**	12. Reasons for later marriage***

*Note.* \* was used as the first writing test; \*\* was used for the second writing test; \*\*\* was used for the final test.

#### 4.2 Data Analysis

The study adopted a mixed-method approach to analyze the data to discover how a writing portfolio could help students to deepen their self-reflection and develop writing proficiency.

**Quantitative Analysis.** Each analytical component and the sum of evaluations assessed by raters were analyzed by conducting a one-way repeated ANOVA in order to examine the change in the average scores between the first, second, and final writing proficiency tests through a bi-semester writing portfolio. In the analysis, “Describe others” was chosen as the first writing test, “Similarities between two countries” was used for the mid-writing test, and “Reasons for later marriage” was selected for the final writing test.

**Qualitative Analysis.** The descriptive responses written in Japanese or English for the five analytical criteria were analyzed using a quantitative text analysis method known as “text mining,” which involves exploring textual data and identifying patterns utilizing computational techniques. The reflections originally written in Japanese were translated into English by the researcher and subsequently reviewed by each respective student to ensure the accuracy of meaning. Final approval of the translations was obtained from the students.

The advantage of text mining lies in its ability to enhance reliability by facilitating data exploration to determine the most frequently occurring words and by providing precise statistical insights (Higuchi et al., 2023). On the other hand, according to Oki (2018), text mining has certain drawbacks, such as “the risk of overlooking important information” and “the risk of misinterpreting the text’s

content” (p. 259). To mitigate these risks, the analysis underwent a review by research collaborators.

For the text mining process, KH Coder (<https://kncoder.net/>) was used to analyze the patterns in descriptive replies following the methodology outlined by Higuchi et al. (2023). The analytical procedure consisted of two main steps: The first was to verify the results for the words automatically extracted by KH Coder; and the second was to identify and analyze specific concepts that were of particular interest to the researchers. The frequent occurrence results from the initial stage were used as a reference for word extraction. Modifications were made to the default settings of co-occurrence network analysis, including changing the aggregation unit to sentences and adding negation auxiliary verbs to the part-of-speech-based word selection criteria. Additionally, adjustments were made to visualize stronger co-occurrences with thicker lines and to display the results in grayscale. In the analysis, there was a focus on co-occurrence patterns and concordance, with particular attention paid to the relationships between strongly connected words. Through this, we examined the changes in concepts and perceptions regarding the advising process among research participants. After quantitative and qualitative analysis, the results were triangulated in order to find the effect of the writing portfolio.

## **5. Results**

### **5.1 How Did Students’ Writing Proficiency Develop Through Bi-Semester Self-Assessments by Compiling a Portfolio?**

A one-way repeated-measures ANOVA was conducted to examine the change in writing proficiency through bi-semester self-assessments by compiling a portfolio. Only “Organization and Coherence” did not present any significant difference between the first, second, and final tests, but the other four analytical components and the total sum showed a significant difference between the conditions (Table 2). Bonferroni’s multiple comparison test was also employed, and it was found that the scores increased significantly with each successive stage, i.e., from the initial to the intermediate stage, indicating an overall upward trend in writing proficiency throughout the academic year. Notably, the improvement from the initial to the intermediate stage was particularly marked across all evaluated criteria except “Organization and Coherence.” It was found that the scores of the total sum, “Task fulfillment” and “Vocabulary,” might improve from the first to the second test, but the second and final tests did not present any notable improvements. In other words, compared to the first test, the scores of the second and final tests did not show any major development. The improvement of scores in “Grammar” was not found between the second and final test, but there was an increase between the first and the final test. In short, writing proficiency might develop, especially from the beginning to the end of the first semester, but such improvement was not clearly found from the middle to the end of the second semester.

**Table 2***The Results of One-way ANOVA*

	Occasion	<i>n</i>	<i>m</i>	<i>SD</i>	F	<i>p</i> value	Multiple Comparison
Task Fulfillment	First	42	3.07	.34	3.64	.030	First < Mid
	Mid	42	3.36	.62			First <
	Final	42	3.36	.69			Final
Grammar	First	42	3.38	.49	6.20	.003	First <
	Mid	42	3.52	.51			Final
	Final	42	3.74	.45			
Vocabulary	First	42	2.79	.87	26.91	<.001	First < Mid
	Mid	42	3.98	.87			First <
	Final	42	3.79	.90			Final
Organization & Coherence	First	42	4.93	.46	1.00	.370	<i>n.s.</i>
	Mid	42	5.00	.00			
	Final	42	5.00	.00			
Total Sum	First	42	14.17	1.15	28.16	<.001	First < Mid
	Mid	42	15.86	1.46			First <
	Final	42	15.88	1.23			Final

*Note.* The number of students who completed all three writing tasks was 42 out of 55 (data are missing for the remaining participants).

## 5.2 How Did Students' Self-Reflection Evolve Through Bi-Semester Self-Assessments by Compiling a Portfolio?

In accordance with the four analytical components, and the overview self-assessment, the dominant descriptive patterns and word co-occurrences in students' responses were examined.

**Task Fulfillment in the First and Second Semesters.** Table 3 displays the ten most frequently used words in students' self-assessments over the two semesters. The top five words relating to task fulfillment remained largely consistent, with the exception of "content," which rose to fifth place in the second semester. This suggests a shift in students' writing focus from merely completing assignments to placing increased emphasis on the quality and depth of their written content.

**Table 3***Most Frequently Appearing Words Related to Task Fulfillment*

Ranking	First semester	Frequency	Ranking	Second semester	Frequency
1	task	30	1	task	32
2	think	24	2	submission	24
3	write	16	3	write	22
4	submission	15	4	think	16
5	myself	11	5	content	13
6	class	11	6	myself	9
7	accomplish- ment	11	6	accomplish- ment	9
8	sentence	11	8	hand in	8
9	every time	9	8	sentence	8
10	the second term	8	10	work on	7
10	designation	8			
10	time	8			
10	can write	8			
Sum		170			148

The KH Coder map for the first semester illustrates co-occurrence relationships among words in students' responses, with "write," "task," "think," and "accomplishment" exhibiting strong interconnections. In the second semester, although "task" and "submission" remained frequent, the co-occurrence map reveals more intricate interrelations, indicating that students' engagement evolved toward a more interconnected and reflective approach. "Writing" is directly associated with "the opening," which further links to "expression," "sentence," and "grammar." A noteworthy connection between "satisfaction" and "myself" is observed in the first semester but does not persist into the second.

**Organization/Coherence in the First and Second Semesters.** The main difference in the coding for organization/coherence between the end of the first semester and the end of the second lies in the appearance of "sentence" and "topic" codes in the first term, whereas "awareness" and "difficulty" codes emerge in the second, with "sentence" disappearing from the top ten. This suggests that at the end of the first term, learners were still heavily focused on individual sentences and topics, which represented significant challenges in their writing. By the end of the second term, however, they had developed the ability to reflect on their work from different perspectives, indicating a shift in their cognitive approach. The KH Coder map for the first semester shows the reflection on structure/coherence in the first semester (see Appendix 1), with the most frequently used words being "task," "write," "submission," "class," "time," and

“deadline.” In the second semester, however, reflections on “structure”, “content”, and “coherence” became prominent. Additionally, codes related to the structural aspects of writing, such as “conclusion” and “introduction,” appeared, and words like “reason,” “concreteness,” “assertion,” “opinion,” and “theme” also emerged, reflecting how the writers worked to express their ideas more effectively. This suggests that, over the course of the year-long academic writing class, students demonstrated significant development in their ability to reflect and improve as writers (Appendix 2).

**Table 4**

*Most Frequently Appearing Words Related to Organization/Coherence*

Ranking	First semester	Frequency	Ranking	Second semester	Frequency
1	write	48	1	write	33
2	sentence	46	2	content	32
3	content	25	3	structure	29
4	think	22	4	think	27
5	structure	21	5	coherence	18
6	coherence	17	6	writing	11
7	topic	14	7	class	8
7	writing	14	8	feel	7
9	can write	12	9	awareness	6
			9	difficult	6
Sum		219	Sum		240

**Lexical Usage in the First and Second Semesters.** The findings indicate that by the end of the second semester, students had become more aware of the importance of appropriate vocabulary usage in writing, which is reflected in their frequent use of the term “check.” In contrast, during the first semester, their primary focus appeared to be on completing and submitting their writing tasks, with less attention given to precise word choice. Over time, however, a noticeable shift occurred in their approach: Students moved from emphasizing task completion to engaging more thoughtfully with the nuanced use of vocabulary. This progression suggests a developing awareness and an improvement in their ability to select vocabulary that is more contextually appropriate.

**Table 5***Ranking of the Most Frequently Appearing Words Pertaining to Lexical Usage*

Ranking	First semester	Frequency	Ranking	Second semester	Frequency
1	task	30	1	word	35
2	think	24	2	vocabulary	30
3	write	16	3	use	21
4	submission	15	4	think	14
5	myself	11	5	usage	12
6	class	11	6	write	10
7	accomplish- ment	11	7	many	9
8	sentence	11	7	check	9
9	every time	9	7	appropriate	9
10	the 2 <sup>nd</sup> term	8	10	expression	8
10	designated	8			
10	time	8			
10	can write	8			
Sum		170	Sum		157

**Grammatical Accuracy in the First and Second Semesters.** According to the KH Coder analysis, the most frequent terms for grammar-related self-reflection in the first semester were “grammar,” “many,” and “mistakes,” while in the second semester, “grammar,” “think,” and “write” were predominant, with “mistake” dropping to sixth place. Notably, the first semester includes specific grammatical terms such as “tense,” “subject,” “present,” and “verb,” which disappear in the second semester. Additionally, “review” appears in the first semester but is replaced by broader terms such as “accuracy,” “comprehension,” “sentence,” and “writing” in the second. This shift suggests a developmental trajectory in students’ understanding—from isolated grammar points to a more integrated and conceptual approach to grammatical accuracy.

**Table 6***Most Frequently Appearing Words Related to Grammatical Accuracy*

Ranking	First semester	Frequency	Ranking	Second semester	Frequency
1	grammar	53	1	grammar	43
2	mistake	27	2	think	12
3	many	18	3	write	11
4	think	13	4	usage	9
5	forget	13	5	class	9
6	tense	12	6	mistake	7
7	review	11	7	a little	7
8	subject (S)	9	8	use	6
9	exam	8	8	myself	6
10	mind	8	8	accuracy	6
10	present	8	8	sentence	6
11	verb	7	8	text	6
			8	comprehension	6
Sum		187	Sum		134

The KH Coder map for the first semester reveals that “grammar” was often paired with “not,” suggesting frequent expressions of negative self-evaluation. Other specific terms like “plurality,” “forget,” and “preposition” also indicate a focus on discrete grammatical items. In contrast, the second semester reflections note habitual revision and reconnecting with high school grammar knowledge. Students’ reflections thus evolved from identifying surface-level errors to engaging in a deeper analysis of their writing process and underlying linguistic comprehension.

In summary, at the end of the first semester (the second writing assessment), students tended to pay attention to specific components such as sentence and language use through completing a writing portfolio. It was also found that their awareness shifted towards the quality of writing and broader perspectives, including mentioning those of writers. Therefore, it is suggested that students might develop a more sophisticated understanding of themselves as writers, and their metacognitive knowledge or awareness had developed by the end of the second semester.

**Table 7**

*Changes in Learners' Awareness of Academic Writing Through Bi-Semester Self-Assessment*

	Task fulfillment	Organization/coherence	Lexical use	Grammatical accuracy	Overview
First semester	Completion of the task	Sentence topic	General awareness of tasks	Specific grammatical items	English grammar, focus on language use
Second semester	Quality, substance of the written context	Awareness, difficulty	More specific awareness	Broader understanding of grammar and its application, learning attitudes towards grammar	Substance of writing

A comparison of students' first and second semester reflections on their writing activities reveals a shift in awareness towards writing in English (Table 7). Notably, students recognized the challenges of writing, as indicated by "difficult" appearing in fifth place in the reflections of the second term, while "difficulty" was absent from the first term's rankings. Additionally, terms like "content" and "myself" emerged in the second term, suggesting a greater focus on the substance of writing and the writer's personal perspective. Unlike the second semester, the first showed more frequent use of terms related to "English" and "grammar." These differences highlight a significant shift in students' awareness, moving from a focus on language use to a more sophisticated understanding of themselves as writers.

## 6. Discussion

### 6.1 How Did Students' Writing Proficiency Improve Through Bi-Semester Self-Assessments by Compiling a Portfolio?

The results of the quantitative analysis elucidated the effect of the self-assessment of the first semester through a portfolio, but an improvement in writing proficiency was not found at this stage. It is assumed that self-reflection through a portfolio presented an opportunity to metacognitively observe their writing ability and learning strategy. While the compiling of a writing portfolio helped students to push forward positively to the next step, it might not have affected the

development of their writing proficiency. This is because the writing tasks in the second semester needed social background knowledge, while those of the first semester tended to be more personal topics. The lack of improvement between the second and final writing assessment might have been caused by the level of difficulty of the tasks. Though no clear increase in the assessment scores from the end of the first to the end of the second semester was found, students' perspectives toward their writing seemed to shift from specific awareness—such as in lexical matters—to more writer-focused aspects like the quality of their writing. Reflecting on the whole year's writing tasks led students to a different type of awareness from the first semester. This deeper reflection might make students more conscious of the difficulty involved in writing better, more skillfully, and more coherently. However, it cannot be denied that self-assessment showed the limits to enhancing writing proficiency, though it worked well in helping students to develop by themselves—in the study, teachers evaluated students' writing tasks, but they did not make any descriptive comments on their portfolio. Self-assessment is useful for deepening self-reflection, yet once a certain level has been reached, further progress may be best supported by diagnostic feedback and guidance from a teacher, because no development of writing proficiency was found after the second writing assessment.

## **6.2 How Did Students' Self-Reflection Evolve Through Bi-Semester Self-Assessments by Compiling a Portfolio?**

Based on a comparison of self-assessments in the first and second semesters, it was found that individuals can become aware of their progress, feel satisfied or reflect on their development, and set long-term goals rather than focusing on short-term objectives by observing their growth over a certain period. In particular, experiencing growth over an extended period fosters a sense of fulfillment and accomplishment, because reflective practice through a portfolio encouraged students to expand cognition and reinterpret what they were aware of. Such nonlinear and recursive experience helped students to gain self-confidence and self-esteem; therefore, it is believed that the compilation of a portfolio enabled the enhancement or deepening of self-evaluation.

Some students shared reflective comments regarding the process of compiling their portfolios. One student noted, "I became aware of my writing development and weaknesses through assembling my writing tasks in the portfolio, as it allowed me to trace the short history of my assembling my writing tasks in the portfolio. I could revisit my earlier thoughts, even though I still face difficulties in constructing sentences". This suggests that the portfolio functioned as a tool for metacognitive reflection, enabling the student to review their developmental trajectory and identify future directions for improvement.

On the other hand, no development of writing proficiency was found in terms of quantitative analysis, though the results of the qualitative analysis reflected the different perspectives between the first and second semester. It is

considered that, in order to catalyze a breakthrough from the stagnated growth observed after the previous term, constructive intervention by others is necessary. In other words, self-assessment has limited benefits in terms of improving writing proficiency, because it encourages students to focus only on the positive aspects or to criticize the drawbacks of their writing. In short, collaborative efforts between students and teachers could provide the former with catalytic interventions. In short, it may be beneficial to prepare more specific interventions by teachers in order to deepen learners' reflection and self-assessment. By incorporating such reflective cooperation into a portfolio, individual learners can engage in more effective self-reflection, which may contribute to a higher degree of self-regulation.

## 7. Conclusion

The study found that the changes in self-reflection over two semesters in academic writing classes helped students to become more sophisticated writers. Furthermore, the portfolio allowed learners to be aware of their progress as writers and facilitated self-accomplishment and satisfaction. We need to create a more focused and targeted approach to portfolios with positive intervention on the part of others. By adding such reflective help by others to a personal portfolio, individuals can deepen their self-reflection process, ultimately enhancing their ability to self-regulate and manage their own learning and growth. In particular, the study did not ask teachers to make any comments to students on their self-reflection after each semester. It is anticipated that the effect of teachers' intervention on writing portfolios will be investigated as a further study.

## References

- Amiran, E., & Mann, J. (1982). *Written composition, grades K-12*. Literature Synthesis and Report. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=a3fcd3f8cf9bfa833fee9b051fc41a3308f34657>
- Barrett, H. C. (2005). Researching electronic portfolios and learner engagement. The reflect initiative. <http://www.jstor.org/stable/40015496?origin=JSTOR-pdf>
- Chesney, S., & Marcangelo, C. (2010). 'There was a lot of learning going on'. Using a digital medium to support learning in a professional course for new HE lectures. *Computers & Education*, 54, 701–708. <https://doi.org/10.1016/j.compedu.2009.09.027>
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin.
- Earl, L. M., & Katz, S. (2008). Getting to the core of learning: Using assessment for self-monitoring and self-regulation. In S. Swaffield (Ed.), *Unlocking assessment: Understanding for reflection and application* (pp. 90–104). Routledge.

- Higuchi, K., Nakamura, Y., & Shu, K. (2023). *Ugokashi te manabu! Hajimete no text mining – free software wo mochiita jiyu kijyutsu no keiryō text bunseki* [Learning by doing: An introduction to text mining - Quantitative analysis of open-ended responses using free software]. Nakanishiya Publisher.
- Hirose, K. (2003). Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing, 12*(2), 181–209. [https://doi.org/10.1016/S1060-3743\(03\)00015-8](https://doi.org/10.1016/S1060-3743(03)00015-8)
- Hobbs, V. (2007). Faking it or hating it: Can reflective practice be forced? *Reflective Practice, 8*, 405–417. <https://doi.org/10.1080/14623940701425063>
- Housego, S., & Parker, N. (2009). Positioning ePortfolios in an integrated curriculum. *Education + Training, 51*, 408–421. <https://doi.org/10.1108/00400910910987219>
- Klenowski, V. (2010). Portfolio assessment. In P. Perterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (3rd ed., pp. 236–242). Oxford.
- Korthagen, F. A. J. (1999). Linking reflexion and technical competence: The logbook as an instrument in teacher education. *European Journal of Teacher Education, 22*(2/3), 191–207. <https://doi.org/10.1023/A:1010986404527>
- Lam, R. (2018). *Portfolio assessment for the teaching and learning of writing*. Springer.
- Lipstein, R. L., & Renninger, K. A. (2007). Putting things into words: The development of 12–15-year-old students' interest for writing. In P. Boscolo, & S. Hidi (Eds.), *Motivation and Writing: Research and School Practice* (pp. 113-140). Elsevier.
- Mulvey, B. (2016). Writing instruction: What is being taught in Japanese high schools, why, and why it matters. *The Language Teacher, 40*(3), 3–8. [https://jalt-publications.org/sites/default/files/pdf/the\\_language\\_teacher/40\\_3\\_tlt.pdf#page=5](https://jalt-publications.org/sites/default/files/pdf/the_language_teacher/40_3_tlt.pdf#page=5)
- Oi, Y. (2019). Japanese high school English teachers' perspectives on classroom writing assessment criteria: A needs analysis. *The Bulletin of the Graduate School of Education of Waseda University, 27*(1), 159–176. <http://hdl.handle.net/2065/00063295>
- Oki, H. (2018). Text mining – Tairyo no kijyutsushiki anketo wo bunseki suru [Text mining approach to the analysis of large-scale open-ended survey responses ]. In A. Hirai (Ed.), *Kyoiku/shinri/gengo kei kenkyu no tameno data bunseki – Kenkyu no haba wo hirogeru tokei shuho* (pp. 258–285). Tokyo Shoseki.
- Paris, S. G., & Paris, A. H. (2001). Classroom application of research on self-regulated learning. *Educational Psychologist, 36*(2), 89–101. [https://doi.org/10.1207/S15326985EP3602\\_4](https://doi.org/10.1207/S15326985EP3602_4)

- Pelliccione, L., & Dixon, K. (2008). *ePortfolio: Beyond assessment to empowerment in the learning landscape*. Hello! Where are you in the landscape of educational technology. Ascilite 2008, Melbourne. <https://www.ascilite.org/conferences/melbourne08/progbook/progbook-p1-5.pdf>
- Phillips, R., McNaught, C., & Kennedy, G. (2012). *Evaluating e-learning: Guiding research and practice*. Routledge.
- Roberts, P. (2018). Developing reflection through an ePortfolio-based learning environment: Design principles for further implementation. *Technology, Pedagogy and Education*, 27(3), 313–326. <https://doi.org/10.1080/1475939X.2018.1447989>
- Roberts, P., Farley, H., & Gregory, S. (2014). Authentic assessment of reflection in an ePortfolio: How to make reflection more ‘real’ for students. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and reality: Critical perspectives on educational technology*. Proceedings Ascilite Dunedin 2014 (pp. 441–445). Retrieved from <https://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1834&context=ecuworkspost2013>
- Roberts, P., Maor, D., & Herrington, J. (2016). ePortfolio-based learning environments: Recommendations for effective scaffolding of reflective thinking in higher education. *Journal of Educational Technology and Society*, 19(4), 22–23. <https://researchportal.murdoch.edu.au/esploro/outputs/journalArticle/ePortfolio-Based-learning-environments-Recommendations-for-effective/991005544119907891/filesAndLinks?index=0>
- Rogers, R. R. (2001). Reflection in higher education: A concept analysis. *Innovative Higher Education*, 26, 37–57. <https://doi.org/10.1023/A:1010986404527>
- Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, 35(5), 535–550. <https://doi.org/10.1080/02602930903541015>.
- Shepherd, C. & Hannafin, M. (2011). Supporting preservice teacher inquiry with electronic portfolios. *Journal of Technology and Teacher Education*, 19, 189–207. [https://www.researchgate.net/profile/Michael-Hannafin/publication/279464623\\_Supporting\\_Preservice\\_Teacher\\_Inquiry\\_with\\_Electronic\\_Portfolios/links/58ff57fca6fdcc8ed50da28a/Supporting-Preservice-Teacher-Inquiry-with-Electronic-Portfolios.pdf](https://www.researchgate.net/profile/Michael-Hannafin/publication/279464623_Supporting_Preservice_Teacher_Inquiry_with_Electronic_Portfolios/links/58ff57fca6fdcc8ed50da28a/Supporting-Preservice-Teacher-Inquiry-with-Electronic-Portfolios.pdf)
- Strampel, K., & Oliver, R. (2010). They think they are learning, but are they? Strategies for implementing Web 2.0 to positively impact student learning. In C. H. Steel, M. J. Keppell, P. Gerbic, & S. Housego (Eds.), *Curriculum, technology & transformation for an unknown future*. Proceedings Ascilite Sydney 2010 (pp. 924–935). <http://ascilite.org.au/conferences/sydney10/procs/Strampel-full.pdf>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

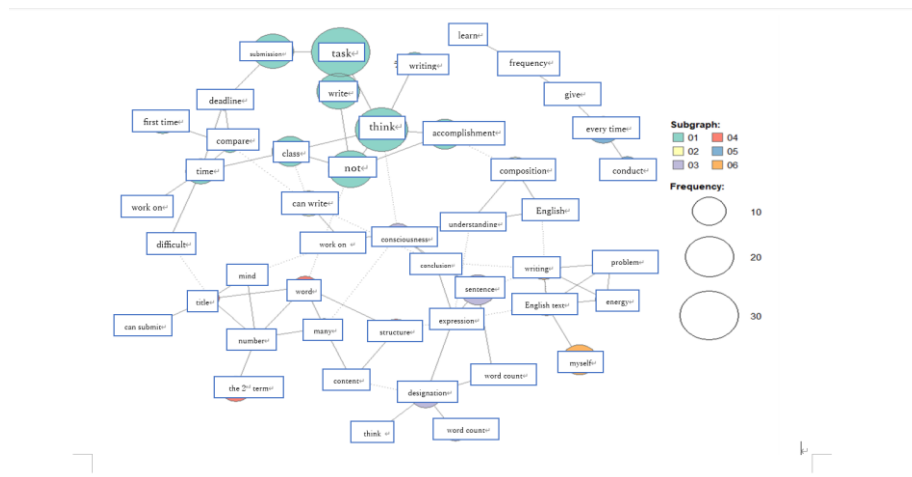
Yin, M. (2014). Portfolio assessment in the classroom. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. II, pp. 659–676). Wiley.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Erlbaum.

## Appendices

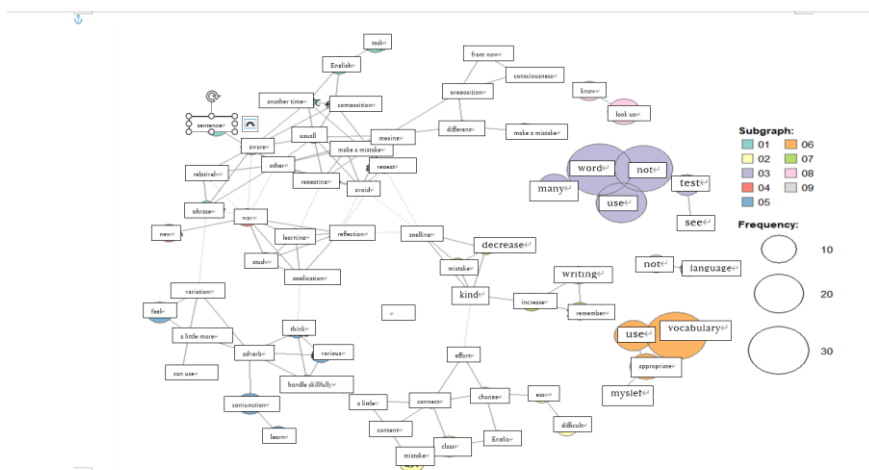
### Appendix 1

#### *KH Coder Map for Lexical Usage in the First Semester*



### Appendix 2

#### *KH Coder Map for Lexical Usage in the Second Semester*



英語学習に対する自己効力感尺度と“英語力”の関連：  
日本の英語教育研究で使用される尺度の傾向と  
社会的比較の影響

Aspects of English Ability Reflected in Self-Efficacy Scales:  
Trends in Self-Efficacy Scales in Japanese EFL Studies  
and the Influence of Social Comparison

藤田 恵里子

(日本大学)

Abstract

Self-efficacy, known to influence academic achievement, has been widely studied in English education in Japan, often to assess instructional effectiveness. This paper first reviews 81 articles to analyze the use of self-efficacy scales in this field, revealing three issues: 1) participants, 2) social comparison, and 3) the indicators for English ability. Incorporating these three points, the second half of the paper examines the correlations between three self-efficacy scales and four indicators of English ability. The scales are: A) Matsunuma (2006), the most frequently used scale; B) Matsunuma with social comparison; and C) Pintrich et al. (1991), originally developed for college students. The results suggest that A) is broadly applicable for measuring self-efficacy related to various aspects of English ability among students with a wide range of proficiency levels, B) is effective in capturing self-efficacy in English proficiency among CEFR A1 learners, and C) is more suited to measure self-efficacy for academic achievement among CEFR A2 learners. Given that most Japanese university students are at CEFR A2 or higher, C) may be the most appropriate for use in college settings. These findings highlight the importance of carefully selecting self-efficacy scales based on learner characteristics when evaluating the effectiveness of English education methods.

*Keywords:* 自己効力感, 尺度, 社会的比較, 英語習熟度, 学業成績

1. はじめに

自己効力感は学習に関わる多くの側面と関連があり、学業成績にも影響することが知られている (松沼, 2004 他)。日本の英語教育の分野においても、指導の効果を検証する際に自己効力感に焦点を当てたものが多く見られる (合田・奥田, 2009 他)。効果を適切に検証するためには、使用する尺度が意図した側面を測定している必要がある。本稿前半では、日本の英語教育分野において用いられている自己効力感尺度を概観する。後半では、前半における気づきを元に用意した 3 つの自己効力感尺度がそれぞれ英語力のどの側面と関連するのか、さらに英語習熟度によってどのような違いがあるのかを検証する。

## 2. 先行研究

### 2.1 自己効力感

自己効力感とは、心理療法の効果を予測する目的で Bandura (1977 他) が提唱した概念で、自身が特定の成果を得るための一連の行動を成功裏に遂行することに対する自信を意味する。高い自己効力感があれば、行動を開始、継続し、困難に直面しても粘り強く努力することができ、さらに難易度の高いタスクを選択するため、より良い成果につながる。その成功体験がさらに自己効力感を高めるという好循環を生む。そのため、学業成績とも強い相関があることが知られている (松沼, 2004; Caprara et al., 2011; Multon et al., 1991; Pajares & Miller, 1994; Pintrich & De Groot, 1990; Siegel et al., 1985)。特に, Multon et al. (1991) は複数の研究のメタ分析を行い、学業成績の低い学習者の方が高い学習者よりも自己効力感と学業成績の間に強い相関を示すことを明らかにした。さらに、自己効力感は人生を決定づける、職業選択や進路決定 (Betz & Hackett, 1983; Hackett, 1985; Lent & Hackett, 1987) にも影響することが示されている。また、自己調整学習、動機づけも自己効力感と関連することが知られている (Pintrich & De Groot, 1990; Zimmerman, 2002; Zimmerman & Martinez-Pons, 1990)。このように、自己効力感は学習の根底にあり、成果を左右すると言っても過言ではないだろう。

英語教育の分野でも、各技能に対する自己効力感が当該技能と関連することが示されている (リスニング: 渡・中島, 2020; Mills et al., 2006; Rahimi & Abedini, 2009, リーディング: Fitri et al., 2019; Mills et al., 2006, スピーキング: Zhang & Ardashaeva, 2019; Zhang et al., 2020, ライティング: Erkan & Saban, 2011; Öztürk & Saydam, 2014; Woodrow, 2011)。

さらに、日本国内の英語教育分野では、指導を通して自己効力感を高める試みが行われ、それぞれ意図した効果が得られている (合田・奥田, 2009; 中山・松沼, 2013; 新本, 2020; 濱田, 2011; 牧野, 2013, 2016, 2021; Fukihara, 2019 他)。その効果を正しく検証するためには、使用尺度が意図した側面を適切に測定していることが不可欠である。

英語学習に関する自己効力感の尺度に焦点を当てた研究は海外において複数見られるが (Gan et al., 2022; Kutuk et al., 2022; Mendoza-Torres et al., 2023; Ocak & Olur, 2018; Sağlam & Arslan, 2018; Wang et al., 2014; Wang & Sun, 2024; Zhang et al., 2023), 日本国内の英語教育分野において学習者を対象とした自己効力感尺度の作成、妥当性・信頼性の検証を目的とした研究は松沼 (2006), 藤田 (2024), 中垣他 (2024) に限られる。しかし、藤田 (2024) は英語学習者の辞書検索行為という限定的な行為に対する自己効力感を測定することを目的としており、中垣他 (2024) は松沼 (2006) を小学生に使用するために文言を調整したのみであった。このように日本の英語教育分野においては、自己効力感に対する関心は高いが、その尺度に対する関心はあまり高くないというのが現状である。

### 2.2 日本国内の英語教育分野において使用されている自己効力感尺度

#### 2.2.1 使用されている自己効力感尺度及び件数

日本の英語教育分野において使用されている自己効力感尺度の傾向を明らかにすべく、以下の検証を実施した。使用言語の影響を回避するため、日本語で実施されたことが想定される、日本国内で実施、発表された研究に限定した。まず、CiNii Research<sup>1</sup> に「英語、自己効力感」の2語を入力して検索を行った。内容の確認ができるように「本文あり」の条件を選んだところ、2025年8月24日時点で、107件が該当した。その中から、発表要旨 (12件)、英語教育分野以外の論文 (3

件), 重複 (1 件), 自己効力感に言及しているものの検証は行っていないもの (9 件), 大学紀要の異なる論文内に英語と自己効力感がそれぞれ含まれているもの (1 件) の合計 26 件を除外し, 81 件を検証の対象とした。これらの論文において, 参考にした, あるいは使用したとして言及があった尺度をまとめたところ, 表 1 の結果が得られた。その際, 一つの論文内でしか言及されていない尺度及び尺度が明記されていない場合は, 「その他」としてまとめ, 引用元や参考文献が示されておらず, 筆者が一から作成したと思われる尺度については「自作」とした。また, Pintrich and De Groot (1990) を引用したとし, 松沼 (2006) への言及はないが, 松沼 (2006) と質問項目が完全に合致した 1 本については, 松沼 (2006) として積算した。さらに, 森 (2004) を参照したと書かれてるものの, 質問項目がほとんど合致しない自作の質問項目になっている論文も 1 本あったが, 言及があることから森 (2004) として数えた。なお, 一つの論文内で複数の尺度に言及していた場合もあったため, 総数は合致しない。

表 1 の結果から, 日本の英語教育分野で使用されている尺度としては, 圧倒的に松沼 (2006) が多く (18 件), Pintrich and De Groot (1990) 及び森 (2004) 各 6 件が続く。詳細は 2.3 において後述するが, 松沼 (2006) 及び森 (2004) はいずれも Pintrich and De Groot (1990) を日本語訳して作成されている。本分析は対象者を日本人とした英語教育分野の研究に限定していることから, Pintrich and De Groot (1990) よりも日本語に訳され, 英語学習に対して使用する目的で妥当性・信頼性の検証が行われている松沼 (2006) の方が引用件数が多いというのは首肯できる結果であろう。

表 1

日本の英語教育分野の論文で使用されている自己効力感尺度

自己効力感尺度	本数	自己効力感尺度	本数
松沼 (2006)	18	伊藤 (2009)	2
自由記述	15	水本 (2011)	2
自作	14	面接	2
Pintrich and De Groot (1990)	6	その他	22
森 (2004)	6	合計	87

### 2.2.2 上位 3 尺度を用いた研究の調査対象者の比較

次に特に引用頻度が高かった 3 つの尺度, 松沼 (2006), Pintrich and De Groot (1990), 森 (2004) を用いた研究の調査対象者を表 2 に示した。表 2 に明らかなように, その多くは大学生を対象としていることがわかる。多くの研究者が大学に属していることから, 自身の指導対象である大学生を調査対象とする傾向があると推察される。

表 2

松沼 (2006), Pintrich and De Groot (1990), 森 (2004) を使用した研究の調査対象

	大学生	高校生	中学生	小学生	合計
松沼 (2006)	15	2	0	1	18
Pintrich and De Groot (1990)	3	2	1	0	6
森 (2004)	4	0	2	0	6
合計	22	4	3	1	30

### 2.2.3 上位 3 尺度を用いた研究で使用されている英語力の指標

これら 3 つの尺度を用いた研究の中で英語力と自己効力感の関連に言及しているものは 10 件あった。英語力の指標には、学業成績 (1 件)、学期末テストのような到達度テスト (1 件)、TOEIC のような英語習熟度テスト (8 件) を用いた検証が混在していた。習熟度テストは、出題範囲が決まっておらず、一般的な英語運用能力を反映しているのに対し、到達度テストは出題範囲が限定されているため高得点であっても必ずしも英語運用能力が高いことを意味しない可能性がある。また、学業成績には課題や授業参加度といった英語力以外の要素も含まれている。このように、先行研究で使用されている英語力の指標が反映している側面にはばらつきが見られた。

## 2.3 松沼 (2006), Pintrich and De Groot (1990), 森 (2004) の尺度の比較

### 2.3.1 各尺度の作成背景

次に日本の英語教育分野において引用頻度が高い 3 種類の尺度の作成背景を明らかにしたい。両日本語尺度の元となった Pintrich and De Groot (1990) は、英語教育に特化したものではなく、動機づけの方向性、自己調整学習、学業成績の関連を検証する目的で、動機づけ、認知方略使用、メタ認知に関する複数の尺度を元に作成され、56 項目で構成されていた。中学生を対象に 7 件法で回答させ、因子分析を行った結果、9 項目が自己効力感を測る質問項目として抽出された。

森 (2004) は、大学生を対象に、現在の自己効力感、中学時代及び現在の英語学習方略の関係を明らかにすることを目的としていた。この際、自己効力感を測定するために、Pintrich and De Groot (1990) で自己効力感として分類された 9 項目を筆者自身が日本語訳し、6 件法で実施した。日本語訳に際して、「被験者が理解するには難しい、あるいは誤解を招きやすい項目には言葉を追加するなどした」(p.46) との記載がある。このことから、本尺度は調査対象者である大学生を念頭に調整されたと考えられる。

松沼 (2006) は 3 つの尺度の中で唯一尺度作成を目的とし、妥当性・信頼性の検証を行っている。この尺度は先述の Pintrich and De Groot (1990) の 9 項目を筆者、中学校、高等学校の英語教員の 3 名で日本語訳し、現職の英語教諭の助言を得て作成された。さらに、「あくまで、個人が、ある事態に対処する時、それをどの程度効果的に処理できると考えているかという認知を問題としているのであって、他者との比較に関する認知を問題としているのではない」(p.91) ことから、他者との比較の文言 *Compared to other students in this class* は削除した旨が記載されている。高校 1 年生を対象に検証が行われ、英語の学期末試験 (学業成績)、英語の学習時間 (努力の持続性)、英語の課題提出数 (内発的動機づけ) を外部基準として用いて、妥当性が認められている。

### 2.3.2 質問項目の比較

表3に、松沼 (2006)、Pintrich and De Groot (1990)、森 (2004) の質問項目を同一の質問と思われるものを並べて提示する。日本語の2つの尺度は Pintrich and De Groot (1990) を元に作成されているため共通点も多いが、次のような違いが見られた。森 (2004) においては、Pintrich and De Groot (1990) では全項目に見られる this class / the subject のような当該科目に限定する文言が削除されている。一方、松沼 (2006) では、他者との比較を促す Compared with other students in this class に相当する文言が削除され、項目5と項目9が同義になったことで1項目少ない8項目になっている。

この文言の違いを原因として、Pintrich and De Groot (1990) は全項目が当該科目に限定した質問群、森 (2004) は全項目が一般的な英語力や学習能力に関する質問群、松沼 (2006) は当該科目に限定した問 (項目2, 3, 4, 7) と一般的な英語力や学習能力に関する問 (項目1, 5, 6, 8) で構成された質問群というように違いが生じている。それぞれ一般的な英語力や学習能力に関する問は英語習熟度と、当該科目に限定した問は学業成績と関連が高いことが予測される。

表3

松沼 (2006)、Pintrich and De Groot (1990)、森 (2004) の質問項目の比較

	松沼 (2006)	Pintrich and De Groot (1990)	森 (2004)
1	私は英語が得意だと思う。	<u>Compared with others in this class</u> , I think I'm a good student.	他の人と比べると、自分はよい学習者であると思う。
2	私は英語の授業で教えられたことを理解することができると思う。	I'm certain I can understand the ideas taught in this course.	教えられる内容を、自分は理解できる方だと思う。
3	私は英語で良い成績をとることができると思う。	I think I will receive a good grade in this class.	自分はよい成績をとれると思う。
4	私は英語の授業で与えられた課題に適切に答えることができると思う。	I am sure I can do an excellent job on the problems and tasks assigned for this class.	授業で出された問題や課題を、自分はいまよくこなせると思う。
5	私の英語の学力はすぐれていると思う。	<u>Compared with other students in this class</u> I expect to do well.	他の人と比べると、自分はよくやれると思う。
6	私は英語の学習内容についてたくさんを知っていると思う。	<u>Compared with other students in this class</u> I think I know a great deal about the subject.	他の人と比べると、自分は授業で学習する内容についてよく知っていると思う。
7	私は英語の学習内容を習得できると思う。	I know that I will be able to learn the material for this class.	授業のレベルについていけると思う。
8	私は英語の勉強方法を知っていると思う。	My study skills are excellent <u>compared with others in this class</u> .	自分の英語学習能力は、他の人に比べてすぐれていると思う。
9	—	I expect to do very well in this class.	自分は、授業でうまくやれると思う。

注. 下線は他者との比較に関する文言。番号は便宜的に付与した。

ここまでに、現状日本の英語教育分野において使用頻度が高い3つの自己効力感尺度の使用状況と各尺度の作成背景を概観し、以下3つの気づきを認めた。まず、尺度を構成する質問項目には、一般的な学習能力や英語力に関する項目と当該科目に限定した項目があり、それぞれ関連がある英語力の側面に違いがあることが予測されるが、これらの尺度を用いた研究において使用されている英語力の指標も習熟度テスト、学業成績、到達度テストとばらつきが見られた。次に、日本語版の元となった Pintrich and De Groot (1990) は中学生、松沼 (2006) は高校生、森 (2004) は大学生を対象とすることを念頭に作成されていたが、必ずしも作成時想定の対象者に対して使われているとは限らなかった。最後に、松沼 (2006) では、他者との比較を促す文言が削除されていた。以降では、特に対象者、他者との比較の2観点についてさらに深く考察する。

#### 2.4 Pintrich の大学生向け MSLQ (Motivated Strategies for Learning Questionnaire)

調査対象者に関して、上述の Pintrich and De Groot (1990) の著者の一人である、Paul Pintrich が同僚と共に大学生を対象として作成した尺度 Motivated Strategies for Learning Questionnaire (MSLQ) が存在する。その作成経緯については Duncan and McKeachie (2005) に詳しい。Pintrich は 1986 年から 5 年間ミシガン大学において、大学生の学習全般に関して研究を行い、MSLQ を作成した。MSLQ は動機づけ尺度 (31 項目) と学習方略尺度 (50 項目) の 2 要因合計 81 項目で構成されている。動機づけ尺度は因子分析の結果、9 つの下位尺度に分類され、そのうち自己効力感に関する質問項目とされたのが 8 項目であった。その後、3 度の検証を経て、最終版が MSLQ manual (Pintrich et al., 1991) としてまとめられている。表 4 に、中学生を対象として作成された Pintrich and De Groot (1990) と大学生を対象に作成された Pintrich et al. (1991) を、含まれるキーワードが似通っているもの同士を並べて記載する。両質問群は、全項目に this class / this course / the subject のような文言を含み、当該科目に限定した質問のみで構成されているという点で一致している。

表 4

Pintrich et al. (1991) と Pintrich and De Groot (1990) の質問項目の比較

	Pintrich et al. (1991)	Pintrich and De Groot (1990)
1	I believe I will receive an excellent grade in this class.	I think I will receive a good grade in this class.
2	I'm certain I can understand the most difficult material presented in the readings for this course.	—
3	I'm confident I can learn the basic concepts taught in this class.	I'm certain I can understand the ideas taught in this course.
4	I'm confident I can understand the most complex material presented by the instructor in this course.	I know that I will be able to learn the material for this class.
5	I'm confident I can do an excellent job on the assignments and tests in this course.	I am sure I can do an excellent job on the problems and tasks assigned for this class.
6	I expect to do well in this class.	I expect to do very well in this class.
7	I'm certain I can master the skills being taught in this class.	My study skills are excellent <u>compared with others in this class.</u>

8	Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this class.	<u>Compared with other students in this class I expect to do well.</u>
9	—	<u>Compared with others in this class, I think I'm a good student.</u>
10	—	<u>Compared with other students in this class I think I know a great deal about the subject.</u>

注. 下線は他者との比較に関する文言。番号は便宜的に付与した。

相違点としては、一方では learn となっていたものが understand になっていたり (項目 3, 4), very などの強調が付与されていたり (項目 6) といった些末な文言の違いや焦点が未来か現在かといった違い (項目 7) が見られるが、特徴的な違いは2つにまとめることができる。まず、大学生用の Pintrich et al. (1991) では「最も難解な教材 (the most difficult / complex material)」を理解できる自信を問う項目が含まれていることである (項目 2, 4)。特筆したい違いとしては、他者との比較を前提とする文言が大学生用に作成された Pintrich et al. (1991) には見られないことである (項目 7, 8, 9, 10)。例えば、項目 8 については、大学生用では自身のスキルを含めた環境要因を判断の前提としているのに対し、中学生用では他者との比較を前提としている。この違いは松沼 (2006) があえて他者との比較に関する文言を削除したという点と合致するが、青年期以降は他者との比較の影響を受けないという傾向があるのだろうか。この点について、2.5 においてさらに考察する。

## 2.5 社会的比較 (social comparison) と自己効力感

本節では、松沼 (2006) 及び Pintrich et al. (1991) から削除されていた他者との比較という観点について考察したい。他者との比較に関して Festinger (1954) が提唱した社会的比較 (social comparison) という理論がある。Festinger (1954) によれば、人は生来自身の能力や意見を評価したいという衝動を持っており、客観的かつ非社会的な方法がない場合、他者との比較に基づいてその評価を行う。そして、比較対象としては自身と類似性が高い他者を選ぶ傾向があるとしている。自己効力感とは、自己評価に基づく自身の能力に対する認識であるため、一見他者とは無関係に自己完結しているもののように見える。しかしながら、社会的比較理論に基づけば、客観的かつ非社会的な方法がない場合、自己の能力に関する評価も他者との比較を参考に行われるはずである。すなわち、自己効力感を考える際に社会的比較は排除できない要素であることが予測される。

事実、自己効力感の提唱者であるバンデューラ自身も社会的比較も自己効力感に影響する重要な要素の一つとして認めおり (Bandura, 1993; Pajares, 1997), その影響が実証されている (Bandura & Jourden, 1991; Bouffard-Bouchard, 1990; Inagaki, 2022; Shunk, 1983;1985; Schunk & Hanson, 1985)。

Schunk は児童を対象とした一連の研究を通して、他者の成功を観察することが自己効力感につながるという vicarious experience<sup>2</sup> の影響を検証している。Schunk (1983) は、小学生を対象に、算数の問題に取り組み、他者との比較に関する情報が自己効力感及び算数スキルの向上に寄与し、さらに目標設定を組み合わせることでその効果が増大することを明らかにした。また、vicarious experience の一例であるモデリングも社会的比較の形態の一つであるとし、類似の他者の成功体験を観察することで自身の自己効力感も向上することが示されている (Schunk, 1985; Schunk & Hanson, 1985)。児童の発達段階による違いとして、7 ~ 8 歳までは社会的比較の影響はないが、

年齢が上がるにつれて社会的比較が自己評価に与える影響が大きくなることが明らかになっている (Ruble et al., 1980)。

次に、大学生以上を対象とした研究に目を向ける。Bandura and Jourden (1991) は大学院生 (21 ~ 49 歳) を対象に、タスクが完了する度に自身と他者の成績を比較して示しつつ、架空の人材マネジメントタスクに全 18 回取り組ませた。その過程で、①自身と他者の成績が同等、②自身の方が優れている、③当初は自身の方が劣っているが次第に追いつく、④当初は他者と同等の成績だが次第に自身が劣っていく、という 4 種のパターンの成績を意図的に協力者に提示した。その結果、④の成績パターンを提示された協力者のみが著しく自己効力感の低下を示した。また、Bouffard-Bouchard (1990) は、大学生に認知タスクに取り組ませた後、教員は各学生に対して他学生と比較して成績が良い、あるいは悪いのいずれかのフィードバックを与えた。その結果、他学生よりも成績が良いというフィードバックを得た学生は自己効力感が向上し、その後のパフォーマンス、自身の解答の正否に関する判断、粘り強さに影響が見られた。実験研究だけでなく、調査研究によっても社会的比較と自己効力感の関係が示されている。Inagaki (2022) は、英語専攻の日本人女子大学生を対象に質問紙調査を行い、継時的比較 (temporal comparison: 過去の自己との比較, Albert, 1977) 及び社会的比較はいずれも自己効力感及び自信と関連があることを示した。また、徳岡・前田 (2011) は大学生を対象に質問紙調査を行い、自己効力感及び英語成績 (TOEIC の点数) が高い学習者は社会的比較と継時的比較の両方を行う傾向を明らかにした。

以上のように、児童だけでなく、大学生であっても他者との比較が自己効力感、ひいては行動や成績にも影響することが明らかになった。特に、Festinger (1954) が指摘するように人間は生来他者との比較をするものであり、自身と類似性が高い他者を比較対象として選ぶ傾向があることから (Collins, 1996; Festinger, 1954)、教室環境においては、クラスメイトとの比較による影響は排除できない要素であろう。大学の教室環境において、先行研究の実験環境のように他者との優劣が明示的に提示されることは想定しがたいが、学習者は協同学習や授業内の行動・発言を通して、無意識的に他者と比較していることが予測される。したがって、大学の教室環境においても、社会的比較が自己効力感に影響する可能性は十分にあるだろう。

### 3. 研究目的

以上を踏まえて、本稿は、英語教育分野における自己効力感尺度として、採用数最多で妥当性、信頼性の検証が行われている松沼 (2006) を質問紙 A、松沼 (2006) に他者との比較を促す文言を付記したものを質問紙 B、大学生を対象として作成された Pintrich et al. (1991) の日本語訳を質問紙 C として 3 つの尺度を用意し、以下の検証を行うことを研究目的とする。

- 1) 各尺度は英語習熟度及び学業成績とそれぞれどのような関連があるのか。
- 2) また、その関連は英語習熟度によってどのような違いが見られるのか。

### 4. 方法

#### 4.1 参加者

本研究では、日本の関東圏にある私立大学において、筆者が担当する英語必修科目を受講している 1, 2 年生 (各 2 クラス) を対象とした。両学年とも卒業に必要な必修科目として、リスニング・スピーキングを中心とする LS 科目と、リーディング・ライティングを中心とする RW 科目の 2 科目を各週 1 回受講しており、いずれも GTEC Academic (ベネッセ) の当該技能のスコアを

元に3レベル (CEFR B1, A2, A1) にクラス分けしている。調査対象であった1年生2クラスはいずれもA1レベルで、日常会話を焦点としたLS科目であった(各29名)。2年生2クラスはともにA2レベルで、1パラグラフ程度のエッセイを焦点としたRW科目(26名)とプレゼンテーションを焦点としたLS科目(31名)であった。計115名が履修登録をしていたが、英語力の指標に関するデータが1つ以上欠損している学生を除外したところ、100名(1年生49名、2年生51名)が分析対象となった。また、質問紙を3回の授業に分けて実施した結果、質問紙Aは92名、質問紙Bは86名、質問紙Cは88名から回答が得られた。なお、調査に先だって、調査協力は任意であり、途中で中止・中断できること、一旦同意した場合も取り消すことができること、入手した情報は個人を特定できない形で保存、使用する旨を説明し、同意を得た。

## 4.2 材料

### 4.2.1 質問紙

本調査を行うにあたって、質問紙A: 松沼 (2006)、質問紙B: 松沼 (2006) に他者との比較の文言を付与したもの、質問紙C: Pintrich et al. (1991) の日本語訳の3種類の尺度(各8項目)を用意した。全質問紙において、「1: まったく当てはまらない ~ 6: とても当てはまる」のリッカートスケール6件法で回答させた(資料)。基準となる質問紙Aとしては、信頼性、妥当性の検証が行われていること、採用件数最多であることから、松沼 (2006) を採用した。

質問紙Bを作成するにあたり、他者との比較を促す文言以外の要素はそろっていた方がより正確な比較結果が得られると考え、松沼 (2006) を土台とした。また、比較対象との類似性が高い方が強い影響を与えるため (Collins, 1996; Festinger, 1954)、本稿で比較を促す対象も類似性が高い「クラスメイト」として、松沼 (2006) の全項目に「クラスメイトと比較して」を付与して作成した。

質問紙Cに関しては、2.4で言及した、大学生対象に作成されたPintrich et al. (1991) において自己効力感尺度として分類された8項目を採用した。日本の英語教育分野の論文において日本語訳が公開されている例が見られなかったため<sup>3</sup>、筆者が以下の手順で作成した。まず、Pintrich et al. (1991) を筆者自身が日本語に訳し、次にChatGPTを用いてバックトランスレーションを行い、元の英文との間に齟齬がないことを確認した。次に日本語として不自然な点がないか確認するため、ChatGPTに「修正は必要最小限とし、日本語として不自然な表現があれば修正してください」というプロンプトを入力し、抽出された情報を参考に日本語訳に手を加えた。その後改めてChatGPTを用いてバックトランスレーションを行い、Pintrich et al. (1991) の英文と齟齬がない内容になっていることを確認した。最後に、Pintrich et al. (1991) においてthis class / this courseとなっていた文言を質問紙Aとそろえて、「英語」に修正した。

なお、各質問紙に関して主成分分析を用いて次元性、クロンバック $\alpha$ を用いて内的整合性を検証したところ、質問紙Aは因子負荷量.69 ~ .88、寄与率.67、 $\alpha = .93$ 、質問紙Bは因子負荷量.81 ~ .94、寄与率.79、 $\alpha = .96$ 、質問紙Cは因子負荷量.71 ~ .89、寄与率.66、 $\alpha = .93$ であり、いずれの質問紙も十分な次元性及び内的整合性が認められた。特に質問紙Bは因子負荷量もクロンバック $\alpha$ も高かったのは、「クラスメイトと比較して」という同一の文言が全ての質問項目に含まれていたことが影響しているかもしれない。

## 4.2.2 英語力に関する指標

各自己効力感尺度が英語力に関するどの側面と関連するかを検証するために、英語習熟度及び学業成績の指標を用意した。具体的には、英語習熟度の指標として、調査対象校においてクラス分けのために使用している①GTEC Academic (4技能各250点、1000点満点)の4技能の合計点を採用した。学業成績の指標としては、筆者自身が担当する科目における、②素点(課題、授業参加度、各種試験を含む複数の評価項目の合計100点満点)、③焦点(当該科目で焦点とした技能の定着度を測る学期末試験)、④筆記(当該科目内で学習した文法・語彙の知識の定着を測る学期末筆記試験)を用いた。③焦点に関しては、1年生LS科目(焦点:日常会話)は、教員と1対1の会話試験、2年生RW科目(焦点:エッセイ)は英文エッセイ、2年生LS科目(焦点:プレゼンテーション)はグループプレゼンテーションの得点をそれぞれ用いた。③焦点及び④筆記は、②素点の一部であり、各100点満点に換算して分析に用いた。

## 4.3 手順

各調査対象者には、基本的に3種類の質問紙すべてに回答をさせた。その際、実施の順番が影響しないよう配慮し、先に実施した質問紙の影響を最小限にするために各質問紙間に1か月以上の期間を設け、第1回授業、第6回授業、第11回授業で実施した(表5)。特に、クラスメイトとの比較を促す質問紙Bについては、クラスメイトとの接触が十分ではないと考えられる第1回授業では実施せず、第6回と第11回において実施した。

表5

各質問紙実施時期

	第1回授業	第6回授業	第11回授業
クラス1(2年生LS)	質問紙A	質問紙C	質問紙B
クラス2(2年生RW)	質問紙C	質問紙B	質問紙A
クラス3(1年生LS)	質問紙A	質問紙B	質問紙C
クラス4(1年生LS)	質問紙C	質問紙A	質問紙B

## 5. 結果と考察

### 5.1 各項目の平均値の比較

GTECの4技能の合計点(各技能250点、合計1000点満点)及び学業成績関連3項目(各100点満点)の平均値を表6に、各質問紙の平均値を表7にそれぞれ示した。

英語力に関して、調査対象者全体の①GTECの平均値336.11( $SD=111.39$ )で高低群に分けたところ各群50名になった。各群の①GTECの平均値は、高群(H群)は437.16( $SD=38.87$ )、低群(L群)は235.06( $SD=52.13$ )であり、対応なしの*t*検定を行ったところ両群間の英語力に有意な差が認められた( $p<.001$ )。同様に、②素点( $H=84.04, L=78.12, p<.001$ )、③焦点( $H=77.53, L=68.81, p=.03$ )、④筆記( $H=65.91, L=58.25, p=.002$ )の全項目においてH群の方が統計的に有意に高いという結果になった。

表 6

英語力 4 観点の平均値 (記述統計)

	①GTEC		②素点		③焦点		④筆記	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
H 群	437.16	38.87	84.04	8.97	77.53	10.40	65.91	11.57
L 群	235.06	52.13	78.12	8.34	68.81	25.26	58.25	12.40
全体	336.11	111.39	81.08	9.12	73.17	19.71	62.08	12.53

注. H 群:  $n = 50$ , L 群:  $n = 50$ 

表 7

各質問紙の平均値 (記述統計)

	質問紙 A			質問紙 B			質問紙 C		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
H 群	45	3.07	0.84	44	2.93	1.04	42	3.26	1.03
L 群	47	2.52	1.00	42	2.82	1.08	46	2.75	1.03
全体	92	2.79	0.96	86	2.87	1.06	88	2.99	1.06

注. 質問紙 A ~ C は最大値 6

質問紙に関しては、回答者にばらつきがあるため、回答の平均を従属変数、群 (H / L) 及び質問紙 (A ~ C) を固定効果、学生 ID をランダム効果に指定し、線形混合モデルを用いて分析した。その結果、群間 ( $p = .02$ )、質問紙間 ( $p = .04$ ) には有意差が見られたが、交互作用 ( $p = .18$ ) は見られなかった。ボンフェローニの補正を用いて多重比較を行った結果、質問紙 C は質問紙 A よりも有意に高い ( $p = .03$ ) ことがわかった。各質問紙の推定周辺平均を群ごとに表 8 に示す。

表 8

推定周辺平均 (線形混合モデル)

		平均値	標準誤差	自由度	95% 信頼区間	
					下限	上限
質問紙 A	H	3.01	0.14	162.59	2.72	3.29
	L	2.53	0.14	157.80	2.25	2.81
質問紙 B	H	3.04	0.15	167.76	2.75	3.33
	L	2.83	0.15	170.45	2.54	3.12
質問紙 C	H	3.26	0.15	170.50	2.97	3.55
	L	2.73	0.14	160.13	2.44	3.01

線形混合モデルにおいて交互作用は非有意であったが、補足的に、記述統計の平均値を用いて、各質問紙において H/L 群間に有意差がないか対応なしの  $t$  検定で検証した。その結果、質問紙 A (H = 3.07, L = 2.52,  $p = .005$ ) 及び 質問紙 C (H = 3.26, L = 2.75,  $p = .002$ ) に関しては、H 群の方が統計的に有意に高かったが、質問紙 B (H = 2.93, L = 2.82) については有意差が認められなかった。

この違いは、質問紙 B では習熟度が同等であることが想定されるクラスメイトと横の比較をしているのに対し、比較対象の指定がない質問紙 A 及び C では配置されたクラスの優劣、つまり同学年の全学生と縦の比較をしていることに起因する可能性が考えられる。

また、H/L 群内において質問紙 A ～ C 間に有意差が見られるペアがないか線形混合モデルを用いて分析した結果、両群ともに有意差が見られる組み合わせはなかった。同様に、各質問紙に関して回答順による有意差も認められなかった。しかし、他者との比較の有無で対応がある質問紙 A と質問紙 B に焦点を当てると、記述統計上は H 群は他者比較なしの方が高いのに対して、L 群は他者比較ありの方が高かった (表 7)。同一クラスの学生と比べた場合、H 群は自身がそれほど優れているわけではないと認識するのに対し、L 群は自身の方が優れていると認識する傾向があるとと言えるだろう。

## 5.2 質問紙 A ～ C と英語力の 4 観点の相関

次に、調査対象者全体及び英語習熟度別 H/L 各群に関して、英語力に関する 4 観点①GTEC、②素点、③焦点、④筆記と質問紙 A ～ C の間の相関関係を明らかにするために、スピアマンの相関係数を求めた (表 9)<sup>4</sup>。

特徴的な点は、③焦点は、調査対象者全体において質問紙 B と弱い相関 ( $r = .21, p < .05$ ) があるのみで他の組み合わせでの相関は見られないのに対し、④筆記は、L 群において質問紙 B 及び C と相関が見られないが、他の組み合わせでは比較的高い相関を示していることである。このことは、調査対象とした全科目の焦点となる技能がアウトプットであるのに対して、④筆記は文法や語彙の知識の定着を測るインプットであったことと関連があるかもしれない。すなわち、調査対象者はこれまでインプット中心の英語学習をしてきたために、自己効力感もインプットの能力を根拠としている可能性がある。そのため、アウトプットの能力と自己効力感尺度の間に相関が見られなかったと考えられる。以降、質問紙別に傾向を検証する。

質問紙 A は、対象者全体において、①GTEC ( $r = .31, p < .01$ )、②素点 ( $r = .32, p < .01$ )、④筆記 ( $r = .33, p < .01$ ) とそれぞれ相関が見られた。H 群に関しては、②素点 ( $r = .39, p < .01$ ) と④筆記 ( $r = .30, p < .05$ ) は相関するが①GTEC とは相関が見られなかった。一方、L 群は①GTEC ( $r = .36, p < .05$ ) 及び④筆記 ( $r = .32, p < .05$ ) と相関が見られるが、②素点とは相関が見られなかった。

質問紙 B は、対象者全体では②素点 ( $r = .26, p < .05$ )、③焦点 ( $r = .21, p < .05$ )、④筆記 ( $r = .35, p < .001$ ) と相関が見られた。L 群では①GTEC ( $r = .46, p < .01$ ) のみではあるが、他と比べて高めの相関が見られた。H 群は④筆記 ( $r = .48, p < .01$ ) とは比較的高めの相関が見られたものの、①GTEC も②素点も相関が見られなかった。

質問紙 C は L 群においては、相関が見られる英語力の側面はなかったが、対象者全体では①GTEC ( $r = .27, p < .05$ )、②素点 ( $r = .32, p < .01$ )、④筆記 ( $r = .27, p < .05$ ) と相関が見られた。H 群では②素点 ( $r = .44, p < .01$ ) と④筆記 ( $r = .51, p < .001$ ) と相関が見られ、いずれも質問紙 A よりも相関係数が高かった。

表9

英語力関連4観点と各質問紙の相関係数(群別)

	①GTEC			②素点			③焦点			④筆記		
	全	H	L	全	H	L	全	H	L	全	H	L
質問紙A	.31**	-.13	.36*	.32**	.39**	.11	.20	.27	.02	.33**	.30*	.32*
質問紙B	.18	.15	.46**	.26*	.28	.29	.21*	.21	.16	.35***	.48**	.25
質問紙C	.27*	.06	.16	.32**	.44**	.15	.19	.12	.17	.27*	.51***	.04

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

質問項目の内容から、質問紙A及びBは習熟度と学業成績の両方と、質問紙Cは学業成績と関連することが予測されたが、本稿の結果をまとめると、各質問紙が反映する側面は習熟度によって異なる可能性が示唆された。具体的には、学習者の英語習熟度に関わらず、最も多くの英語力の側面と関連する自己効力感を測定できるのは質問紙Aであった。そして、H群の成績と関連する自己効力感を測定するには大学生用に作成された質問紙Cが、L群の習熟度には他者との比較を促す質問紙Bがより適切である可能性が示唆された。逆に、H群の習熟度、L群の学業成績と関連する自己効力感の測定に有効な尺度は本稿で検証対象とした3つの中にはなかった。また、学生の多くは自身のインプット能力を根拠に自己効力感を構築している可能性があり、本稿の3つの尺度ではアウトプットの能力に対する自己効力感を測定することは難しいことが示唆された。

社会的比較の影響に関連して、H群において質問紙Aは②素点と関連するのに対して、質問紙Bは関連しないことから、H群に関しては他者との比較を促す文言がむしろ自己効力感の適切な判断の妨げになる恐れがある。逆に、L群に関しては、他者との比較を含む質問紙B( $r = .46, p < .01$ )の方が比較を含まない質問紙A( $r = .36, p < .05$ )よりも習熟度との相関が高いことから、他者との比較が自身の英語力を適切に認識する上で有効な手がかりとなっている可能性がある。あるいは、①GTECの結果が他者との比較に強く影響していると言えるかもしれない。

質問紙Cは、本稿におけるH群において質問紙Aよりも僅かながら②素点と高い相関が見られ、L群においてはどの英語力の側面とも相関が見られなかった。すなわち、習熟度の低い学習者の自己効力感を測定するには適さないが、習熟度の高い学習者の成績に関する自己効力感はより適確に測定する可能性が示唆された。仮に、一般的な大学生の英語力が本稿のH群と同等のCEFR A2以上である<sup>5</sup>とすれば、多くの先行研究で採用されている質問紙Aよりもむしろ質問紙Cの方が僅かながら高い精度で大学生の成績に対する自己効力感を測定する可能性があるだろう。

## 6. まとめ

本稿前半では、日本の英語教育分野において用いられている自己効力感尺度を概観し、松沼(2006)、Pintrich and De Groot(1990)、森(2004)が多く採用されているが、1) 英語力の指標、2) 対象者、3) 他者との比較の3観点において検討が必要であることが明らかになった。後半では、これらの要因を考慮し、質問紙A(松沼, 2006)、質問紙B(他者比較)、質問紙C(大学生対象)の3種類の自己効力感尺度と英語力に関する4観点(習熟度、当該科目の素点、当該科目で焦点とした技能の学期末試験、文法・語彙の知識を中心とした学期末筆記試験)の相関を習熟度別に検証した。その結果、学生の多くは自身のインプットの能力を根拠として自己効力感を構築しており、本稿で採用した3つの尺度ではアウトプットに対する自己効力感は測れない可能性があることが

わかった。さらに、測定したい英語力の側面や習熟度によって、使用が適した尺度が異なることも明らかになった。具体的には、質問紙 A は習熟度や英語力の側面にかかわらず、万遍なく自己効力感を測定でき、質問紙 B は習熟度の低い学習者の習熟度を、質問紙 C は習熟度の高い学習者の成績を反映する傾向が見られた。しかしながら、一般的な大学生の英語力が本稿の H 群と同等の A2 レベルであると仮定すれば、大学生の成績と関連する自己効力感尺度としては、そもそも大学生を対象に作成された、質問紙 C がより適している可能性が示唆された。社会的比較の影響に関しては、習熟度の低い学習者にとっては自己効力感の適切な測定及び向上の両面から有益である一方で、習熟度の高い学習者にとっては不要、場合によっては自己効力感の適切な測定の妨げになる恐れがあることがわかった。

教育的示唆としては、L 群は他者との比較があった方が自己効力感を正確に把握できるだけでなく、自己効力感が高く示されることが明らかになった。Multon et al. (1991) が自己効力感と学業成績の相関は学力が高い学習者よりも低い学習者の方が強いことを示していたことから、習熟度の低い学習者ほど情意面の影響を受けやすい可能性がある。そのため、L 群に関しては、社会的比較が自己効力感、ひいては英語力の向上に影響する可能性を考慮し、グループワークやペアワークの組み合わせに配慮する必要があるだろう。また、質問紙 A ～ C の回答の平均値を比較した結果から、クラスメイトとの比較を促されない限り、配置されたクラスのレベルに基づいて自身の英語力を判断する傾向が見られた。この点も鑑みると、自身が配置されたクラスの優劣を知ることが自己効力感、その後の英語学習に対する姿勢も左右しうするため、レベル開示には慎重になる必要があるかもしれない。

研究における今後の課題としては、本稿及び松沼 (2006) において検証した、質問紙 A と学期末筆記試験の相関に関して、本稿でも有意ではあったものの ( $r = .33, p < .01$ ), 松沼 (2006,  $r = .69, p < .01$ ) ほど高くはなかった。この差が、大学生と高校生の違いに起因するものか、習熟度の違いに起因するのかはさらなる検証が必要である。同様に、大学生を対象とする尺度として質問紙 C が質問紙 A よりも適切かを判断するためには、協力者数を増やし、幅広い習熟度の学習者を対象とした検証を要する。

質問紙 B は社会的比較の影響を検証するために、全ての質問項目に「クラスメイトと比較して」という文言を付与した。しかしながら、Pintrich and De Groot (1990) においては、4 つの質問項目にしか当該の文言は付与されていなかった。Pintrich and De Groot (1990) に忠実にこれら 4 つの質問項目のみに他者との比較を促す文言を付与した場合、異なる結果が得られたかもしれない。また、英語教育分野に限定しない社会的比較に関する先行研究においては、大学生においても社会的比較は自己効力感に影響するという結果を示していた。しかし、本稿の H 群においては社会的比較の影響は見られず、仮に一般的な大学生が本稿の H 群と同等の CEFR A2 レベルであるとすれば、社会的比較と自己効力感の関連は日本人大学生の英語学習場面では当てはまらない可能性がある。徳岡・前田 (2011) の知見を勘案すると、英語力が高い学習者は社会的比較だけでなく継続的比較も行うために、社会的比較の影響が希薄化することを意味するのかもしれない。あるいは、先行研究では、授業内でタスクに従事させる過程で他者の到達具合を明示的に提示していたのに対し、本稿では他者との比較を促す指示を与えたのみであったことが影響した可能性もある。授業における観察を根拠とした主観的な印象に基づく情報のみでは、客観的な情報を与えられたときと同等の効果を得られないのかもしれない。さらに、本稿では英語習熟度に基づいて 2 群に分けて比較を行ったが、H 群は全て 2 年生、L 群は 1 人を除いて全て 1 年生であった。そのため、

H/L 群間の違いは学年による違いであった可能性もある。Ruble et al. (1980) は年齢と共に社会的比較の影響は強くなると報告していたが、一定の年齢を過ぎると影響を受けづらくなる傾向が生じる可能性も検討する必要があるだろう。本調査では、日本の大学生の英語教育分野における自己効力感尺度に対する他者比較の影響に関して、さらに検証を要する課題が複数残された。

本稿の結果だけでは断定的なことを述べるには十分とは言えないが、指導の自己効力感に対する効果を測定する際は、学習者の特性、検証したい英語力の側面によって尺度を使い分けるといった配慮が必要になるかもしれない。例えば、CEFR A2 以上の学生を対象として、英語成績への効果を検証する際は、一般的に使用されてきた質問紙 A よりも本稿で作成した質問紙 C の方が高い精度の結果が得られる可能性も視野に置く必要があるだろう。今後は本稿において課題となった観点をさらに検証し、具体的な尺度選択の基準を提案したい。

## 注

- <sup>1</sup> J-Stage の収録記事約 580 万件 (2025 年 5 月現在) に対して、CiNii Research は 2023 年 4 月時点で約 1 億件以上であり、その後も増加中である旨記載があるため (オープンアクセスリポジトリ推進協会, 2023), 収録記事数が多い CiNii Research において検索を行った。
- <sup>2</sup> 自己効力感の提唱者である Bandura (1995) は、自己効力感には、active mastery experience (自分自身の成功体験), vicarious experience (他者の成功の観察), verbal persuasion (周囲からの口頭の支援), physiological and affective states (身体的, 感情的状態) の 4 つの源があるとしている。
- <sup>3</sup> Pintrich et al. (1991) の MSLQ を日本語訳し、信頼性・妥当性の検証を行った研究に宮部他 (2016) 及び Nomura et al. (2023) があるが、いずれも英語以外を対象としていた。具体的には、前者は看護学生の学業成績、後者は、医学部生の Problem-Based Learning 授業をそれぞれ検証対象としていた。英語教育の分野では、Onoda (2017) において英文読解、自己調整方略、自己効力感の関連を検証するために MSLQ を用いたとの記載があるが、使用した訳は公開されていない。
- <sup>4</sup> 本調査で示された、例えば  $r = .21, p < .05$  は一般的な相関係数としてはそれほど高くはないかもしれない。しかし、心理尺度作成を目的とした太幡 (2020) において、外部基準との相関が  $r = -.15 \sim -.52, p < .05$  であったことで妥当性を確認できたと結論付けていることを勘案すれば、本稿で見られた相関係数は各尺度が各英語指標と関連すると判断する上で十分であると言えるだろう。
- <sup>5</sup> 2023 年時点で高校生の 50.6% が CEFR A2 レベル、19.8% が B1 レベルに達していた (文部科学省, 2024) ことに鑑みると大学生の約 7 割は A2 レベル以上の英語力を有していると考えられる。

## 引用文献

伊藤崇達 (2009). 『自己調整学習の成立過程：学習方略と動機づけの役割』北大路書房。  
オープンアクセスリポジトリ推進協会. (2023.9.25). 「CiNii Research の現況と展望について」  
<https://jpcoar.repo.nii.ac.jp/records/2000300>

- 合田美子・奥田雅信.(2009). 「自己調整学習サイクルにおける目標設定と自己効力感」 『リメディアル教育研究』 第4巻1号, 80–87. [https://doi.org/10.18950/jade.4.1\\_80](https://doi.org/10.18950/jade.4.1_80)
- 太幡直也.(2020). 「嘘をつくことに対する認識尺度の作成」 『心理学研究』 第91巻, 34–43. <https://doi.org/10.4992/jjpsy.91.18237>
- 徳岡大・前田健一.(2011). 「大学生の英語学習と社会的比較・継時的比較」 『広島大学心理学研究』 第11号, 107–116. <https://doi.org/10.15027/32388>
- 中垣州代・文鐘聲・小山内秀和.(2024). 「小学校外国語におけるICTを用いた市内2校交流実践—自己効力感尺度の分析から—」 『畿央大学紀要』 第21巻, 59–65. <https://doi.org/10.24482/0002000102>
- 中山誠一・松沼光泰.(2013). 「再帰属訓練法は英語学習に対する自己効力感を向上させるか」 『城西大学語学教育センター研究年報』 第7巻, 23–32. [https://doi/10.20566/18801919\\_7\\_23](https://doi/10.20566/18801919_7_23)
- 新本省悟.(2020). 「リメディアル教育が必要な大学生に対するディクテーション指導と自己効力感に関する研究」 『リメディアル教育研究』 第14巻, 73–77. <https://doi.org/10.18950/jade.2019.12.28.01>
- 濱田陽.(2011). 「取り組みやすいシャドーイング法の開発」 『リメディアル教育研究』 第6巻, 71–78. [https://doi.org/10.18950/jade.6.1\\_71](https://doi.org/10.18950/jade.6.1_71)
- 藤田恵里子.(2024). 「英語学習者向け辞書検索行為に対する自己効力感尺度作成の試み」 『JACET 関東支部紀要』 第11号, 5–25. [https://doi.org/10.57365/jacetkanto.11.0\\_5](https://doi.org/10.57365/jacetkanto.11.0_5)
- 牧野眞貴.(2013). 「英語が苦手な大学生の自己効力感を高める授業づくり」 『リメディアル教育研究』 第8巻1号, 172–180. [https://doi.org/10.18950/jade.8.1\\_172](https://doi.org/10.18950/jade.8.1_172)
- 牧野眞貴.(2016). 「英語リメディアル授業におけるスピーキング指導と自己効力感の関係についての一考察」 『関西英語教育学会紀要』 第39巻, 1–15. [https://doi.org/10.18989/selt.39.0\\_1](https://doi.org/10.18989/selt.39.0_1)
- 牧野眞貴.(2021). 「リメディアル学生の英作文に対する自己効力感を高める取り組み」 『関西英語教育学会紀要』 第44巻, 81–90. [https://doi.org/10.18989/selt.44.0\\_81](https://doi.org/10.18989/selt.44.0_81)
- 松沼光泰.(2004). 「テスト不安, 自己効力感, 自己調整学習及びテストパフォーマンスの関連性—小学校4年生と算数のテストを対象として—」 『教育心理学研究』 第52巻4号, 426–436. [https://doi.org/10.5926/jjep1953.52.4\\_426](https://doi.org/10.5926/jjep1953.52.4_426)
- 松沼光泰.(2006). 「英語自己効力感 (ESE) 尺度の作成」 『早稲田大学大学院教育学研究科紀要』 第14巻1号, 89–97.
- 水本篤.(2011). 「自己調整語彙学習における自己効力感の影響」 『関西大学外国語学部紀要』 第5号, 35–56. <https://kansai-u.repo.nii.ac.jp/records/10174>
- 宮部明美・富樫千秋・佐久間夕美子・佐藤千史.(2016). 「日本語版MSLQ (Motivation Scales) の信頼性と妥当性の検討」 『日健医誌』 第25巻 (S), 276–286. [https://doi.org/10.20685/kenkouigaku.25.supplement\\_276](https://doi.org/10.20685/kenkouigaku.25.supplement_276)
- 森陽子.(2004). 「大学生の自己効力感と英語学習方略の関係」 『日本教育工学会論文誌』 第28号, 45–48. <https://doi.org/10.15077/jjet.KJ00003730663>
- 文部科学省.(2024). 「令和5年度『英語教育実施状況調査』概要」 [https://www.mext.go.jp/content/20240527-mxt\\_kyoiku01-000035833\\_1.pdf](https://www.mext.go.jp/content/20240527-mxt_kyoiku01-000035833_1.pdf)

- 渡寛法・中島宏治. (2020). 「リメディアル大学生の英語学習態度とパフォーマンスの関係—自己効力感・内発的価値・自己調整と TOEIC リスニングスコア—」 『リメディアル教育研究』 第 14 卷, 51–59. <https://doi.org/10.18950/jade.2020.04.30.01>
- Albert, S. (1977). Temporal comparison theory. *Psychological Review*, 84(6), 485–503. <https://doi.org/10.1037/0033-295X.84.6.485>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28(2), 117–148. [https://doi.org/10.1207/s15326985ep2802\\_3](https://doi.org/10.1207/s15326985ep2802_3)
- Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). Cambridge University Press. <https://doi.org/10.1017/CBO9780511527692>
- Bandura, A., & Jourden, F. J. (1991). Self-regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology*, 60(6), 941–951. <https://doi.org/10.1037/0022-3514.60.6.941>
- Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *The Journal of Social Psychology*, 130(3), 353–363. <https://doi.org/10.1080/00224545.1990.9924591>
- Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior*, 23(3), 329–345. [https://doi.org/10.1016/0001-8791\(83\)90046-5](https://doi.org/10.1016/0001-8791(83)90046-5)
- Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology*, 81(1), 78–96. <https://doi.org/10.1348/2044-8279.002004>
- Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 119(1), 51–69. <https://doi.org/10.1037/0033-2909.119.1.51>
- Duncan, T. G., & McKeachie, W. J. (2005). The making of the motivated strategies for learning questionnaire. *Educational Psychologist*, 40(2), 117–128. [https://doi.org/10.1207/s15326985ep4002\\_6](https://doi.org/10.1207/s15326985ep4002_6)
- Erkan, D. Y., & Saban, A. I. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in Turkish tertiary-level EFL. *The Asian EFL Journal Quarterly*, 13(1), 164–192. <https://shorturl.at/REv7v>
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140. <https://doi.org/10.1177/001872675400700202>
- Fitri, E., D. R., Sofyan, D., & Jayanti, F. G. (2019). The correlation between reading self-efficacy and reading comprehension. *Journal of English Education and Teaching*, 3(1), 1–13. <https://doi.org/10.33369/jcet.3.1.1-13>
- Fukihara, A. (2019). Self-efficacy change of university slow-learners of English. *The Review of Osaka University of Commerce*, 15(1), 517–540. [https://ouc.repo.nii.ac.jp/record/822/files/fukihara.191-192\\_33.pdf](https://ouc.repo.nii.ac.jp/record/822/files/fukihara.191-192_33.pdf)

- Gan, Z., & Yan, Z., & An, Z. (2022). Development and validation of an EFL speaking self-efficacy scale in the self-regulated learning context. *The Journal of Asia TEFL*, 19, 35–49. <https://doi.org/10.18823/asiatefl.2022.19.1.3.35>
- Hackett, G. (1985). Role of mathematics self-efficacy in the choice of math-related majors of college women and men: A path analysis. *Journal of Counseling Psychology*, 32(1), 47–56. <https://doi.org/10.1037/0022-0167.32.1.47>
- Inagaki, Y. (2022). The relationships between self-confidence, self-efficacy, motivation, and social and temporal comparisons among foreign language learners. *JACET Journal*, 66, 39–56. [https://doi.org/10.32234/jacetjournal.66.0\\_39](https://doi.org/10.32234/jacetjournal.66.0_39)
- Kutuk, G., Putwain, D. W., Kaye, L. K., & Garrett, B. (2022). The development and preliminary validation of a new measure of self-efficacy: Questionnaire of self-efficacy in learning a foreign language. *International Journal of Applied Linguistics*, Advance online publication. <https://doi.org/10.1075/itl.21031.kut>
- Lent, R. W., & Hackett, G. (1987). Career self-efficacy: Empirical status and future directions. *Journal of Vocational Behavior*, 30(3), 347–382. [https://doi.org/10.1016/0001-8791\(87\)90010-8](https://doi.org/10.1016/0001-8791(87)90010-8)
- Mendoza-Torres, M. E., Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., Rivera-Lozada, O., Sucapuca-Sucapuca, F. E., & Cunza-Aranzábal, D. F. (2023). Psychometric properties and invariance of an English self-efficacy scale for university students in Peru. *Frontiers in Psychology*, 14, 1–13. <https://doi.10.3389/fpsyg.2023.1187342>
- Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety and their relation to reading and listening proficiency. *Foreign Language Annals*, 39(2), 273–292. <https://doi.org/10.1111/j.1944-9720.2006.tb02266.x>
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38(1), 30–38. <https://doi.org/10.1037/0022-0167.38.1.30>
- Nomura, O., Soma, Y., Kijima, H., & Matsuyama, Y. (2023). Adapting the motivated strategies for learning questionnaire to the Japanese problem-based learning context: A validation study. *Children*, 10(1), 154. <https://doi.org/10.3390/children10010154>
- Ocak, G., & Olur, B. (2018). The scale development study on foreign language speaking self-efficacy perception. *European Journal of Foreign Language Teaching*, 3(1), 50–60. <https://doi.org/10.5281/zenodo.1170932>
- Onoda, S. (2017). Self-efficacy, metacognitive self-regulation strategy use, and L2 English reading skills. *Media, English and Communication*, 7, 51–66. [https://doi.org/10.11293/james.7.1\\_51](https://doi.org/10.11293/james.7.1_51)
- Öztürk, G., & Saydam, D. (2014). Anxiety and self-efficacy in foreign language writing: The case in Turkey. *Başkent University Journal of Education*, 1(2), 10–21. <http://buje.baskent.edu.tr/index.php/buje/article/view/32>
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr, & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1–49). JAI Press. <https://doi.org/10.1108/S0749-742320140000018015>

- Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology*, 86(2), 193–203. <https://doi.org/10.1037/0022-0663.86.2.193>
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40. <https://doi.org/10.1037/0022-0663.82.1.33>
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Technical Report No.91-8-004). National Center for Research to Improve Postsecondary Teaching and Learning. <https://files.eric.ed.gov/fulltext/ED338122.pdf>
- Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-ROYAL*, 3(1), 14–28. <https://www.ajindex.com/dosyalar/makale/acarindex-1423909865.pdf>
- Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebel, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology*, 16(2), 105–115. <https://doi.org/10.1037/0012-1649.16.2.105>
- Sağlam, D., & Arslan, A. (2018). The development of English language skills self-efficacy scale for higher education students. *International Journal of Psycho-Educational Sciences*, 7(2), 1–15. <https://files.eric.ed.gov/fulltext/EJ1252951.pdf>
- Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology*, 8(1), 76–86. [https://doi.org/10.1016/0361-476X\(83\)90036-X](https://doi.org/10.1016/0361-476X(83)90036-X)
- Schunk, D. H. (1985). Self-efficacy and classroom learning. *Psychology in the Schools*, 22(2), 208–223. [https://doi.org/10.1002/1520-6807\(198504\)22:2<208::AID-PITS2310220215>3.0.CO;2-7](https://doi.org/10.1002/1520-6807(198504)22:2<208::AID-PITS2310220215>3.0.CO;2-7)
- Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, 77(3), 313–322. <https://doi.org/10.1037/0022-0663.77.3.313>
- Siegel, R. G., Galassi, J. P., & Ware, W. B. (1985). A comparison of two models for predicting mathematics performance: Social learning versus math aptitude–anxiety. *Journal of Counseling Psychology*, 32(4), 531–538. <https://doi.org/10.1037/0022-0167.32.4.531>
- Wang, C., Kim, D., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System*, 44, 24–33. <https://doi.org/10.1016/j.system.2014.01.015>
- Wang, Y., & Sun, P. P. (2024). Development and validation of scales for speaking self-efficacy: Constructs, sources, and relations. *PLOS ONE*, 19(1), Article e0297517. <https://doi.org/10.1371/journal.pone.0297517>
- Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System*, 39, 510–522. <https://doi.org/10.1016/j.system.2011.10.017>
- Zhang, J., Zhang, L. J., & Zhu, Y. (2023). Development and validation of a genre-based second language (L2) writing self-efficacy scale. *Frontiers in Psychology*, 14, 1–13. <https://doi.org/10.3389/fpsyg.2023.1181196>

- Zhang, X., & Ardasheva, Y. (2019). Sources of college EFL learners' self-efficacy in the English public speaking domain. *English for Specific Purposes*, 53, 47–59. <https://doi.org/10.1016/j.esp.2018.09.004>
- Zhang, X., Ardasheva, Y., & Austin, B. W. (2020). Self-efficacy and English public speaking performance: A mixed method approach. *English for Specific Purposes*, 59, 1–16. <https://doi.org/10.1016/j.esp.2020.02.001>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41(2), 64–70. [https://doi.org/10.1207/s15430421tip4102\\_2](https://doi.org/10.1207/s15430421tip4102_2)
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82(1), 51–59. <https://doi.org/10.1037/0022-0663.82.1.51>

## 資料

### 質問紙 A : 松沼 (2006)

- 1) 私は英語が得意だと思う。
- 2) 私は英語の授業で教えられたことを理解することができると思う。
- 3) 私は英語で良い成績をとることができると思う。
- 4) 私は英語の授業で与えられた課題に適切に答えることができると思う。
- 5) 私の英語の学力はすぐれていると思う。
- 6) 私は英語の学習内容についてたくさんを知っていると思う。
- 7) 私は英語の学習内容を習得できると思う。
- 8) 私は英語の勉強方法を知っていると思う。

### 質問紙 B : 他者との比較

質問紙 A の各項目の冒頭に「クラスメイトと比較して、」を付記

### 質問紙 C : Pintrich et al. (1991) の日本語訳

- 1) 私は英語の課題及びテストをうまくこなすことができる自信がある。
- 2) 私は英語において素晴らしい成績を収めることができると信じている。
- 3) 私は英語のクラスでうまくやれると信じている。
- 4) 私は英語の教員が提示する最も難解な教材も理解できる自信がある。
- 5) 私は英語のクラスで習う基本的な概念を身につける自信がある。
- 6) 私は英語のクラスで習うスキルを自分のものにすることができる自信がある。
- 7) 英語の難易度, 教員, 自分のスキルを鑑みて, 本科目でうまくやれると思う。
- 8) 私は英語のクラスで提示される最も難解な読み物を理解できる自信がある。

## 文章作成支援ツールの使用による英作文の書き直しに対する 日本人英語学習者の認識

### Japanese English Learners' Perceptions of English Writing Revisions Using Writing Assistance Tools

甲斐 順  
(帝京大学)

#### Abstract

This study examines Japanese university students' perceptions of written corrective feedback (WCF) and the revisions of English essays when using writing assistance tools such as DeepL Write or Google Translator. It also explores which tools students employ during the revision process. Among thirty-two sophomores in a general English course, twenty-seven responded to a questionnaire about their essay writing practices. The results indicate that students prefer writing essays on their own, without relying on tools, recognize the benefits of revising with assistance tools, and generally enjoy writing activities. Most respondents reported improvement in their writing skills through revisions with such tools, and many learners showed a tendency to depend on a particular tool throughout the revision process.

**Keywords:** 書記訂正フィードバック, 書き直し, ルーブリック, 文章作成支援ツール, アンケート

#### 1. はじめに

第二言語ライティング指導における書記訂正フィードバック(WCF)は, その是非をめぐって1990年代の論争 (Ferris, 1999; Truscott, 1996, 1999) を経て, 2000年代の最初の10年間にはWCFの有効性を示す研究が見られるようになる (Bitchener, 2008; Bitchener et al., 2005; Ellis et al. 2008; Sheen, 2007)。近年ではWCFにより学習者のライティングの正確性の向上につながることやWCFの効果が持続することがわかってきている (Hyland et al., 2018; 鈴木, 2024)。WCFに関する研究は, 直接的WCFや間接的WCFといった種類だけでなく, WCFを受けた後の行動の効果に関する議論がなされるようになってきている (齋藤・鈴木, 2022)。WCFを受けた後, 学習者に英作文を書き直させることで, 新たな作文の正確性の向上に効果があることを示唆する研究も見られる (Shintani et al., 2014)。近年ではテクノロジーの進歩に伴い, 英作文の添削に活用する研究事例も見られるようになり, Google 翻訳やDeepL Writeといった文章作成支援ツールを用いた英作文指導及びフィードバックに関する研究成果も報告されつつある (Barrot, 2023; 甲斐, 2025; 新美・梅木, 2024)。ただし, 筆者の知る限り, 教員からのWCFを受けて, これらの文章作成支援ツールを使用し, 学習者に英作文の書き直しを行った研究は見られない。本研究では, 教員からのWCFを受けて, 文章作成支援ツールを使用し, 英作文の書き直しを行った日本人英語学習者の認識を明らかにすることを目的として行い, 調査の結果, どのような特徴が見られるのかを報告する。

## 2. 先行研究

### 2.1 WCF の書き直しに関する研究

近年、日本人英語学習者を対象とした WCF の書き直しに関する研究が報告されてきている。

青山 (2019) は、複数回与える WCF が、同じ英作文の書き直しにおける自己訂正率と別の新しい英作文における正確さにどのような影響を与えるのかを文法項目 (仮定法と直説法) に焦点を当てて高校生を対象に研究を行っている。熟達度に応じて熟達度高位群と低位群に分け、それぞれの群で直接的訂正フィードバック (DCF) 群、メタ言語的訂正フィードバック (MCF) 群、統制群を設定し、相対的効果を検証した。pretest, post-test1, post-test 2 それぞれにおける書き直しの機会を設け、合計 3 回の書き直しを実施した。なお、pretest となる英作文課題終了後、DCF 群には初めから学習者に正しい形式を与えるプリントが与えられ、MCF 群には文法説明と使用例と学習者自らが正しい形式を考えるフィードバックシートが与えられた。結果として、熟達度高位群では、DCF と MCF 共に書き直しへの効果が確認されたが、新しい英作文には確認されなかった。一方、熟達度低位群では、DCF は書き直しに対してすぐに効果を及ぼしたが、新しい英作文における正確さの向上はもたらさなかった。しかし MCF は、与えるごとに書き直しにおける自己訂正率を向上させ、3 度目には DCF と同程度の自己訂正率を獲得した。さらに新しい英作文では、2 度目を与えた後の事後テストにおいて、DCF と比較し、正確さの向上を導いた。青山の研究では、MCF 群にフィードバックシートを配付し、学習者が誤っている箇所を参照できるようなチェックポイントが設けられており、目標文法事項の事前説明の単なる反復にならないように工夫されているという点で注目に値する。ただし、1 文レベルの和文英訳にとどまっており、まとまりのある文章を書かせた場合にはどのような MCF を行えばよいのかといった点にまでは踏み込んでいない。また学習者が書き直しを通じて、どのように WCF を感じているかまでは明らかにされていない。

齋藤・鈴木 (2022) は、WCF 後の振り返りや書き直しを含む行動を説明する WCF 認知モデルを提案し、(1) 高校時代と大学時代における WCF 後の行動としての振り返りや書き直しの経験、(2) WCF の種類 (直接訂正と間接訂正) に対する態度としての好みと有効性の認知、(3) WCF エピソードに基づいた思考の程度に関して、大学生を対象にアンケートを実施した。その結果、振り返りが大学時代よりも高校時代のほうがなされる傾向にあること、間接訂正よりも直接訂正を好むものの、必ずしも間接訂正の有効性を認知していないわけではないということ、書き直しがないと WCF について考える程度がそうでない場合と比べて相対的に低減することが明らかとなった。齋藤・鈴木の研究は、WCF 後の振り返りや書き直しを検討することの価値を示しているが、実際に学生に英作文を書かせたり、書き直しをさせたりといった調査までは行っていない。

坂本・古家 (2023) は、高校生が書いた自由記述式英作文をルーブリックで評価・フィードバックすることによって学習改善の効果があるかどうかを実証的に検証している。論題に対して、第 1 次ドラフティングを書かせ、ルーブリックに従って採点し、ルーブリックとともに返却後、同じ論題で書き直しをさせた。その結果、ルーブリックを用いたフィードバックは、生徒の学習改善を促すのにある程度有効な手段となる可能性があることが示唆された。ただし、教員に対するアンケートの結果、調査に使用したルーブリックが使いにくいという課題もあることが判明した。坂本・古家の研究では英作文において学習者が書き直しを行う上で、ルーブリックは一定程度効

果があることを示している。なお、この研究では、1つの論題に対して、1次ドラフティングと2次ドラフティングしか書かせていないので、複数の論題に対して、同じ手続きで複数回英作文を学習者に書かせた場合、同様の結果が導き出されるかはわからない。

## 2.2 文章作成支援ツールを使用した WCF に関する研究

より近年になり、Grammarly や DeepL Write といった文章作成支援ツールが普及してきたこともあり、ライティング活動において、文章作成支援ツールを使用した指導、書き直しに関する研究も見られる。本項で触れる先行研究では、ChatGPT、DeepL Write、Google 翻訳、Grammarly、Transable が扱われているので、ここで簡単に各ツールについて触れておきたい。ChatGPT は米国 OpenAI 社によって開発された生成 AI で、大規模言語モデルが用いられ、膨大なテキストデータから単語の意味や文脈を事前に学習することで、入力された情報に対して確率的に次の単語や文章を予測・生成する仕組みを持つ、ブラウザ上で直接利用できるアプリケーションである (水本, 2024)。翻訳、要約、パラフレーズ、文章生成、コーディングやアノテーションのサポートを得意とする (水本, 2024)。DeepL Write は、AI を活用したライティングアシスタントで、文章力を向上させることが可能で、文法、スペル、句読点の修正に加え、代替語や表現の提案を通じて、文章作成をサポートする (DeepL, n.d.)。Google 翻訳は利用するアプリを問わず、テキストをコピーしてタップするだけで、文字入力、音声入力、手書き入力での翻訳が可能ツールである (Google, n.d.)。Grammarly は、作成された文章について、文法や表記、文章の一貫性や明瞭さ等を添削し、誤りや不適切な表現に赤や青の下線が引かれ、修正案が提示される (新美・梅木, 2024)。Transable は杉山 滉平氏が開発したもので、後述するように複数の AI ツールが組み込まれており、利用者がそれらの支援を受けながら英文を書き進めることができる (山下ら, 2024)。

まず、新美・梅木 (2024) の研究から見ていく。新美・梅木は、「学習者の英語ライティングの語彙や文法的側面の改善を促すための方策として、Grammarly を援用した修正活動をカリキュラムに組み込み、実践を通じた学習者の英語ライティングの変容について理解を深めること」(p. 71) を目的として取り組んだ実践を報告している。大学生を対象に2つの英作文のトピックに関して、ピアフィードバック、Grammarly を使用させた訂正活動を経ることで、学生の英語ライティングは、ピアフィードバック活動後には内容や構成面が向上し、Grammarly を用いた修正活動後には、文法や表記法が大幅に向上したことが判明した。新美・梅木の実践では、Grammarly の使用・修正方法についての教員からの説明はあるが、教員から学生への英作文に関しての WCF は特に行われていない。また、新美・梅木が認めているように、学生が Grammarly の使用を通じて何を感じ、何を学んだのかについての質的調査については行われていない。

山下ら (2024) は、Grammarly や DeepL 翻訳、ChatGPT といった複数の AI ツールを搭載した英文作成支援ツール Transable を大学生に使用させた研究を報告している。3回の授業の一部を使い、次のような手順で行った。1) 授業内の 20 分で学生に英語の質問文に対する意見をマイクロソフト社のワードファイルに何も参照させずに書かせた。2) その後、Transable を起動させ、書いた英文を入力し、GTEC と CEFR 基準で評価させた。3) 授業の宿題として、授業内に書いた英文を Transable を使用させて適宜修正、「表現を調べる」箇所 ChatGPT に質問するなどして加筆させた。4) 修正後の英文を再度 GTEC と CEFR で評価させ、初稿、修正稿それぞれの英文と評価をワークシートに記録させた。また、Transable の機能を用いた修正稿に修正部分を赤で記させ、修正

理由や使用感を毎回記録させた。調査の結果、自動採点・評価の点で課題はあるが、自動添削の点では有用性が確認され、学生の満足度も高かった。自力で学生がまず英文を書いて Transable の支援を受けて書き直しをさせるという点で、主体的な学習であると言える。ただし、一連のライティング活動では、教師からの WCF は特に行われていない。

書き直しをさせた研究ではないが、甲斐 (2025) は学習者が英作文を行う際に、DeepL Write を活用することを勧め、アンケートを通じて学習者に感想を書かせた実践を報告している。使用教科書の章ごとの writing セクションの論題について、大学生に半期間で 6 回英作文を書かせた。初回と最終回を除く 4 回のライティング活動では DeepL Write を活用することを勧め、英作文を書かせた。甲斐は学生の英作文をパソコンに入力し、DeepL Write に添削させ、元の英作文と DeepL Write による訂正版の両方が比較できる形で印刷した。その用紙と学生が提出した原文にコメント及び DeepL Write が未修正の部分で訂正した上で、学生にフィードバックした。AI ツールの利用の有無や授業でのライティング活動に関するアンケートを実施し、回答結果から AI ツールの活用に関する肯定的な意見が多く見られたが、使い方には工夫が必要であることがわかった。また、アンケートの結果から、自力で書くことを重視し AI ツールを敢えて活用せずに英作文を行っている学生も見られた。甲斐の研究は複数回のライティング活動を行い、フィードバックに関しては DeepL Write の訂正候補について学習者に検討させているが、書き直させる段階までは実施していない。また、学習者は書き直しのないライティング活動を肯定的にとらえているが、書き直しをさせた場合、学習者がどのように感じているのかといった質的研究までは行っていない。

先行研究から、文章作成支援ツールは学習者が書き直しを行う上で、文法や表記法の向上や学習者の使用に関する高い満足感や肯定感を得ることを助けるといったことが示唆されている。先行研究では文章作成支援ツールを用いた英作文活動が見られるが、教員からの WCF を受けた後で、学習者が文章作成支援ツールを用いて書き直しを行っているという研究は報告されていない。また、学生が文章作成支援ツールを使用した際にどのような感想を持ち、ツールに対する好みがあるかといった点に関する調査は行われていない。

### 3. 研究の目的

本研究では、第二言語学習者に複数の論題で英作文を書かせ、教員からのフィードバックを基に文章作成支援ツールを使って書き直しをさせることで、WCF に関する学習者の認識を調査することを目的として行う。そこで次の 4 つの研究課題を設定した。

- 研究課題 1 学習者は 1 度目に自力で英作文を書き、フィードバックを受けて文章作成支援ツールを使用して書き直しを行い、どのように感じているか。
- 研究課題 2 学習者は書き直しをする際に、どのような文章作成支援ツールを使用しているか。また、特定のツールを使用する傾向が見られるか。
- 研究課題 3 学習者は授業で取り組んだ英作文について楽しいと感じたか。
- 研究課題 4 学習者は授業を通じて書く力がついたと感じたか。

## 4. 研究方法

### 4.1 調査対象者

日本の私立大学で英語を専攻としない文系学部に所属する大学2年生で、筆者が指導する後期の一般英語クラスの32人を調査対象者とした。このクラスの授業は、英語の4技能(聞く、読む、話す、書く)を向上させることを目的としている。半期15回の授業で、1回あたりの授業時間は90分である。この大学の一般英語科目は、習熟度クラスで編成されている。2年次のクラスは、1年次のアチーブメントテストの結果に基づいており、筆者の担当クラスは上位クラスに割り当てられた学生で構成されていた。英検の取得状況は2級が7人、準2級が11人、3級が4人、英検未取得者が10人であった。なお、調査対象者となる学生からは、本研究の目的を説明し、協力への了承を得ている。

### 4.2 手順

一般英語クラスの使用教科書は Lynch and Shitori (2018) *Trend watching 2* である。全20章で各章は、Pre-Reading Vocabulary Task, Reading, Comprehension Questions, Vocabulary Practice, Read and Listen, Writing から構成されている。後期授業では、Chapter 9 から開始し、2週で1つの章を終えるようにし、上記タスクの順で実施した。参考までに、Chapter 9 は、Waiting in Line for Delicious New Food で日本人が新規開店の店に数時間並ぶ様子に関する内容である。この章のWriting セクションの設問は “Would you wait in line for two or three hours to try food at a new restaurant? Explain your answer.” である。

ライティングの活動は、後期授業15回のうち、8回実施した。8回のうち、奇数回は学生が自力で、偶数回は筆者からの語彙・文法などの正確性に焦点を当てたフィードバックを基に書き直しを行うという手順を踏んだ。別の言い方をすると、題目は4種類あり、1種類につき学生は1回自力で英作文を行い、筆者からの間接的WCFの返却に基づき、2回目に書き直しを行い、筆者が直接的WCFを加えて返却するという流れで実施した。

奇数回では、携帯電話や辞書等を参照させずに英作文に取り組みさせた。まず、学生に英作文用紙を配付した。英作文用紙は、各課の質問と「携帯や辞書等を使わずに書いてみよう。」という指示文、英作文作成用の数行に渡る罫線、「構成・内容」、「理由・説明」、「語法・文法」の3つの観点に対して4つの基準を設けたループリック、学籍番号と名前の記入欄から構成されていた(付録)。学生は90分の授業のうち15分程度を自力で英作文に取り組み、教員が英作文を回収した。授業後に、間接的WCFとして英作文の誤りのある箇所の下線を引く(例、Second, it is reasonable for me. Reebok is cheap than NIKE, adidas and so on.) とともに、ループリックの各観点の該当項目に○をつけ、作文の内容に関してコメントを記入した。

偶数回に当たる翌週、学生に英作文を返却した。その際、「前回の原稿を書き直しましょう。DeepL Write 等を使って構いません。使ったツール名は必ず書いてください。」という指示文、英作文作成用の数行に渡る罫線、ループリック、学籍番号・名前の記入欄からなる英作文用紙も併せて配付した。よく書けている英作文をマイクロソフト社のパワーポイントを用いて全体に紹介した後、DeepL Write 等の文章支援ツールを使用させ、返却された英作文の書き直しを行わせた。よく書けている英作文を全体に紹介したのは、学習者の作文の中から模範となる作文を選び、クラスで紹介することがあらゆるレベルの学習者にとって有意義であると思われるからである(鈴

木, 2017)。また, 学習者にとっては紹介されることで, 書く意欲の向上や自信につながるといった指導上の効果も期待して全体に紹介した。書き直しには 90 分の授業のうちの 15 分から 20 分程度の時間を与えた。その後, 英作文を回収し, 授業後に誤っている箇所については直接的 WCF を施す (例, Reebok is cheaper than NIKE, Adidas and so on. →Reebok's sneakers are cheaper than those of other brands such as Nike or Adidas.) とともに, ルーブリックの該当項目に○をつけ, 英作文の内容に関する簡単なコメントを書いて, 次の週に返却した。返却時には, 共通して見られる誤り, および学生の名前を伏せた上でよく書けている作文をパワーポイントを用いて紹介した。よく書けているものを紹介した理由は先述した通りである。

英作文作成用紙内にあるルーブリックについては, 初回の授業でこのような観点で英作文を見ると簡単に触れた。ただし, 「必ず参照するように」といった明確な指示は出さなかった。

### 4.3 データ収集と分析方法

研究課題 1,3,4 については, 15 回目の授業時にライティング活動に関するアンケートを実施した。回答にあたって, 成績評価とは一切関係ないこと, 無記名であること, 回答するかどうかは自由参加であることを伝え, できるだけ協力するように依頼した。

アンケートは 3 つの設問で構成されていた。研究課題 1 を調査するために, 設問 1 では「2 週間にわたり 1 つのトピックの英作文を行いました。まず, 第 1 週の授業で, 自力で英作文を書いてもらいました。第 2 週には返却された教員のコメントや下線部を参照して, DeepL Write 等を用いて書き直しを行い, 再提出してもらいました。このことについて, どう思いましたか。」とたずね, 自由記述で回答させた。研究課題 3 を調べるために, アンケートの設問 2 で「この授業で取り組んだ英作文について楽しかったですか。」, 研究課題 4 については, 設問 3 で「この授業で取り組んだ英作文を通じて英語を書く力はついたと思いますか。」とたずね, それぞれ選択形式の 5 件法 (「とても当てはまる」, 「やや当てはまる」, 「どちらともいえない」, 「あまり当てはまらない」, 「全く当てはまらない」) でたずね, 回答してもらった。

アンケートの自由記述以外の設問 (設問 2,3) については単純集計し, 必要に応じてグラフ化した。自由記述の設問 (設問 1) に対する回答については, 似たような回答が複数含まれている場合, カテゴリーに分類し, 分析を行った。

研究課題 2 については, 8 回のライティング活動のうち, 書き直しを行った偶数回で使用したツール名を書かせた。使用したツール名が書いてある場合は具体名を, 書いていない場合は不明として, 学習者ごとに, マイクロソフト社のエクセルに 4 回分のデータを入力・保存し, 使用状況を分析した。なお, 学習者が欠席した場合は何も入力しなかったが, 出席した回数での使用状況を把握するため分析に活用した。

## 4. 結果

調査対象者 32 人のうち, 無記名のアンケートに回答したのは 27 人であった。アンケートの設問 1 は, 奇数回で自力で英作文を書いた後, 教員からのフィードバックを受けて, DeepL Write 等を使用して書き直しを行ったことについて自由記述で回答するものであった。記述から次の 4 つのカテゴリーに分類することができた。

### (1) 自力による英作文の重要性の認識

記述式アンケート回答者27人のうち、7人がこのカテゴリーに該当する回答をしていた。学生は、最初に自力で英作文を書くことの重要性を認識していることがわかる。自分の力で書くことで、どのように表現するかを考える機会が得られ、文法や語彙の理解が深まると感じている。特に、最初の週に自力で書いた後、教員からのフィードバックを受けることで、自分の弱点を認識し、改善点を明確にすることができたと述べている。以下に3例ほど自由記述を示す。

- ・まずは自力で作文を行い、自身の力を知ることができたことと、知らない単語、熟語、表現を調べて自然とよく使う表現を学ぶことができた。単語を使う際にも作文をすることでスラスラとまではいかないが、以前より速く言葉が出てくるようになった。・何も使わずに英作文に取り組むことは、頭をしっかりと使いながら取り組めると実感できたし、その次にアプリを使って書くとより良い文が書けるようになってると感じられて良かった。・同じテーマで2回書くことによって、自分の実力を理解したり、英単語、文法、同じ日本語の文でも複数の表現の仕方があることを知るきっかけになり、とても役立ちました。第1週はとても書くのが大変でしたが、書く能力を向上させるために、必要な刺激だと思いました。一方で、自力で書くことの負担感を訴える声も見られた。以下に2例ほど示したい。

- ・いきなり英作文を書くのは自分には難しいと感じたので、日本語で作文する時間が少しほしいと思いました。
- ・自力で書くこと自体が難しいのでそこに向けた授業をするべきだと思った。文章での表現で使う文法やコツなど。教科書の文章を読むのは読解であり、読解に時間を使っても書けるようになっているとは思えなかった。また文章を頭の中でイメージできてもある単語がわからないといったことが多いので調べるのが最初はなしというのはきつかった。

### (2) フィードバックと書き直しの学習効果

記述式アンケート回答者27人のうち、9人がこのカテゴリーに該当する回答をしていた。学生は教員からのフィードバックを基に、書き直しを行うことで、自分の誤りを認識するとともに、英語力がついたり、正しい用法が定着していることがわかる。以下に4例ほど自由記述を示す。

- ・第1週は自力、第2週は直すべき点や改善点を書いてもらって返却される方法はとても良いなと思った。しっかり自分はどこがダメなのか理解してアプリ等使って書き直しできるのは自分に英語力が身につくなと感じた。・最初に何も見ずに自力で書いたものを2回目返却されたときに間違えてしまっている部分を認識できたし、それをDeepL等使用して正しい文法で解き直すことができたので、頭の中に残りやすくて良かったと思いました。・間違っている部分を自分で調べることによって、何が違っていたのかを学ぶことができ、書く力を伸ばすことにつながりました。・非常に英作文の力がついたと感じている。良い取り組みだと思った。英作文をする機会は受験生の時だけだと思っていたが、大学でも本格的に学ぶことができたため、良い機会であったと感じた。ただ、自力で英作文をするのは、まだ難しいと感じているため、単語力、文法力などを今後も引き上げていきたいと思った。

一方で次のように、書く力の伸長を認識しているものの、書き直し後にどのように自分の中で定着させるかについて課題を感じ、3回目に自力で英作文を行うことの必要性を挙げているコメントも見られた。

- ・自力で英作を行うことは実力が伸びているような気がした。2週目の DeepL Write で書き直した後、どのように自分に身につけさせるかが難しいと感じた。3回目にもう1度自力で英作を行うのが良いと思った。

#### (3) 文章作成支援ツールの活用と学習効果

記述式アンケート回答者27人のうち、9人がこのカテゴリーに該当する回答をしていた。DeepL Write や Google 翻訳といった文章作成支援ツールを使用することで、学生は自分の書いた文章を見直し、正しい文法や表現を学ぶことができたと感じている。特に、文章作成支援ツールを使うことで、間違いを確認し、次回の作文に活かすことができた点が評価されている。以下に3つの例を示したい。

- ・自分で書くものと、同じテーマでサイトを使って書くスタイルは斬新かつ効果的だとやってみて感じた。自分で書いた自分の文章を添削しているようで覚えやすい？感じがした。・最初に自分の力で答えて後から翻訳を使うことによって間違いを確認することができたので良いなと思いました。・自分だけで英作文の練習はしにくいから勉強になった。今後、DeepL Write を使って英作文の練習を続けていきたい。

中には、DeepL Write 等を書き直しに使用することを肯定的にとらえている一方で、筆者のフィードバックにおけるコメントの文字が判読できないといったコメントも見られた。丁寧に学生が判読できるようにコメントを書く必要があるだろう。こういった例は、1例しか見られなかったが、以下に示しておきたい。

- ・DeepL Write 等を使わないで、自分の力で書き、それを先生が修正し、再度、書くのは英語の力を伸ばすのに必要なことだと思うが、時々先生の文字が読めないことがあった。

#### (4) ルーブリックに対する肯定的評価

記述式アンケート回答者27人のうち、2人がこのカテゴリーに該当する回答をしていた。アンケートではルーブリックに関する設問を設けていなかったが、自由記述からルーブリックを肯定的に見ていることを窺わせるコメントが見られた。具体的には以下の2例が該当する。

- ・自分が考えたけど自力で文章化できなかった所を含めて、どう直すべきなのか、分かったので、とても直しやすく、次の英作文で気をつけるポイントが分かって、良かったです。また、1回目の評価から2回目で評価が上がる所も完全な自力ではありませんが、嬉しく感じるポイントの1つでした。

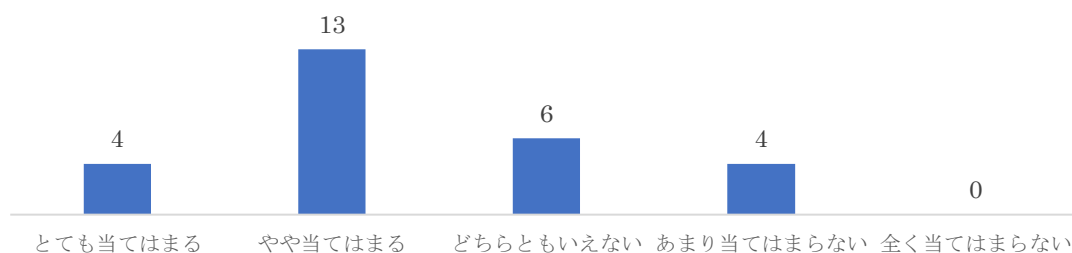
- ・1回書いて終わりではなく、先生の評価や外部システムをもとにもう1度書く機会があることが良かったなと思いました。ここはこうすればもっと良くなるな、など自分で気がつくことができました。

英作文用紙には「ルーブリック」という言葉は用いていなかった(付録)ので、1つ目のコメントに見られる「1回目の評価から2回目で評価が上がる所」、2つ目のコメントの「先生の評価」という箇所は、英作文用紙に記載しておいた表、つまりルーブリックの構成要素のことを指すと思われる。筆者は先述したように学習者が自力で書いた奇数回の英作文については、内容に関するコメントと誤っている箇所に下線を施し、ルーブリックの該当項目を○で囲んで返却するだけであった。ルーブリックが学習者の英作文を書くのに一定程度の効果があることを示唆するコメントである。

自由記述から、「学生の自力による英作文・提出」→「教員による添削とフィードバック」→「ツール活用の書き直し英作文・再提出」→「教員による再添削とフィードバック」といった段階的な英作文指導が効果的であることが示唆された。教員が文章作成支援ツールの適切な活用方法を教えることで、学習効果を最大化できるとともに、語彙力・文法力の強化を継続的に支援する必要性があることも判明した。一方で、教員が最初に書かせる段階で文章作成支援ツールの使用を制限することやフィードバックにおける教員の文字の可読性に課題があることも判明した。

続いて、研究課題3,4に関するアンケート結果について報告したい。アンケートの設問2は授業で取り組んだ英作文について楽しかったかどうかをたずねる設問であった。図1はその結果を表したものである。

図1  
授業で取り組んだ英作文の楽しさ (n = 27)

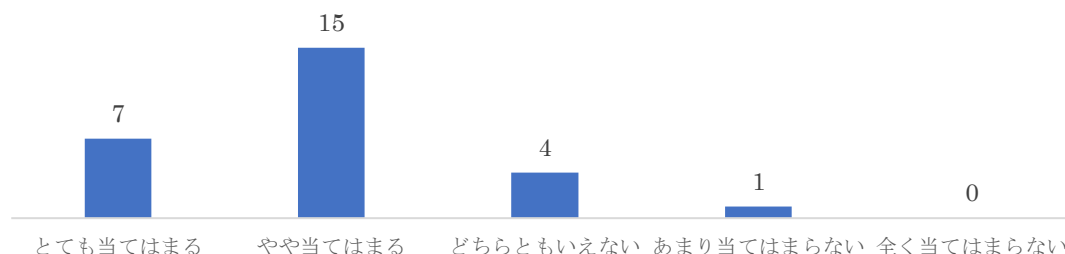


結果は「とても当てはまる」と「やや当てはまる」を合わせて17人で、約60%の学生が英作文を楽しかったと実感していた。一方、「あまり当てはまらない」と「全く当てはまらない」の合計が4人で約15%の学生は英作文を楽しいとはとらえていなかった。残る6人は「どちらともいえない」と回答した。書き直しを行っていない甲斐(2025)でも英作文の楽しさに関するアンケート調査を行っているが、回答者28人中、肯定的評価が約60%で、否定的評価が25%であったことからすると本研究では否定的評価が減っている。

アンケートの設問3は授業で取り組んだ英作文を通じて英語を書く力がついたかどうかに関する設問で、図2はその結果を表したものである。

図2

授業での英作文を通じて英語を書く力がついたという実感 (n = 27)



結果は「とても当てはまる」と「やや当てはまる」を合わせて22人で、約80%の学生が書く力がついたと実感していた。一方、「あまり当てはまらない」と「全く当てはまらない」の合計が1人しかいなかった。残りの4人が「どちらともいえない」と回答していた。書き直しを行っていない甲斐(2025)でも英語を書く力に関するアンケート調査を行い、回答者28人中、肯定的評価が約64%で、否定的評価が約21%であったことからすると本研究ではいずれも改善が見られる。

これら2つの結果から、文章作成支援ツールを使用させての「書き直し」がライティング活動にプラスの効果が見られることが示唆される。

次からは、学生が書き直しをした際に使用していた文章作成支援ツールの使用状況について見ていきたい。表1は、8回のライティング活動のうち、書き直しを行った各回で使用したツール名を書かせた結果をまとめたものである。

表1

書き直しにおける使用ツール名と各回の人数

	書き直し1	書き直し2	書き直し3	書き直し4	
DeepL Write	12	13	12	16	
Google 翻訳	10	10	12	12	
Line 通訳	1	1	1	0	
Google 翻訳+Line 通訳	0	1	0	0	
不明	4	4	1	1	合
計	27	29	26	29	

多くの学生がDeepL Writeまたは、Google 翻訳を使って書き直しを行っている。DeepL Writeを使用している学生数がGoogle 翻訳を使用している学生数より上回る回数が多い(書き直し1, 書き直し2, 書き直し4)が、ほぼ人数は拮抗している。

各学生が書き直しをする際、どのツールを使用し書き直しを行っていたかをまとめたのが表2である。なお、表2のデータにはアンケート回答者以外の学生も含む、各回の課題提出者全員を対象としているため、表1の合計数と表2の合計数が異なっている。

表2

使用していたツールと人数

使用したツール	人数
DeepL Write のみ	12
Google 翻訳のみ	9
DeepL Write → Google 翻訳	2
Google 翻訳 → DeepL Write	1
Line 通訳 → DeepL Write	1
不明 → DeepL Write	2
不明 → Google 翻訳	3
不明 → Google 翻訳+Line 通訳 → Google 翻訳	1
不明	1
合計	32

注. ➡は切り替えたことを表わす。

DeepL Write を最初から最後まで使用し書き直していたのは12人、Google 翻訳を一貫して使用していたのは9人であった。DeepL Write から Google 翻訳に切り替えた学生が2人、逆に Google 翻訳から DeepL Write に切り替えた学生は1人であった。Line 通訳から DeepL Write に切り替えた学生が1人であった。どのツールを使用したのか最初不明だったが、途中から DeepL Write に切り替えた学生が2人、Google 翻訳に切り替えた学生が3人、Google 翻訳及び Line 通訳を経て、Line 通訳、最終的に Google 翻訳へとツールを変えていた学生が1人であった。書き直しの最初から最後まで使用ツール名を記載しない学生が1人(表2の中の「不明」)見られた。

## 6. 考察

本研究では4つの研究課題を設定した。研究課題1「学習者は1度目に自力で英作文を書き、フィードバックを受けて文章作成支援ツールを使用して書き直しを行い、どのように感じているか。」については、アンケートの自由記述から「自力による英作文の重要性の認識」、「フィードバックと書き直しの学習効果」、「文章作成支援ツールの活用と学習効果」、「ループリックに対する肯定的評価」の4つのカテゴリーに分類できた。学生は、①英作文を自力で書いて提出し、②教員からのコメントと誤っている箇所に関する間接的 WCF 及びループリックを参照し、③文章作成支援ツールを活用して書き直し再提出し、④教員からの直接的 WCF を受けるといった段階的な過程を経ることで、自分の語彙や文法などに関する誤りを認識し、英語力がついたと実感していることがわかる。最初のフィードバックの段階で教員からの直接的 WCF を受けない

ことに学生は不満を感じるのではという懸念があったが、むしろ間接的 WCF とルーブリックを基に、文章作成支援ツールを活用して書き直すことを肯定的に受け止めていることが窺える。その一方で、最初から文章作成支援ツールを活用したり、日本語で文章を書いてから英文に直したいという考えを持つ学生もおり、学生の要望に配慮した指導が必要かもしれない。また、フィードバックにおける教員の文字の可読性に課題があることも判明し、より丁寧なフィードバックを心がけることで学習者の英作文に取り組む意欲をそぐことのないように努めることも必要である。

研究課題 2「学習者は書き直しをする際に、どのような文章作成支援ツールを使用しているか。また、特定のツールを使用する傾向が見られるか。」については、DeepL Write と Google 翻訳を利用している学生が多かった。Google 翻訳より DeepL Write を使用している学生の数が多い回も見られたが、人数はほぼ拮抗していた。どちらかのツールを最初から最後まで書き直しをする際に用いている学生も一定程度見られた。その一方で、途中からどちらかのツールに最終的に切り替えている学生も見られた。DeepL Write は前期の授業で活用させていたこともあり、使用していたものと思われる。Google 翻訳は学生が普段から使い慣れ、利便性を感じていることもあり、使用していた可能性がある。今後の展望として、学習者のツールの好みや利便性などを調査することで、書き直しを行わせる際の英作文指導に資するものと思われる。

研究課題 3「学習者は授業で取り組んだ英作文について楽しいと感じたか。」については、約 60%の学生が、自力と文章作成支援ツールを用いた書き直しの英作文活動を楽しかったと実感していた。一方、約 15%の学生は英作文を楽しいとはとらえていなかった。

研究課題 4「学習者は授業を通じて書く力がついたと感じたか。」については、約 80%の学生が書く力がついたと実感していた。否定的な回答は 1 人しかおらず、多くの学生は今回のライティング活動を通じて書く力が伸びたと感じていることがわかる。

学習者が文章作成支援ツールを使用して、英作文の書き直しを行った先行研究では、文法や表記法の向上や学習者の使用に関する高い満足感や肯定感を得ることが示唆されていた。本研究では、文法や表記法の向上といった点について焦点を当てて数値化した実証研究は行っていないので向上したかどうかは言うことができない。ただし、本研究では、学習者自身が、自分の語彙や文法などに関する誤りを認識し、英語力がついたと実感していることが示唆される結果を得ることができたので、これは先行研究を支持する結果と言えるだろう。また先行研究では、教員からの WCF を受けて、学習者が文章支援ツールを用いて書き直しを行っている研究は報告されていなかったが、本研究では、「教員からのフィードバック」と「文章支援ツールによる書き直し」の組み合わせによる有効性を提供している点で、新たな知見を得られたと言えるだろう。

甲斐 (2025) では、学生に書き直しをさせていなかったが、DeepL Write 等の文章作成支援を使用したライティング活動を楽しんでいると感じていたり、書く力がついたと実感していた。本研究では、文章作成支援ツールを使用させて書き直しをさせても、学習者は英作文活動を楽しんでいるとともに、英語を書く力がついたと実感していることがわかる。結果の欄でも紹介したが、「第 1 週は自力、第 2 週は直すべき点や改善点を書いてもらって返却される方法はとても良いなと思った。しっかり自分はどこがだめなのか理解してアプリ等使って書き直しできるのは自分に英語力が身につくなと感じた。」のコメントに見られるように、学習者は自力で書いた英作文を間接的 WCF を受けて、文章作成支援ツールを活用して書き直すことで、書く力を育成することにつながる可能性がある。

## 7. おわりに

本研究では、第二言語学習者に複数の論題で英作文を書かせ、教員からのフィードバックを基に文章作成支援ツールを使って書き直しをさせることで、WCF に関する学習者の認識を調査することを目的として行った。アンケートの結果から、文章作成支援ツールを使用した段階的なWCF が効果的であり、学生は楽しんで英作文活動を行い、書く力がついているという実感を得ていることが示唆された。また、学生が使用していた文章作成支援ツールについては、DeepL Write や Google 翻訳といった特定のものを使用し続ける傾向が見られた。

本研究の限界点をいくつか指摘しておきたい。本研究では、最初に提出させた英作文の誤りの箇所を下線部を施していたので、書き直しの英作文のどこを各学生が訂正したのかは把握できている。しかし、間接的 WCF を受けて、学生自らが誤っている内容を理解し、自力で訂正したのか、それとも文章作成支援ツールを用いた上で、訂正したのかまでは調査することができなかった。学習者の訂正方法、訂正内容を把握することで、文章作成支援ツールをどう使用させたらよいかという指導面に資することになるだろう。また、英作文用紙のルーブリックの該当項目を、学習者がどの程度参照し、どのような印象を持っていたのかまでは検証しきれていない。さらに32人を対象とした小規模な実践のため、今回の結果は一般化はできない。プレテスト・ポストテスト法を用いて、文章作成支援ツールを使用して書き直しを行うことで、実際に学習者の英語力が、どの程度伸びているのかといった検証も必要である。書き直しの回数を増やすことで同様な結果が見られるのかといった点も含めての検証も必要と思われる。今後の研究課題として取り組んでみたい。

## 謝辞

本稿の完成にあたり、貴重なご意見をくださった2名の匿名査読者及び紀要編集委員会に、この場をお借りして厚く御礼申し上げます。

## 引用文献

- 青山聡. (2019). 「筆記による訂正フィードバックの種類と頻度、及び学習者の英語熟達度が単文レベルの英作文の書き直しにおける自己訂正率と新しい英作文の正確さに与える影響」 『日本教科教育学会誌』 42 巻第 2 号, 43–53.
- 甲斐順. (2025). 「DeepL Write を活用した英語ライティング指導と訂正フィードバック」 『JACET 授業学ジャーナル』 第 5 号, 87–106. [https://www.jacetsigdevelopmentaleducation.net/\\_files/ugd/dcd170\\_2792dd77f7e44d45b9004acd70fd8a23.pdf](https://www.jacetsigdevelopmentaleducation.net/_files/ugd/dcd170_2792dd77f7e44d45b9004acd70fd8a23.pdf)
- 齋藤玲・鈴木渉. (2022). 「第二言語ライティングにおける書記訂正フィードバックとその振り返り、書き直しに関する態度と行動: 認知モデルの提案と予備的調査の報告」 『東北英語教育学会研究紀要』 42 巻, 49–64.
- 坂本優子・古家貴雄. (2023). 「高校生を対象とした自由記述式英作文のルーブリックの開発とフィードバック効果の検証」 『教育実践学研究：山梨大学教育学部附属教育実践総合センター研究紀要』 28 巻, 225–236. <https://doi.org/10.34429/00005277>

- 鈴木祐一. (2024). 『あたらしい第二言語習得論—英語指導の思い込みを変える』 研究社.
- 鈴木渉. (2017). 『実践例で学ぶ第二言語習得研究に基づく英語指導』 大修館書店.
- 新美德康・梅木璃子. (2024) 「自動ライティング評価ツール Grammarly を援用した英語ライティング指導の探究的実践」 『リメディアル教育研究』 18 巻, 69–80.  
<https://doi.org/10.18950/jade.2023.07.18.01>
- 水本篤. (2024). 「本編を読む前に知っておきたい 生成 AI 活用のための Q&A」 『英語教育』 73(4), 5–6.
- 山下美朋・山中司・杉山滉平. (2024). 「AI ツールを活かした英語ライティング授業: 英文作成支援ツール Transable を導入して」 『立命館高等教育研究』 24, 75–87.  
<https://doi.org/10.34382/0002000853>
- Barrot, J. S. (2023). Using automated written corrective feedback in the writing classrooms: Effects on L2 writing accuracy. *Computer Assisted Language Learning*, 36(4), 584–607.  
<https://doi.org/10.1080/09588221.2021.1936071>
- Bitchener, J. (2008). Evidence in support of written corrective feedback. *Journal of Second Language Writing*, 17(2), 102–118. <https://doi.org/10.1016/j.jslw.2007.11.004>
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(3), 191–205.  
<https://doi.org/10.1016/j.jslw.2005.08.001>
- DeepL (n.d.). <https://support.deepl.com/hc/ja/articles/6318834492700-DeepL-Write>
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353–371.  
<https://doi.org/10.1016/j.system.2008.02.001>
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1–11. [https://doi.org/10.1016/S1060-3743\(99\)80110-6](https://doi.org/10.1016/S1060-3743(99)80110-6)
- Google (n.d.). <https://translate.google.com/about/?hl=ja>
- Hyland, F., Nicolás-Conesa, F., & Cerezo, L. (2018). Key issues of debate about feedback on writing. In R. M. Manchón, & P. K. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 433–452), De Gruyter Mouton.
- Lynch, J., & Shitori, K. (2018). *Trend watching 2*. Seibido.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283. <https://doi.org/10.1002/j.1545-7249.2007.tb00059.x>
- Shintani, N., Ellis, R., & Suzuki, W. (2014). Effects of written feedback and revision on learners' accuracy in using two English grammatical structures. *Language Learning*, 64(1), 103–131.  
<https://doi.org/10.1111/lang.12029>
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327–369. <https://doi.org/10.1111/j.1467-1770.1996.tb01238.x>
- Truscott, J. (1999). The case for “The case against grammar correction in L2 writing classes”: A response to Ferris. *Journal of Second Language Writing*, 8(2), 111–122. [https://doi.org/10.1016/S1060-3743\(99\)80124-6](https://doi.org/10.1016/S1060-3743(99)80124-6)





# 機械翻訳の使用が日本人大学生の英語ライティング学習に与える影響：

## 脳機能イメージングの観点から

### Exploring the Impact of Machine Translation Use on English Writing Among Japanese University Students: A Neuroimaging Approach

佐竹 幸信

(東京都市大学)

#### Abstract

The rapid evolution of machine translation and generative AI in recent years—driven by big data—has resulted in output that closely approximates natural human language. Against this backdrop, a growing body of literature has advocated for the integration of such technologies into English language education. Nevertheless, practical classroom-based applications remain limited, and few studies have provided concrete evidence to substantiate the pedagogical benefits of machine translation. This study aims to evaluate the effectiveness of machine translation as a tool for learning English writing by examining cerebral hemodynamic responses and learners' actual writing performance. Japanese university students with TOEIC scores in the low 600s were divided into two groups: those who used machine translation and those who did not. The analysis revealed that activating the brain region associated with language production—specifically, Broca's area—may be essential for fostering writing proficiency. Although the use of machine translation was found to have limited influence on the activation of this area, it showed potential utility in facilitating vocabulary acquisition. These results suggest that while machine translation may not fully support the cognitive processes underlying syntactic production, it can serve as a complementary tool promoting language learning.

**Keywords:** ライティング学習, モデル文提示, 教師による添削, 機械翻訳, 脳機能イメージング

#### 1. はじめに

モデル文提示の英語ライティング学習の研究は少なからず存在するが、生成 AI や機械翻訳が生成した英文をモデル文として提示し、その学習効果を測った研究は数少ない。しかし昨今の生成 AI や機械翻訳は、文法等の誤りもほとんど認められず、インターネット上に溢れている英語の慣用的な使用法をベースとしているため、例えば教科書に掲載されているモデル文よりも、ある意味、海外で実際に使用されている「生きた」英語のロール・モデルとなり得るのではないかと考えた。本研究では、従来の教師によるフィードバック型のライティング学習と比べ、このような「AI 支援型モデル文提示」のライティング学習にどのような教育的効果があるかを、実際のライティング・パフォーマンス及び脳血流量の観点から調査した。

## 2. 先行研究

### 2.1 第2言語(外国語)教育における機械翻訳や生成AIの使用について

昨今、機械翻訳や生成AIを第2言語(外国語)学習に応用することを論じる先行研究が増えているが(Son et al., 2025), 筆者の考えによれば, それらは大きく3つのグループに分類される。まず, 第2言語(外国語)学習における生成AIや機械翻訳利用についての筆者の考えを述べるものである。例えば Betal(2023) や Wei(2023), Choukaier(2024) は, AIテクノロジーは, 個別最適化学習, 即時のフィードバックや学習評価, 自律学習, 動機づけ等といった点で第2言語習得にとって効果的であると論じている。Kristiawan et al. (2024) は英語教育へのAIの導入は, 学習結果を促進し, 動機づけを高めるツールや戦略を提供する可能性が十分であると述べ, Almelhes(2023) は今後はチューターとしてのAIが果たす役割が大きくクローズアップされるだろうと予想している。こうした論調を代表する形で, Godwin-Jones(2024) は, AIは“learner-tool-environment”(pp. 7-8)の3者の関係性に大きな変革をもたらし, この現象は言語教育にとっての“dynamically shifted ecosystems”(p. 8)を象徴していると述べている。日本では, 山中(2019)はAIに端を発する「英語教育解体論」を展開しているが, 木村(2024)は, むしろAIと共存したり, それを取り込むことで, 英語教育は新たな可能性を示していくことになるだろうと指摘している。

2つ目は, 生成AIや機械翻訳利用に関する学習者の意識・姿勢・習慣等の調査である。例えば Kumar(2023)は, パキスタンの大学生のAI利用に関する意見や実践について調査を行い, こうした学習法は効果的かつ動機づけを高め, 積極的に推進すべきであると主張している。日本でも, 例えば佐藤(2023)の調査では, 全体的に機械翻訳が授業に導入されても英語学習は必要と感じる学生の割合は高いことが示された。一方, 横野(2023)は, 今後AI通訳・翻訳の発達とともに, 英語を教養科目としてのみ学ぶ大学生の英語学習意欲が低下し続ける可能性を危惧し, そのためにも教養課程での英語の授業の目的や方法の再考の必要性を訴えている。

3つ目は, 生成AIや機械翻訳を利用した実践研究であり, 多くがそのような学習法が有効であることを指摘している(Gayed et al., 2022; Lee & Davis, 2024; Syifaiddin & Yuliansyah, 2023; Ziwen & Hongwei, 2025)。例えば Zhang and Huang(2024)は, 大規模言語モデルは第2言語教育に革新的な変革をもたらす可能性を秘めていると主張し, Wei(2023)は, AIを搭載した学習アプリは, 学習を個人のレベルやニーズに合った, より面白く, 手軽に行える身近な体験とし, 内発的動機づけや自己効力感を高め, 自律学習の促進に効果的であると論じている。一方, Karataş et al. (2024)は, トルコの学生にChatGPTを利用した英語学習を行わせ, 文法, ライティング, 語彙習得に関してこうした学習法が有効であると指摘している。日本でも, 例えば山下他(2024)は, Grammarly, DeepL, ChatGPTを搭載した英作文支援ツール Transable を授業の中に導入した結果を報告し, 岩中(2023)は, ライティング支援ツール Criterion®のフィードバックの効用について論じている。

### 2.2 モデル文提示型ライティング学習法と教師によるフィードバック型ライティング学習法

遠山(2017)は, モデル文提示のタイミングと英語ライティング学習の関係性について調査した。日本人の大学生及び大学院生38名を, (a) 書く活動を行った後モデル文を見る群, (b) モデル文を見てから書く活動を行う群, (c) モデル文を見ずに書く活動を繰り返す群の3群に分類し, Schmidtの気づき仮説及びSwainのアウトプット仮説をベースとして, 以下の仮説を形成した。すなわち, (c)群は書いた時に感じる自分の知識の穴に対する気づきしかないものの, (b)群は加えて初めにモデル文を見ているため形式への気づきがあり, (a)群は自分が犯した誤りに対する

気づきも促されるため、3 群の中で (a) 群が最もライティング学習が促進すると予測した。しかし調査の結果、各群において予想された気づき自体は観察されたものの、(a) 及び (b) 群は (c) 群よりも有意に正確性の向上が見られたが、(a) 群と (b) 群の間で有意な差は確認されなかった。つまりモデル文提示自体の効果はあっても、異なるタイミングで提示することの効果はないと言える。遠山はさらに考察を進め、(c) 群でも確認された穴への気づきはライティング学習にとってさほど重要ではなく、また、(a) 群のみに誤りへの気づきが認められたが、(a) 群と (b) 群間に正確性の点で有意差が生じなかったことから、誤りへの気づきもそれ程重要ではなく、すなわち、(a) 群及び (b) 群ともに観察された形式への気づきが最も重要な要素である、と結論づけている。

田辺 (2023) は、日本人の中学生約 200 名に対し、教科書に掲載されているモデル文を定着させる活動 (段落毎にカードを作成し、つなぎ言葉を意識しながら内容を捉え、まとまりのある英文に並べ替える、等) を実施し、その後何度か段階的に書く練習をさせた上で、事後ライティング及び質問紙調査を実施した。その結果、モデル文の提示により形式への気づきが促され、また、内容、構成、正確さの点で格段にパフォーマンスが向上していることが確認された。とりわけ内容に関しては、伝えたい内容はあるもののそれをどう伝えればよいか分からなかった際に、モデル文から言語形式の気づきを得たことで表現の幅が広がったのではないかと、また構成に関しても、当初は一文ずつ書くことばかりに注意が向き、全体の構成までは考えられなかったが、モデル文を手本とすることで構成に関する気づきを得られたのではないかと田辺は推測している。

国外では、Ayyash and Khalaf (2016) が、これから書こうとしているトピックに関わる別の英文を読むことが、いかに学習者のライティング・パフォーマンスの向上に寄与するかを検証した。パレスチナの大学生 67 名を実験群と統制群に分け、実験群にはライティング前にそのトピックに関わる内容の英文エッセイを読ませ、両群のパフォーマンスを比較した。その結果、実験群のパフォーマンスの方が有意に優れ、とりわけ結論部で再度 *main idea* を強調したり、具体的な理由や詳細な例を提示する点で優れていることが確認された。Ayyash and Khala は、こうした *pre-writing activity* は、読んだものを自身のライティングのモデルとすることで、ライティング・スタイルや新しい語彙、構成、テクニック、読みやすさ等について学ぶことができるだけでなく、使用する語彙や表現、扱うトピックの幅を広げることにも繋がり得ると指摘している。

一方、Abbuhl (2011) は、自身の研究の結果に基づき、ライティング学習にもたらすこうしたモデル文を読む効果は、学習者の英語習熟度に拠らずに観察されるが、モデル文を読むだけでは学習者は“*target rhetorical features*” (p. 8) にまで意識が向かないことも多く、モデル文提示とともに“*explicit instruction*” (p. 9) を行う必要性も訴えている。

従来の教師によるフィードバック型ライティング学習に関しては、米村 (2018) は、日本人高校生 60 名を、直接的フィードバックをもらうグループ、間違いのみを指摘し訂正は学習者に委ねる間接的フィードバックをもらうグループ、修正はせず学習者自身が誤りを見つけて訂正させるグループ、全くフィードバックを行わずに次のライティングを行わせるグループの 4 グループに分類し、事前・事後ライティングを実施して、その効果を比較検証した。調査の結果から、前者 2 つのグループ、とりわけ間接的フィードバックのグループが、他のグループと比べ、事後ライティングにおいて正確性の向上が認められた。その理由として米村は学習経験を挙げている。すなわち、本実験の参加者は英語学習を始めてまだ 3 年と学習経験が浅く、Kellogg (2008) の言う「知識伝達モデル (*Knowledge-Telling*)」の段階にあり、このモデルではライティングは学習者の知識のみに依拠しており、知識の再構成は不可能であるという。文法的正確性が向上した領域につ

いて、項目毎に検証したところ、「動詞時制」の領域で顕著に正確性の向上が確認されたが、「動詞時制」といった比較的単純な文法間違いは、間違いを指摘されるだけでも、比較的容易に自身の知識の範囲内で修正が可能であり、これが正確性を押し上げた要因ではないかと米村は推測している。一方、前置詞の誤りは上昇しており、これは参加者が前置詞の用法に関する知識が極めて限定されており、結果的に知識の再構成が機能しなかったことによるものと考えられる。

望月 (2024) は、非明示的な直接的フィードバックを日本人大学生 24 名に与え、フィードバックが与えられる前後の英文を、流暢さ (文字数)、正確さ、複雑さの点で比較検証した。望月によれば、学習者が書いた英文を修正後 Google Classroom を通して返却すると、どこがどのように修正されたかが分からなくなってしまうため、この点で本フィードバックは「非明示的」であるという。調査の結果から、流暢さと複雑さは有意な向上が見られたが、それらについては他の要因によるものが大きいと推測され (例えばフィードバック後のライティングでは文字数が指定されたり、ライティングを行う直前に読んだ英字新聞の記事内の表現をそのまま使用した可能性がある点等)、正確さについては有意な変化は認められなかった。その要因として望月は二つの可能性を指摘している。一つは、この種のフィードバックに対する学習者の理解の限度である。たとえフィードバックが与えられても、何がどのように修正されたのかを理解しない限り、正確さの向上には繋がらない。また、ライティングを電子端末上で行ったことも一つの要因と考えられる。例えば Word 等のアプリを利用すると、綴りや文法等の誤りは自動的に訂正されてしまい、そうした誤りに対する気づきは促されない。この種のフィードバックが文法の正確さに及ぼす影響については、米村 (2020) も、日本の高校生 45 名を対象に、非焦点化フィードバックの観点から考察している。焦点化フィードバックとは、特定の文法項目に絞ったフィードバックのことで、最初からどの項目についてのフィードバックであるかが分かっているため認知的負荷が低く、特に習熟度が低い学習者にとって有効である。一方、非焦点化フィードバックとは、対象とする文法項目が限定されないフィードバックのことで、米村はそうした非焦点化された間接的フィードバックを、英語学習を始めて 3 年目の学習者に与え、直接的フィードバックよりも文法の正確性の向上が確認されたと報告している。文法項目が限定されていないとはいえ、学習経験を 3 年間積み上げた学習者は、その経験から間接的フィードバックの意味するところをある程度推測することができ、直接的フィードバックより高い教育的効果が得られたのではないかと推測している。

西尾・猪井 (2012) は、日本の中学生 95 名を対象に、英語ライティング学習に有効な修正フィードバックの種類と英語習熟度の関係性について調査している。本調査では、教師による口頭でのフィードバックを、メタ言語フィードバック (正しい文法的形式を直接与えるのではなく、間接的に示し、学習者が自ら正しい形式を産出できるように促す暗示的修正方法)、リキャスト (学習者が間違えた後で、エラーを正しく言い直す明示的修正方法)、意味内容をベースとしたフィードバック (形式ではなく、意味内容に着目させ、結果的に形式への気づきを促す修正方法) の 3 種類のフィードバックに分け、どの種類のフィードバックが過去形のアウトプットの正確性に寄与するかを調査した。結果、1 番目のフィードバックが他の 2 つのフィードバックに比べ、長期にわたり有意に維持されたことが判明したが、この種のフィードバックは習熟度が高い学習者のみ有効で、習熟度が中程度の学習者に対してはリキャストの方が有効であることが判明した。その要因として、言語的な知識や技能を身につけている学習者には、ある言語形式に注目しアウトプット活動を通して自らの言語的なエラーに気づき、修正を促すメタ言語フィードバックは有効

だが、言語的知識はあるものの、それをまだ活用する段階に至っていない学習者には、まずは正しい形式を伝え、それを反復練習させるリキャストが有効である点が指摘された。

## 2.3 脳血流と第2言語教育研究

19世紀に脳機能に関する研究が始まり、ブローカ野、ウェルニッケ野、角回・縁上回という言語機能を担う3つの脳の領域が特定された(藤本・田浦, 2011)。それに伴い、飛田・湯舟(2014)も指摘しているように、これまでは筆記試験やアンケートの結果といった経験則に基づき学習成果の要因や学習の際の認知過程を推測するといった方法が主流であったが、脳のこの3領域の活動状態を調べることで、より直接的な証拠やデータを入手することが可能となった。これらは脳機能イメージングと呼ばれる手法によって行われ、脳波(Electroencephalogram)や機能的磁気共鳴画像法(Functional Magnetic Resonance Imaging)といったものがあるが、中でも近赤外線分光法(near-infrared spectroscopy [NIRS])は、非侵襲的、そして何より学習者があるタスクに従事している際の脳血流量の測定が可能である点で(Farrukh et al., 2025)、高額な点や煩雑さの点で研究自体の数は少ないものの、第2言語教育研究においては比較的好んで利用されてきた。例えば湯舟(2011)は、チャンクで区切ったリーディングのトレーニングを積むことで、脳への負担がいかに軽減されるかをNIRSを用いて検証した結果、文法解析の作業を司るブローカ野の血流量が減り、速く読むために必要な、音韻符号化の高速化に関連する聴覚性言語野のウェルニッケ野での上昇が見られたことを受け、リーディングの自動化が促進されたと推測している。また、飛田・湯舟(2014)は、NIRSを使って学習者特性と教材の相関関係を調べ、表面上は学習しているように見えても、実は脳が活性化していない場合もあることを指摘した。Farrukh et al. (2025)は、パキスタンの大学生にあるイメージを提示し、それを言語化するタスクを行わせたところ、“language production”(p.1)を担う左脳に脳血流の上昇が認められたと報告している。大石・木下(2008)は、学習者に第1言語と第2言語のリスニング課題を行わせ、NIRSを用いてその際の脳血流量を測定した。まず、脳血流量の変化は実際には何を意味しているかという問いに対し、脳血流量と注意量は正の相関があるとする先行研究の結果を踏まえ、課題に振り向ける注意量を示しており、さらに言えば、課題が難しくなれば当然のことながら注意量も増え、課題の難易度とも正の相関があるとした。これを仮説として、第2言語によるリスニング課題に従事している学習者の脳血流量を測定したところ、実際に第1言語のそれよりも高いことが示され、第2言語を伴う課題の方が、情報処理の自動化の度合いが低く、結果、より多くの注意量を必要とすると指摘している。

## 2.4 先行研究を踏まえた本研究のリサーチ・クエスチョン

上記の先行研究から、教師によるフィードバックは、学習者の習熟度をよく踏まえ、そのレベルに即したフィードバックを与えることが重要であり、反対にレベルに適合しないフィードバックを与えると、いくらその中で正しい形式を伝えても、学習者がそれに気づかないことも多いことが確認された。いずれにせよ、学習者のレベルに即したフィードバックを与えることが非常に難しい方法だと言えるだろう。一方、モデル文提示型のライティング学習法も、気づきをベースとしている点では同じだが、学習者の習熟度がある程度高い場合は、提示されたモデル文から、自身のニーズに合わせ好きなだけ吸収できる点で、教師によるフィードバック型学習法よりも教育的効果が高いと言えるのではないだろうか。一方、上記の先行研究にもあるように、学習者のライティング・パフォーマンスを数値化して、その変化の要因となり得るもの(例えば学習者の

認知過程)を質問紙調査の結果等だけから推測しても、直接的な証拠とは言い難く、根拠としては不十分な場合もあり得る。

以上より、本研究のリサーチ・クエスチョンは以下の通りである。

- (1) 日本の大学生としては比較的英語習熟度の高い学習者に、モデル文提示型のライティング学習法と教師によるフィードバック型のライティング学習法を実施させた場合、どちらの学習法の方が教育的効果は高いか(どちらの方がライティング・パフォーマンスが高いか)。
- (2) モデル文提示型のライティング学習法を実施している時と教師によるフィードバック型のライティング学習法を実施している時とで、両者の間に脳血流量の差異は見られるか。また、差異が見られる場合、両者のライティング学習の際の認知過程(例えば文のどの要素(構文、文法、語彙等)にとりわけ注意を振り向けているか)に具体的にどういった差異があると推測されるか。

また、今回のモデル文提示型の学習法は、従来の方法とは機械翻訳のモデル文を利用する点で異なり、その点で特異な教育的効果が見られる可能性がある。したがって、

- (3) 機械翻訳が生成したモデル文を利用することで特異な教育的効果は見られるか。それは何か。

以上3つを本研究のリサーチ・クエスチョンに据える。

### 3. 方法

調査の際にまとまった量の英文を書いてもらうため、また、両群の条件を同じくするため、TOEIC スコア約 600 点前半の 18~20 歳の商学を専門とする日本人大学生 18 名を対象に、機械翻訳 (Google Translate; 調査当時は今ほど生成 AI が普及していなかった為) を提示して英語ライティング学習を行う参加者 A 群 9 名 (男: 8 名, 女: 1 名) と、利用せずに学習を行う参加者 B 群 9 名 (男: 6 名, 女: 3 名) とに分け、NIRS を用いて脳血流量を測定した。近赤外線は生体組織中のヘモグロビンに吸収されるため、近赤外線の光を脳に照射することで、ヘモグロビン濃度の相対的変化や血中の酸素化状態、また脳血流量自体の変化を測定することが可能である。また、血流量を測定する脳の領域はブローカ野と呼ばれる領域を中心に行った。ブローカ野は「言語中枢」とも称され、文法処理や統語処理を司る脳の領域とされている。利き手が右手の場合、ブローカ野はほぼ 100% 左脳の前頭葉に位置するといわれているが、利き手が左手の場合、必ずしもブローカ野が同部位に位置するとは限らないため、参加者は利き手が右手の学生のみ限定した。

参加者 A 群には、こちらで用意した日本語 (参加者の専門を考慮してビジネス・レター) とそれを Google Translate で英語に翻訳したものを比較し (図 1)、学習した上で (3 分間)、よく似た内容の別の日本語 (図 2) を自力で英訳してもらった (この間辞書は使用可で時間制限なし)。機械翻訳がライティング学習の際のロール・モデルとなり得るかどうかについては検証すべき余地があるが、インターネット上に溢れている英語使用を基に生成され、最近ほとんど誤りも認められない為、本研究では一定のロール・モデルになり得ると判断した。図 1 の日本語と機械翻訳を比較し、その後、図 2 の日本語を英訳している間の参加者 A 群の脳血流量を測定し、同時に、脳血流量のデータとその間の参加者 A 群の動作を比較するため、学習及びライティング中の参加者の様子を許可を得て録画した。その後、両者 (脳血流量のデータと参加者の様子) を見比べながら、英訳した際に何を考えていたか等について参加者 A 群と半構造化インタビューを実施した。

一方、参加者 B 群には、まずこちらで用意した日本語 (ビジネス・レター) (図 1 の上半分の日本語) を自力で英語に訳してもらい (この間辞書は使用可で時間制限なし)、その後それを 20 年

以上大学英語教育に携わっている大学教員に、通常行っている方法と同じ方法 (明示的な直接的フィードバック) で添削してもらうよう依頼し、後日返却した添削を学習した後で (3 分間)、よく似た内容の別の日本語 (図 2) を再度自力で英訳してもらった (この間辞書は使用可で時間制限なし)。教師からの添削を読み、図 2 の日本語を英訳している間の参加者 B 群の脳血流量を測定し、同時に、脳血流量のデータとその間の参加者 B 群の動作を比較するため、学習及びライティング中の参加者の様子を許可を得て録画した。その後、両者を見比べながら、英訳した際に何を考えていたか等について参加者 B 群と半構造化インタビューを実施した。

## 図 1

### 1 回目に英訳する日本語とその機械翻訳 (英文)

< 1 回目 >

私のサンフランシスコ出張中のお手伝いのお申し出、ありがとうございます。これが私のサンフランシスコへの最初の訪問になりますので、5月13日から5月20日の期間でホテルを予約して頂けますでしょうか？今回の出張では、貴社との業務提携についてさらに突き詰めた話し合いが出来ればと考えております。さらにこの機会に、サンフランシスコ近郊にある貴社のいくつかの工場も視察してみたいと思っております。また、可能であれば、あなたのご友人である会計士のヘンリー・スミス氏との打ち合わせも手配して頂ければ有難いです。私の予定は5月15日なら午後、5月18日なら1日中、空いております。ご協力に重ねて感謝申し上げます。

機械翻訳 (Google Translate) で訳したもの

Thank you for your help during my business trip to San Francisco. Since this is my first visit to San Francisco, could you please book a hotel for the period between May 13th and May 20th? On this business trip, we would like to have more in-depth discussions about the business alliance with your company. We would also like to take this opportunity to visit some of your factories near San Francisco. Also, if possible, we would appreciate it if you could arrange a meeting with your friend, accountant Henry Smith. My schedule is open in the afternoon on May 15th and all day on May 18th. Thank you again for your cooperation.

## 図 2

### 2 回目に英訳する日本語

< 2 回目 >

私のロサンゼルス出張中のお手伝いのお申し出、ありがとうございます。これが私のロサンゼルスへの最初の訪問になりますので、10月16日から10月24日までの期間で駅周辺のホテルを予約して頂けますでしょうか？今回の出張では、以前お送り頂いたカタログに掲載されていた商品を実際に拝見させて頂きたいと思っております。さらにこの機会に、それらの商品のいくつかの生産ラインを見学させて頂くことができればと思っております。また、可能であれば、あなたの上司であるジョン・ウィリアムズ氏と次のプロジェクトについて話し合いたいので、打ち合わせを手配して頂ければ有難いです。私の予定は10月19日なら午前、10月22日なら17時以降なら空いております。ご協力に重ねて感謝申し上げます。

## 図 3

### 3 回目に英訳する日本語

< 3 回目 >

私のニューヨーク出張中のお手伝いのお申し出、ありがとうございます。これが私のニューヨークへの最初の訪問になりますので、12月21日から12月27日までの期間で、貴社近くのホテルを予約して頂けますでしょうか？今回の出張では、12月10日付きのメールで頂いた我が社の商品に関するクレームに関して、直接そちらに伺い、詳しくご説明したいと思っております。また、商品を実際に使って、その使用方法についてもデモンストラーションを行いたいと思っております。さらにこの機会に、他製品となりますが、我が社のA製品の大口契約についても詳しいご相談をさせて頂きたいと思っておりますので、貴社の調達部長のマイケル・ミラー氏との打ち合わせを手配して頂ければ有難いです。私の予定は12月23日なら午後、12月26日なら15時以降なら空いております。ご協力に重ねて感謝申し上げます。

両群ともに、上記のライティング学習から約1か月後、長期記憶をはかるために、よく似た内容の別の日本文 (図 3) を再度自力で英訳してもらった (この間辞書は使用可で時間制限なし)。この1か月の間、両群ともに英語の学習について特に制限を設けることはなかったが、最後のライティング・パフォーマンスに直接的な影響を与えるような要素を除去するため、約1か月後に再度ライティングを実施する旨は伝えたものの、具体的に何を書くかについては伝えず、また、日本文や機械翻訳、参加者が書いた英文等は、この時点で全て回収した。

後日、それぞれの群から得た脳血流量のデータを MATLAB R2022b を用いて分析・グラフ化し、その後 IBM SPSS Statistics 26 を用いて、最初の3分間の学習、すなわち参加者 A 群は日本文と機械翻訳の英文を比較している間、参加者 B 群は自分が書いた英文の添削を見直している間の脳血流量に統計的な有意差があるかどうかを検証した。

参加者が訳した英文については、後日 20 年以上日本の大学の英語教育に携わっている英語のネイティブスピーカーの教員に依頼し、(1) Logical connection (文同士が論理的に結びついているかどうか)、(2) Linguistic correctness (文法的に正しい英文が書けているかどうか)、(3) Linguistic appropriateness (ビジネス・レターの形式に沿った書き方のスタイルや用法等が用いられているかどうか)、(4) Vocabulary (ビジネス・レターの形式に沿った表現や言い回しが用いられているかどうか)、(5) Spelling (単語のスペルが正しく書けているかどうか)、(6) How much the goal of the letter is achieved (レターの目的がどれ程達成されているか)、(7) How well the letter is written as a whole の7項目について、5 (Excellent)~1 (Not good at all) の評点法で評価してもらい、両群のライティング・パフォーマンスに統計的有意差があるかどうかを検証した。この7項目については、Schaefer (2008) の分析的評価ルーブリックを基に、本研究の参加者が書く英文が、目的を相手に効率的に伝える必要のあるビジネス・レターである点も考慮して設定した。また、それぞれの群において1回目と2回目のライティング・パフォーマンスに統計的有意差があるかどうかを検証した。

インタビューの内容は全て録音し、後日書き起こした上で、質的データ分析ソフト NVivo 12 を用いて分析を行った。

本実験にあたり、あらかじめ著者が所属する大学の倫理審査委員会の承認を得 (承認番号: 2023-h02)、脳血流量データ、ライティング・パフォーマンスのスコア、インタビュー内容の論文使用について各参加者からインフォームド・コンセントを得た。

## 4. 結果

### 4.1 脳血流量

参加者 A 群、参加者 B 群ともに最初の3分間の学習の間の脳血流量データ (oxy-Hb (酸素化ヘモグロビン量)) は正規分布をなしているとは言えなかったため、Mann-Whitney の  $U$  検定を実施した。結果、両者の血流量に統計的有意差は見られなかったが ( $p = .113 > .05$ )、効果量は大きく、今後さらなる検討が必要であると考えられる (表 1)。

表 1

統制群と実験群の脳血流量の比較

参加者 A 群	参加者 B 群	$p$ 値
---------	---------	-------

<i>M</i> (oxy-Hb)	<i>SD</i>	<i>M</i> (oxy-Hb)	<i>SD</i>		<i>d</i> (効果量)
653.44	31336.37	23194.02	25612.88	0.113	0.79

表 2

両群の脳血流量のカテゴリ毎の人数分布

脳血流量 (oxy-Hb)	参加者 A 群 (N=9)	参加者 B 群 (N=9)
75,001 ~ 100,000	0	1
50,001 ~ 75,000	0	0
25,001 ~ 50,000	1	1
0 ~ 25,000	5	7
-25,000 ~ -1	2	0
-50,000 ~ -25,001	0	0
-75,000 ~ -50,001	1	0

## 4.2 半構造化インタビューと脳血流量の変化

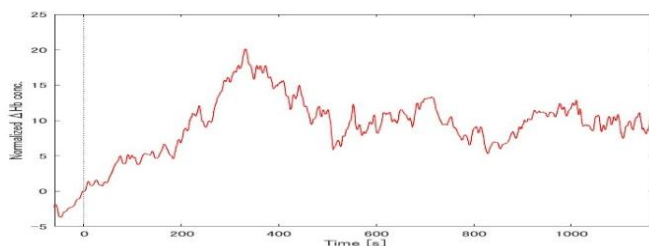
最初の 3 分間 (0-180 秒) の学習中の心理状態について参加者と半構造化インタビューを行った。以下の図は参加者の脳血流量の変化をグラフ化したもので、右のコメントは、この間の心理状態について、当該参加者のインタビュー内容を簡単に要約したものである。トリガ時 (学習を開始した瞬間。参加者 A 群は日本文と機械翻訳とを比較し始めた瞬間、参加者 B 群は返却されたフィードバックを学習し始めた瞬間を指す。グラフ上では横軸 (時間) の 0 で示されているポイント) 以降の血流量の変化を調べるため、トリガ時より 1 分前から測定し始め、データ入手後、加算平均処理 (絶対値ではなく、トリガ時からの相対的な血流量の変化を検証するためのデータ処理)、ベースライン処理 (トリガ時に血流量を 0 に設定するためのデータ処理)、波形強度の正規化 (参加者間の比較を行えるようにするためのデータ処理) を施して脳血流量をグラフ化した。

### 4.2.1 参加者 A 群

#### 【参加者 A 群 1】

図 4

参加者 A 群 1 の脳血流量の変化のグラフ

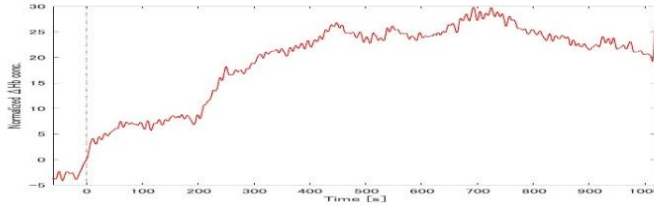


- ・最初の 3 分 (0-180 秒) の比較的早い段階で、全文を読むというよりは、これは後のライティングで使えるなどと思える所だけにスポットライトを当てていくという意識が変わった。
- ・機械翻訳中の語彙や熟語に特に納得がいった。

### 【参加者 A 群 2】

図 5

参加者 A 群 2 の脳血流量の変化のグラフ

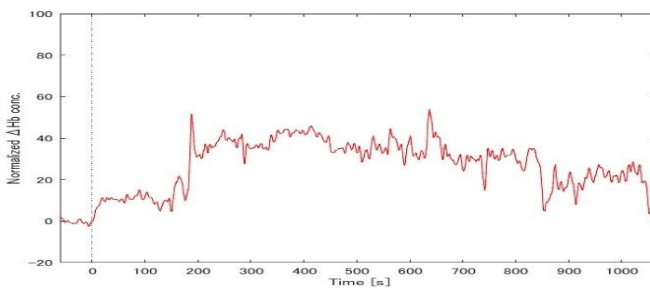


- ・最初の 3 分の学習の間、機械翻訳中の、alliance 等といった、日本語から英語に訳しづらい単語や表現に目が行った。

### 【参加者 A 群 3】

図 6

参加者 A 群 3 の脳血流量の変化のグラフ

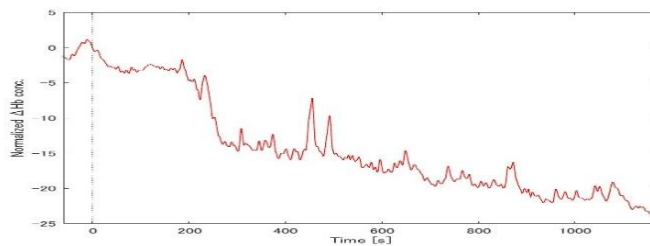


- ・最初の 3 分の学習の間、機械翻訳中に見慣れない表現が多くあり、新鮮に感じ、勉強になった。
- ・この間、文法や構造、構成等についてはほとんど考えなかった。
- ・[最初の 3 分の学習の後] 実際に英文を書き始めてからはかなり頭を使った気がする。

### 【参加者 A 群 4】

図 7

参加者 A 群 4 の脳血流量の変化のグラフ

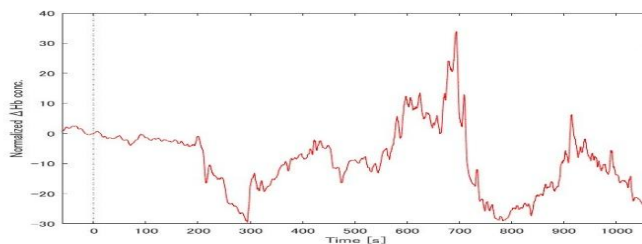


- ・最初の 3 分の学習の間、「近郊にある貴社のいくつかの工場」や「あなたのご友人である会計士のヘンリー・スミス」という表現の語順が気になり、英文を確かめた。

### 【参加者 A 群 5】

図 8

参加者 A 群 5 の脳血流量の変化のグラフ

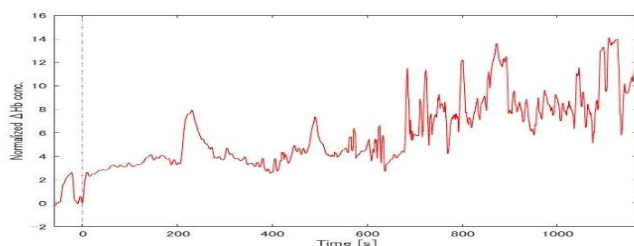


- ・最初の 3 分の学習では、一巡目はざっと機械翻訳を読むだけで、二巡目で細かく確認していった。
- ・この間、more in-depth discussions という表現や、「視察」を visit と表現しているところ等、表現の仕方が特に気になった。

### 【参加者 A 群 6】

#### 図 9

参加者 A 群 6 の脳血流量の変化のグラフ

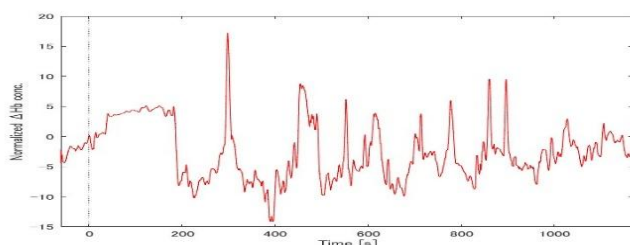


・最初の 3 分間、文の構造というよりは、こういう表現を用いて、こういう単語を使ったら文を作れるのかと思うことがほとんどだった (alliance, in-depth, visit, arrange a meeting, take this opportunity 等)。

### 【参加者 A 群 7】

#### 図 10

参加者 A 群 7 の脳血流量の変化のグラフ

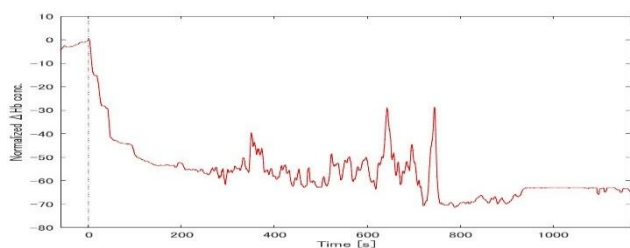


・最初の 3 分間は、機械翻訳をただ受動的にリーディングしているという感覚だった。

### 【参加者 A 群 8】

#### 図 11

参加者 A 群 8 の脳血流量の変化のグラフ

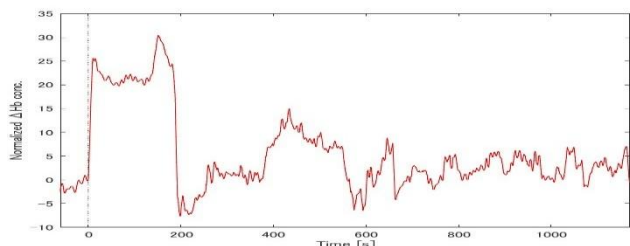


・最初の 3 分の学習の間、英訳しづらい表現を、こう訳すんだと興味を持ちながら見ている感じだった (alliance, in-depth 等)。  
・そういう表現を中心にピックアップして読んでいく感じだった。

### 【参加者 A 群 9】

#### 図 12

参加者 A 群 9 の脳血流量の変化のグラフ



・最初の 3 分の学習の間、分からない単語を対応させていって、特に知らない表現、見慣れない表現に意識が行った。  
・後半くらいから文の流れも意識するようになった。

参加者 A 群は、最初の学習時間の 3 分の間、機械翻訳のモデル文中に使用されている語彙や表現に意識が向いたという声が圧倒的多数を占めた。また、この間、文法処理や統語処理を司るブローカ野の脳血流量が概してそれ程上昇していないことがグラフから読み取れる。

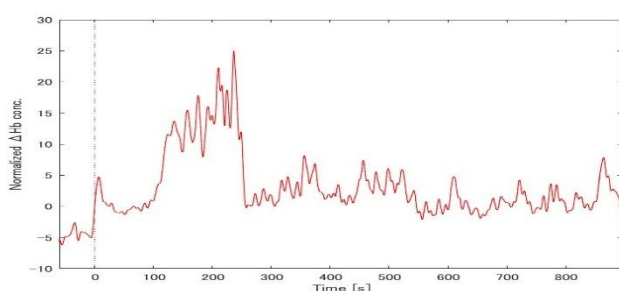
#### 4.2.2 参加者 B 群

一方、参加者 B 群に関しては以下のような結果が得られた。

##### 【参加者 B 群 1】

図 13

参加者 B 群 1 の脳血流量の変化のグラフ

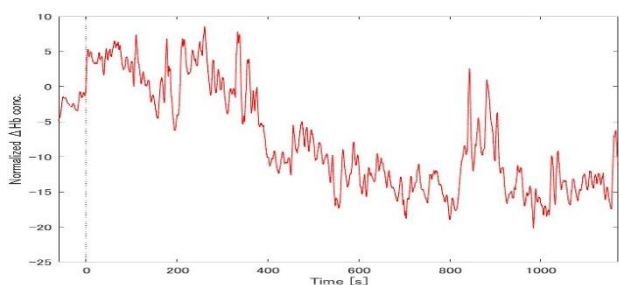


- ・最初の 3 分間の学習時間のうち、120 秒辺りから 180 秒に当たる部分は文法や文の流れ、構成等をじっくり考えた気がする。
- ・前半 (0~120 秒位の間) は修正された箇所だけをさっと見直ただけだった。

##### 【参加者 B 群 2】

図 14

参加者 B 群 2 の脳血流量の変化のグラフ

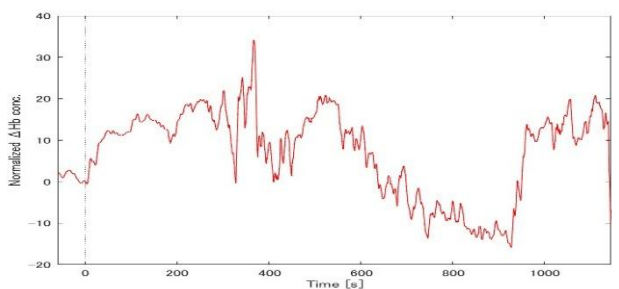


- ・最初の 3 分間の学習時間のうち、0 から 60 秒の間はそれほど深く考えていなかった。
- ・60 秒からもう一度見直していく過程で、文の構造等を意識してじっくり読みこむようになった。
- ・100 秒辺りで、文法や構文的に最も難しい箇所に行き当たった。

##### 【参加者 B 群 3】

図 15

参加者 B 群 3 の脳血流量の変化のグラフ

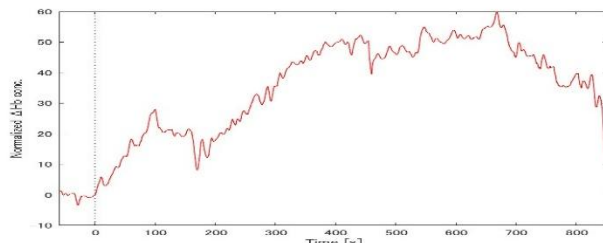


- ・最初の 3 分間の学習時間の間は一つ一つの文の構造等を、丁寧にじっくり考えていった。
- ・先生の添削は構造や文法等についてのもが多く、しっかり時間をかけて確認したい。

### 【参加者 B 群 4】

図 16

参加者 B 群 4 の脳血流量の変化のグラフ

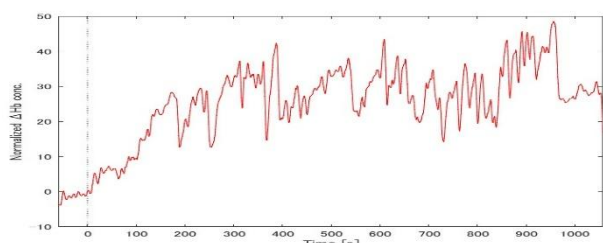


・最初の 3 分間で教師からのフィードバックを読んだが、100 秒辺りで、見慣れぬ appreciate の使い方についてのフィードバックに行き当たり、そこで少し考え込んでしまった。

### 【参加者 B 群 5】

図 17

参加者 B 群 5 の脳血流量の変化のグラフ

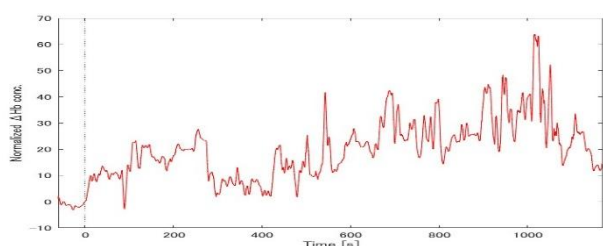


・最初の 3 分間教師からのフィードバックを読んでいる間、in や at の違いといった細かい内容も大事だと思いつつ、どちらかといえば文の流れや文法等といった全体に関わるような内容に目が行った。

### 【参加者 B 群 6】

図 18

参加者 B 群 6 の脳血流量の変化のグラフ

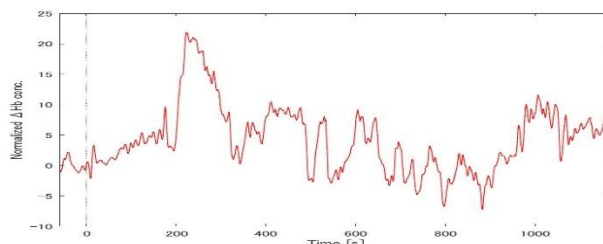


・最初の 3 分間は、全体的に結構時間をかけ、一文一文教師からのフィードバックを丁寧に読んでいった。

### 【参加者 B 群 7】

図 19

参加者 B 群 7 の脳血流量の変化のグラフ

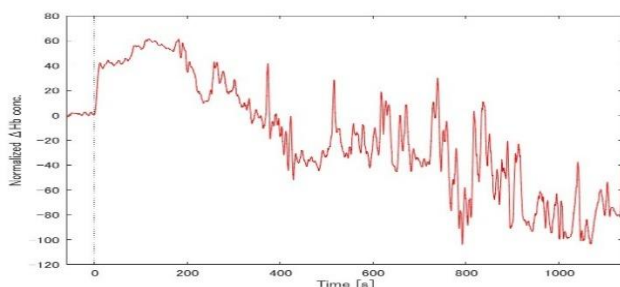


・学習時間の 3 分間の間は、最初の方は文の全体的な直しが多かったので少し頭を使ったが、後半は 1 語や 2 語程度の語彙の訂正が多かったのであまり頭を使わなかった。

### 【参加者 B 群 8】

図 20

参加者 B 群 8 の脳血流量の変化のグラフ

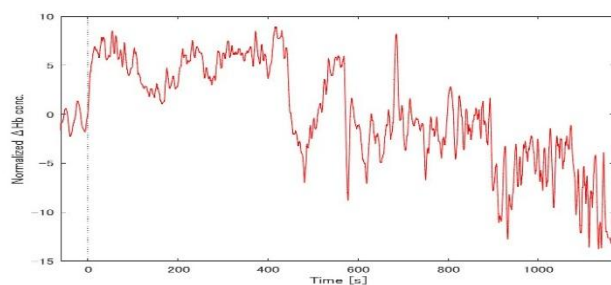


- ・最初の 3 分間、教師からのフィードバックを読みつつ、前に自分がどう書いたかを一文一文思い出していた。
- ・1 分過ぎた辺りで上記の作業を一旦終え、その後はこれから書くことになる英文を自分なりに予測し、シミュレーションをした。

### 【参加者 B 群 9】

図 21

参加者 B 群 9 の脳血流量の変化のグラフ



- ・最初の 3 分の学習時間が終わる直前 (180 秒頃) に、Henry Smith, an accountant who is your friend (教師からのフィードバック) という表現が、文の構造も複雑で、本当にこれで良いのかな? と思い、少し考えこんでしまった。

参加者 A 群と比較して、参加者 B 群は、最初の学習時間の 3 分間、文法や構造、全体的な構成等について色々考えをめぐらしたり、悩んだりしたというコメントが多かった。また、文法や構造について考えあぐねている時に、ブローカ野の脳血流量が上昇していることが確認できる。

## 4.3 ライティング

方法のセクションで述べたように、参加者 A 群に関しては日本文と機械翻訳とを見比べ学習した直後に書いた 1 回目の英文と約 1 か月後に書いた 2 回目の英文、参加者 B 群に関しては添削された英文をよく見直した直後に書いた 1 回目の英文と約 1 か月後に書いた 2 回目の英文を、(1) Logical connection, (2) Linguistic correctness, (3) Linguistic appropriateness, (4) Vocabulary, (5) Spelling, (6) How much the goal of the letter is achieved, (7) How well the letter is written as a whole の 7 項目について、ネイティブスピーカーの教員に 1~5 点の評点法で評価してもらい、両群のスコアに統計的有意差があるかどうかを検証した。項目 (1) の 2 回目及び項目 (5) の 1 回目は両群共に正規分布をなしていないとは言えなかったため  $t$  検定を、それ以外に関しては正規分布をなしているとは言えなかったため Mann-Whitney の  $U$  検定を実施して検証した (表 3)。

表 3

参加者 A 群と参加者 B 群のライティング・パフォーマンス (スコア) の比較

	A 群 (N=9)		B 群 (N=9)		p 値	d (効果量)
	M	SD	M	SD		
(1) Logical connection						
1 回目	4.11	0.78	4.44	0.53	0.436	0.50
2 回目	3.67	0.87	3.67	0.87	1.000	0.00
(2) Linguistic correctness						
1 回目	3.11	0.78	2.67	0.70	0.258	0.59
2 回目	2.67	0.70	3.00	0.70	0.387	0.47
(3) Linguistic appropriateness						
1 回目	3.33	1.00	3.11	0.60	0.666	0.27
2 回目	2.89	0.60	3.22	0.67	0.340	0.53
(4) Vocabulary						
1 回目	4.22	0.67	3.78	0.44	0.190	0.79
2 回目	3.33	0.87	3.89	0.78	0.297	0.67
(5) Spelling						
1 回目	3.89	0.78	3.56	0.88	0.409	0.40
2 回目	4.78	0.44	4.44	0.73	0.387	0.56
(6) How much the goal of the letter is achieved						
1 回目	4.00	0.71	4.00	0.50	1.000	0.00
2 回目	3.33	1.00	3.78	0.67	0.340	0.52
(7) How well the letter is written as a whole						
1 回目	3.22	0.83	3.00	0.50	0.489	0.32
2 回目	3.00	0.87	3.33	0.50	0.258	0.47

表 3 から、(5) Spelling の 2 回目に関しては両群ともに平均値と標準偏差の和が測定値の最大値 5 を超えており、天井効果が発生している可能性があるが、1 回目及び 2 回目ともに、全ての項目において両群のスコアに統計的有意差はないことが分かる。ただし効果量に目を遣ると、(4) Vocabulary の 1 回目が 0.79 と他の項目や回と比べてかなり高く、直前に機械翻訳のモデル文を確認したことが、vocabulary に関して、かなりライティング・パフォーマンスに良い影響を与えたことが理解できる。

一方、各群の 1 回目と 2 回目のライティング・パフォーマンスを比較するため、項目 (1) の参加者 A 群については、1 回目、2 回目ともに正規分布をなしていないとは言えなかったため *t* 検定を、それ以外に関しては正規分布をなしているとは言えなかったため Wilcoxon の符号付き順位検定を実施した (表 4)。

表 4

各群の 1 回目と 2 回目のライティング・パフォーマンス (スコア) の比較

	1 回目		2 回目		<i>p</i> 値	<i>d</i> (効果量)
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
(1) Logical connection						
参加者 A 群 ( <i>N</i> =9)	4.11	0.78	3.67	0.87	0.035	0.84
参加者 B 群 ( <i>N</i> =9)	4.44	0.53	3.67	0.87	0.084	0.65
(2) Linguistic correctness						
参加者 A 群 ( <i>N</i> =9)	3.11	0.78	2.67	0.70	0.102	0.61
参加者 B 群 ( <i>N</i> =9)	2.67	0.70	3.00	0.70	0.317	0.33
(3) Linguistic appropriateness						
参加者 A 群 ( <i>N</i> =9)	3.33	1.00	2.89	0.60	0.046	0.84
参加者 B 群 ( <i>N</i> =9)	3.11	0.60	3.22	0.67	0.564	0.19
(4) Vocabulary						
参加者 A 群 ( <i>N</i> =9)	4.22	0.67	3.33	0.87	0.023	1.14
参加者 B 群 ( <i>N</i> =9)	3.78	0.44	3.89	0.78	0.655	0.14
(5) Spelling						
参加者 A 群 ( <i>N</i> =9)	3.89	0.78	4.78	0.44	0.046	0.84
参加者 B 群 ( <i>N</i> =9)	3.56	0.88	4.44	0.73	0.054	0.76
(6) How much the goal of the letter is achieved						
参加者 A 群 ( <i>N</i> =9)	4.00	0.71	3.33	1.00	0.167	0.47
参加者 B 群 ( <i>N</i> =9)	4.00	0.50	3.78	0.67	0.414	0.27
(7) How well the letter is written as a whole						
参加者 A 群 ( <i>N</i> =9)	3.22	0.83	3.00	0.87	0.317	0.33
参加者 B 群 ( <i>N</i> =9)	3.00	0.50	3.33	0.50	0.257	0.39

表 4 から、参加者 B 群は (1) と (6) 以外の全ての項目において、1 回目から 2 回目のライティングにかけて、統計的有意差はないものの平均点が上昇している一方、参加者 A 群は (5) を除き、全ての項目で平均点が下がっており、かつ (1), (3), (4), (5) の項目で統計的有意差があることが分かる。ただし効果量に目を遣ると、参加者 B 群はほぼ全ての項目で平均点が上がっているとはいえ、その効果はさほど大きくないと言える。特筆すべきは、(5) Spelling のみ、両群ともに平均点が大幅に上昇し (参加者 A 群は統計的有意差もあり)、効果量もかなり大きい点である。また、(4) Vocabulary の参加者 A 群の効果量が 1.14 と、他の項目や群と比べて突出して高く、vocabulary に関しては、前のライティングから 1 か月間を空けたことが、参加者 A 群の最後のライティング・パフォーマンスにかなりの負の影響を与えたことが理解できる。

## 5. 考察

表3より、1回目のライティングでは概して参加者A群の方が平均点が高く、2回目のライティングではそれが逆転してしまっている(参加者B群の平均点の方が高い)ことが確認されたが、 $p$ 値や効果量から、それらの現象が起こる蓋然性は低く、また機械翻訳のモデル文を見たり、教師からのフィードバックを読んだりしたことの影響はとりわけ大きいわけではないと考えられる。ただし、**vocabulary**の項目の1回目は効果量が0.79と、他の項目や回と比べてかなり高い。つまり、機械翻訳のモデル文を見たり、教師からのフィードバックを読んだりしたことの影響は全体としてはさほど大きくないものの、**vocabulary**に関してはかなり大きな影響を及ぼしている、具体的には、直前に機械翻訳のモデル文を読むことが、教師からのフィードバックをもらうことよりも、ライティング・パフォーマンスにかなりの正の影響を与えたことが確認される。このことは、機械翻訳のモデル文中に見慣れぬ単語や表現が多くあり、とても勉強になったという参加者A群のインタビュー内容とも符合する。一方、表4から、参加者B群は1回目のライティングから2回目のライティングにかけて概してパフォーマンスの向上が見られ(統計的有意差はなし)、参加者A群は反対にパフォーマンスの低下が見られた(項目によっては統計的有意差もあり)。とりわけ**Vocabulary**の項目に関しては、参加者A群の効果量は1.14と、他と比べて突出しており、1回目のライティングから1か月間を空けたことが、参加者A群の2回目のライティングにかなりの負の影響を与えていることが確認された。上記のことと考え合わせると、参加者A群は、機械翻訳のモデル文中に多くの知らない単語や表現を見出し、それらを直後の自身のライティングには反映させることができ、結果、表3の(4)**Vocabulary**の1回目の効果量を正の方向にかなり上昇させることに至ったが、その後おそらくそれらを全く、またはほとんど使わず約1か月を過ごした結果、その反動で表4の(4)**Vocabulary**の効果量を負の方向へ大幅に上昇させ、また、表3の(4)**Vocabulary**の2回目の効果量も0.67と、1回目とは反対の方向に比較的大きな上昇を見せるに至ったと考えられる。

これは脳血流量の結果とも符合する。機械翻訳のモデル文中に見知らぬ単語や表現が多くあり、勉強になったと答えた参加者A群の学習時の脳血流量は全体的にそれ程上昇していないのに対し、教師からのフィードバックを読んでいた時の参加者B群の脳血流量は概して上昇している。このことは表1に示した効果量からも裏付けられる。これは参加者B群が、文法や統語等により多くの注意を振り向けた証拠であると考えられる。参加者A群は、たしかに機械翻訳のモデル文中に見慣れぬ単語や表現を見出したものの、参加者B群のように文法や統語等にも細かく注意を払ったわけではない。このことが、1回目のライティングでは大変勉強になったと言った単語や表現が、2回目のライティングでは、表3及び表4の(4)**Vocabulary**の効果量から推測されるように、大半が忘れられてしまった要因と考えられるのではないだろうか。次の参加者A群4のインタビューのコメントはまさにこの点を示唆していると言えるだろう。

たしかに自分で書いているだけだと、同じような、自分に馴染みのある表現しか使わない傾向があります。一方で、[機械翻訳を]見ればたしかに意外性もあるし、あ、こんな風を使うんだと思うことも沢山あるんですけど、その時は[そういう風に感じて]それで終わってしまうんです。それを自分の中で取り込んで使ってみて、初めて覚えられるんだと思いますし、次に使えるんだと思います。

冒頭の先行研究のセクションで挙げたモデル文提示型ライティング学習法について調査した研究 (田辺, 2023; Ayyash and Khalaf, 2016) は、モデル文を提示することで、内容、構成、語彙、正確さ等、多面的なパフォーマンスの向上が見られたが、本研究においても参加者 A 群は、学習直後は概して B 群よりも高いパフォーマンスを示し、中でも語彙に関しては高いパフォーマンスを示した。その理由として、今回のモデル文が比較的短く、かつ機械翻訳をベースにしていることによると考えられる。おそらく参加者達は、日本語訳もつけられた短いモデル文の中で、内容や構成よりも、見慣れぬ語彙や表現の方により注意が向いたのではないだろうか。また、これらの先行研究の多く (田辺, 2023; 遠山, 2017; Ayyash and Khalaf, 2016) は、モデル文提示後も長期に渡ってライティング・パフォーマンスを維持し得ているかどうかについて調べていないが、本研究の結果から、モデル文提示の学習効果は概して一時的で、長期的な定着に至っていないことが確認された。一方、先行研究のセクションで挙げた教師によるフィードバック型ライティング学習法に関する先行研究 (西尾・猪井, 2012; 米村, 2018, 2020) は、概して間接的フィードバックの方が、直接的で明示的なものよりも学習効果が高いことを示しているが、本研究の結果から、たとえ直接的で明示的なフィードバックであっても、再度文法や統語等の観点から自分なりに考え直すことで学習効果が上がる、すなわち長期的な定着へと繋がる可能性があることが確認された。

尚、結果のセクションでも述べたように、表 4 を見ると、(5) Spelling に関して、両群ともに 1 回目から 2 回目のライティングにかけてかなりスコアを伸ばしていることが分かるが、これは 2 回目のライティングでは、自分が確実に綴りを知っている単語のみを使用したことによるものと推測される。もちろん文法等も自身が理解している文法項目を努めて使用するということがあったであろうが、よく理解できていない項目であっても、文脈上使用せざるを得なかったものもあったのではないだろうか。一方、語彙に関しては、自分が確実に綴りを知っている単語に代替することは比較的容易であり、このことが 2 回目の平均値を押し上げた要因であると考えられる。

## 6. 結論

以上から、本研究のリサーチ・クエスチョンへの解答は以下のように示される。

- (1) たしかにある程度英語習熟度の高い学習者は、機械翻訳のモデル文から適切な単語や表現を自力で見つけ出すことができ、その点では教育的効果はあると言えるが、それが必ずしも長期的なライティング・パフォーマンスの向上に繋がるわけではない。一方、教師によるフィードバック型の学習は、文法・統語等にも注意を払いながら英文を読んだり、書いたりする機会が多く、さらにパフォーマンスも維持されており、その点では後者の学習法の方が教育的効果は高いと言える。
- (2) 学習中の脳血流量は、機械翻訳のモデル文提示型より教師によるフィードバック型の方が概して高い傾向がある。このことは、冒頭で挙げた湯舟 (2011) や大石・木下 (2008) の研究の結果を基に考えると、機械翻訳のモデル文提示型のライティング学習を行う学習者は、モデル文中に特異な単語や表現の存在を見出すものの、文法や統語といったものにあまり注意を払わず、一方、教師によるフィードバック型のライティング学習を行う学習者は、教師からのフィードバックを注意深く読み、また、文法処理や統語処理、すなわち文の生成に関わるとされる脳の領域により多くの負荷がかかっている (使用している) ことを示唆していると言える。

(3) 機械翻訳のモデル文からは、特に語彙に関して、かなりの教育的効果が期待できると言える。参加者 A 群の多くが、機械翻訳中に使われていた幾つかの表現を知らなかったとコメントしていることから、おそらくこれらの表現は、日本の中学生や高校生が一般的に学ぶ表現ではなかったと推測される。従来は、いわゆる学校英語以外の英語に触れる機会はかなり限定されていたが、機械翻訳の出現によって、インターネット上の、英語のネイティブスピーカー達が普段使用する、しかし日本人にはあまり馴染みのない英語使用に触れることが可能となった。すなわち機械翻訳は、日本人が習得し得る語彙量を飛躍的に増やす可能性を秘めていると言える。しかし、2 回目のライティングでパフォーマンスを上げた参加者 B 群とパフォーマンスを下げた参加者 A 群間に脳血流量の差異があることから推測できるように、単にモデル文中にその存在を見出すだけで、「文生成」に関わる要素に注意を払わないのでは定着には不十分であると言える。

本研究を通じて、英語ライティング学習において、機械翻訳のモデル文提示型と、従来の教師によるフィードバック型のそれぞれの学習法の効用を一定程度示し得たと考えられるが、本研究には課題も多い。まず参加者数が少ない点が挙げられる。そもそも脳血流量といった現象は、個人的要因（気が散りやすい、神経過敏等）の影響を受けることも多いため、一般化するのは難しい部分もある。一般化を目指すのであれば相当数のサンプルが必要となるだろう。英語習熟度の問題もある。本研究では日本の大学生としては比較的英語習熟度の高い学習者を対象としたが、もっと習熟度が高かったり、低かったりする学習者の場合もはたして同様の事が言えるだろうか。例えば英語習熟度がもっと高い学習者であれば、機械翻訳という「お手本」を見るだけでも、長期的なライティング・パフォーマンスの向上に繋がる可能性も否定しがたい。また、今回は文法処理や統語処理を司るブローカ野のみにフォーカスしたが、機械翻訳を利用した参加者 A 群から、機械翻訳中の意外な単語や表現に注意が向いたというコメントが多数聞かれたため、次回は語彙を司る左角回や左縁上回といった脳の部位にも焦点を当てる必要がある。いずれにせよ、この種の研究はまだ緒に就いたばかりであり、上記のことについてさらに考察を深めていくことで、ライティング学習における機械翻訳のモデル文提示の教育的効果の全体像が掴めていくものと確信している。

### 謝辞

本研究は JSPS 科研費 JP24K04039 の助成を受けたものです。本研究の遂行にあたり、岡村優希さん（早稲田大学教育学部 4 年生 [当時]）、高山祥梧さん（早稲田大学教育学部 4 年生 [当時]）にはデータ収集にご協力頂きました。また、早稲田大学教育・総合科学学術院の伊藤悦朗教授には、早稲田大学先端生命医科学センターに設置されている NIRS の使用を快くご許可頂き、また、NIRS を利用したデータ収集や処理に関する貴重なご助言を賜りました。また、匿名の査読者の先生方からは論文執筆の上で貴重なアドバイスを頂きました。この場を借りて心より御礼申し上げます。

### 引用文献

- 岩中貴裕 (2023). 「英語教育における人工知能 (AI) の活用—その可能性と大学における英語教育の将来についての考察—」 『山口県立大学学術情報』 第 16 号, 109-119.
- 大石晴美・木下徹 (2008). 「第一言語処理と第二言語処理における脳活性状態の違い—日本語と英語のリスニングにおいて—」 『ことばの科学』 第 21 巻, 143-154.

- 木村修平 (2024). 「AI は英語教育を不要にするのか?—プロジェクト発信型英語プログラムの実践から考える—」 『立命館高等教育研究』 第 24 号, 1–16.
- 佐藤眞理子 (2023). 「機械翻訳と共存する英語教育—抵抗感と学習意義の変容についての分析—」 『リメディアル教育』 第 18 号, 81–90.
- 田辺友耶 (2023). 「まとまりのある英文を書く力の育成を目指す授業づくり—教科書本文をモデル文として活用する段階的かつ継続的に書く活動（自己表現活動）を通して—」 『神奈川県立総合教育センター長期研究員研究報告 21』 第 21 集, 1–6.
- 遠山昂希 (2017). 「気づきと第二言語学習に対する異なるタイミングでのモデル文提示の効果」 『中部地区英語教育学会紀要』 第 46 巻, 63–70. [https://doi.org/10.20713/celes.46.0\\_63](https://doi.org/10.20713/celes.46.0_63)
- 飛田ルミ・湯舟英一 (2014). 「近赤外光イメージング装置 (NIRS) による効果的な英語コミュニケーション指導法の検証」 『足利工業大学総合研究センター年報』 第 15 号, 123–129.
- 西尾直美・猪井新一 (2012). 「中学生のライティング学習中における過去形の正確性を高めるための修正フィードバックの効果と学習者の習熟度との関係性について」 『茨城大学教育実践研究』 第 31 号, 191–201.
- 藤本未来・田浦秀幸 (2011). 「第 2 言語習得開始年齢が言語流暢性課題に及ぼす影響: fNIRS 脳イメージング手法によるバイリンガリティー研究」 『言語科学研究』 第 1 巻, 55–90.
- 望月正道 (2024). 「ライティングにおける非明示的な直接的修正フィードバックの効果」 『麗澤レビュー』 第 30 巻, 4–12.
- 山下美朋・山中司・杉山滉平 (2024). 「AI ツールを活かした英語ライティング授業—英文作成支援ツール Transable を導入して—」 『立命館高等教育研究』 第 24 号, 75–87.
- 山中司 (2019). 「大学にもう英語教育はいらない—自身の「否定」と「乗り越え」が求められる英語教育者へのささやかなる警鐘—」 『立命館人間科学研究』 第 38 号, 73–89.
- 湯舟英一 (2011). 「英文速読におけるチャンクとワーキングメモリの役割」 『Dialogue : TALK 紀要』 第 9 巻, 1–20.
- 横野成美 (2023). 「AI 通訳・翻訳の使用が学生の英語学習意欲に与える影響」 『星稜論苑』 第 52 号, 15–26.
- 米村大輔 (2018). 「L2 ライティングにおける修正フィードバックの研究: 日本人学習者の正確性に与える影響に焦点をあてて」 『中部地区英語教育学会紀要』 第 47 巻, 289–296.
- 米村大輔 (2020). 「L2 ライティングにおける日本人学習者の正確性向上の特徴—2 つの CF タイプに焦点をあてて」 『中部地区英語教育学会紀要』 第 49 巻, 165–172.
- Abbuhl, R. (2011). Using models in writing instruction: A comparison with native and nonnative speakers of English. *Sage Journals*, 1–12. <https://doi.org/10.1177/2158244011426295>
- Almelhes, S. A. (2023). A review of artificial intelligence adoption in second-language learning. *Theory and Practice in Language Studies*, 13(5), 1259–1269. <https://doi.org/10.17507/tpsl.1305.21>
- Ayyash, A. A., & Khalaf, R.A. (2016). The effect of reading-to-write approach on EFL undergraduate essay writing. *Arab World English Journal*, 7(3), 256–270.
- Betal, A. K. (2023). Enhancing second language acquisition through artificial intelligence (AI): Current insights and future directions. *Journal for Research Scholars and Professionals of English Language Teaching*, 7(39), 1–8. <https://doi.org/10.54850/jrspelt.7.39.003>

- Choukaier, D. (2024). Integrating AI in English language pedagogy: Innovations and outcomes in teaching English as second/foreign language. *English Administration: Theory and Practice*, 30(5), 3811–3822. <https://doi.org/10.53555/kuey.v30i5.3538>
- Farrukh, F., Nazeer, H., Minhas, H. S., Naseer, N., & Noori, F. M. (2025). Assessing multilingual speakers' language processing through functional near-infrared spectroscopy (fNIRS). *Behavioural Brain Research*, 484, 1–12. <https://doi.org/10.1016/j.bbr.2025.115485>
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055-100061. <https://doi.org/10.1016/j.caeai.2022.100055>
- Godwin-Jones, R. (2024). Distributed agency in second language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5–31. <https://hdl.handle.net/10125/73570>
- Karataş, F., Abedi, F. Y., Gunyel, F. O., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29(15), 19343–19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1(1), 1–26.
- Kumar, O. (2023). Investigating the impact of artificial intelligence AI and technology in English language learning. *Advances in Social Behavior Research*, 3(1), 27–36. <https://doi.org/10.54254/2753-7102/3/2023026>
- Kristiawan, D. Y., Bashar, K., & Pradana, D. A. (2024). Artificial intelligence in English language learning: A systematic review of AI tools, applications, and pedagogical outcomes. *The Art of Teaching English as a Foreign Language*, 5(2), 207–218. <https://doi.org/10.36663/tatefl.v5i2.912>
- Lee, Y. J., & Davis, R. O. (2024). A case study of implementing generative AI in university's general English courses. *Contemporary Educational Technology*, 16(4), 533–540. <https://doi.org/10.30935/cedtech/15218>
- Schaefer, E. (2008). Rater bias patterns in an EFL writing assessment. *Language Testing*, 25(4), 463–492.
- Son, J., Ružić, N. K., & Philpott, A. (2025). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-assisted Language Learning*, 5(1), 94–112. <https://doi.org/10.1515/jccall-2023-0015>
- Syifauddin, M., & Yuliansyah, A. R. (2023). The effect of using AI on students' motivation and anxiety in learning English. *Transtool*, 2(2), –15. <https://doi.org/10.55047/transtool.v2i2.1354>
- Wei, L. (2023). Artificial Intelligence in language instruction: Impact on English learning achievement, L2 motivation, and self-regulated learning. *Frontiers in Psychology*, 14, 1–14. <https://doi.org/10.3389/fpsyg.2023.1261955>
- Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, 10(3), 1–13. <https://doi.org/10.1016/j.heliyon.2024.e25370>
- Ziwen, G., & Hongwei, Y. (2025). The impact of artificial intelligence chatbots on young EFL learners' willingness to communicate. *Curriculum and Teaching Methodology*, 8(1), 112–119. <https://doi.org/10.23977/curtm.2025.080116>

## JACET-KANTO JOURNAL SUBMISSION GUIDELINES

### A. Requirements

1. Contributors and co-authors must be JACET members at the time of submission, with at least one of them being a member of JACET Kanto Chapter.
2. A paper based on a lecture or workshop given at JACET Kanto Chapter events may be submitted as an Invited Paper (by invitation only).

### B. Editorial Policy

1. *JACET-KANTO Journal*, a refereed journal, encourages submission of the following:
  - full-length articles on topics of significance to all English teachers
  - research notes of a practical nature to share findings and insights
  - practical reports on English language teaching in classroom settings.
2. Manuscripts submitted to *JACET-KANTO Journal* must not have been previously published and should not be under consideration for publication elsewhere. Countermeasures may be taken if these conditions are breached.
3. During the peer review process, manuscripts submitted to *JACET-KANTO Journal* must not be withdrawn or used for publication elsewhere without any consultation with the committee.
4. Manuscripts which do not conform to the guidelines will not be considered for review.

### C. Submission Procedure

1. An original manuscript (Microsoft word file; .doc or .docx) should be submitted online.
2. All contributors must complete the Online Submission Form at the time of the submission.  
Refer to: <http://www.jacet-kanto.org/>
3. A manuscript should include title, abstract in English (no more than 200 words), and keywords (no more than 5 keywords).

### D. Guidelines

1. All manuscripts must be in English or Japanese.
2. All submissions to *JACET-KANTO Journal* should conform to the requirements of the *Publication Manual of the American Psychological Association*, 7th edition.
  - 2.1. Use the *JACET-KANTO Journal* Template. Download from <http://www.jacet-kanto.org/>
  - 2.2. Do not exceed 20 pages on A4 paper, including abstract, keywords, references, figures, tables, acknowledgments, and appendices.

2.3. Do not adjust the number of letters per line, lineage, or margins of any page.

#### **E. Copyright**

1. JACET holds the copyright of the articles published in *JACET-KANTO Journal*. Anyone, including the author(s), who wishes to reproduce an article, must obtain permission from JACET. If an article is being reproduced with permission, it should be clearly stated that JACET holds the copyright.
2. The institutional repositories are to use the versions as they are published in *JACET Journal*.

#### **日本語論文執筆要項**

1. *JACET-KANTO Journal* のテンプレート (<http://www.jacet-kanto.org/>) をダウンロードし、書式設定等は変更せずそのまま使用すること。
2. 論文タイトルは和文と英文を併記すること。
3. abstract は英語（200 語以内）で記載すること。
4. 論文は、A4 版横書きで abstract, keywords, 図表, 謝辞や文献を含めて、合計 20 枚以内にする。
5. その他の条件は英語の論文に準じること。

#### **F. *JACET-KANTO Journal* Vol. 14**

1. Deadline: 23:59 JST, August 31, 2026.
2. The result of the first round of review will be announced in November, 2026.
3. *JACET-KANTO Journal* Vol. 14 will be published on March 31, 2027.

JACET-KANTO JOURNAL

投稿規定

(1) 投稿種別 (Submission Category)

種別	内容
論文 Research Paper	<p>英語教育および関連領域について、</p> <ul style="list-style-type: none"> <li>● 先行研究に加えるべきオリジナリティある研究成果が、調査方法（量的・質的・混合法）とデータとともに明確に記述されている。</li> <li>● または、理論の構築やモデル化を目指すために関連領域が包括的にまとめられている。</li> <li>● または、他の研究のレプリケーションを行い、既存の研究に対して更なる既存の研究に対して更なる証拠が付与されている、あるいは反証されている。</li> </ul> <p>Research Papers include empirical and theoretical studies relevant to English language education and related fields. A paper must provide a comprehensive review or replication of current knowledge in a specific area. It must also represent original work that demonstrates a thorough understanding of the relevant literature and includes a detailed description of a theoretical framework, research methods (quantitative, qualitative, or mixed), data analyses, results and discussion. It may also offer new insights into or interpretations of relevant theory-constructing or model-making studies and should suggest future research directions.</p>
研究ノート Research Note	<p>英語教育および関連領域について、新しい事実の発見、萌芽的研究課題の提起、少数事例の提示など、将来の研究の基礎となる内容が明確に記述されている。</p> <p>Research Notes must demonstrate the same level of rigor as Research Papers, though they may reflect earlier stages of study, by presenting preliminary findings or focusing on aspects of larger findings. They should make a practical, useful, and plausible contribution to English language education</p>

	and/or related fields.
実践報告 Practical Report	<p>教育現場における指導実践の目的・意義・内容が具体的，かつ明示的に述べられている。また，当該実践が，他の教育現場での応用性・汎用性を持つものである。</p> <p>Practical Reports should include instructional methods, materials, and techniques for English language learning and teaching featuring clear objectives, significance, and detailed explanations. They must include a discussion of implications and/or applications for practice.</p>

(2) 審査評価項目 (Evaluation Criteria)

	評価項目	例示
a. 分野の適切性 Appropriateness of content and submission category	内容 Content	原稿の内容は本紀要で扱うものとして適切か。 The manuscript is appropriate for <i>JACET-KANTO Journal</i> readers.
	投稿種別 Submission category	投稿種別が適切か。 The manuscript is appropriate to the chosen submission category.
b. 記述の妥当性 Appropriateness of description and writing style	構成 Organization	構成が分かりやすく読み手にとって理解しやすいか。 The content of the manuscript is well organized and written in a way that is accessible to the <i>JACET-KANTO</i> readers.
	倫理的問題 Ethical issues	著作権・倫理的配慮において問題はないか（二重投稿，盗作などを含める）。 The manuscript has no copyright, plagiarism, or other ethical issues, and is not currently under consideration for publication elsewhere.
	言語 Language	表現は適切か。理解困難な表現はないか。 The manuscript uses appropriate and comprehensible language.
	形式とスタイル Formatting and style	投稿規程，テンプレート，および APA (American Psychological Association) スタイル

	ル（第7版）に基づいているか。 The manuscript observes all the <i>JACET-KANTO Journal</i> submission guidelines, including the journal template and APA style (7th edition).
c. 内容の信頼性 Credibility of content	内容に矛盾や誤りはないか。論理の展開に無理はないか。 The manuscript has no inconsistencies, inaccuracies, or illogical development in the content.
d. 独創性または新規性 Originality, creativity, novelty	従来にない新しい考え方，理論，実践，手段，事例のいずれかが，意義のある研究成果とともに示されているか。 The manuscript offers new insights, theory, practice, method, or case studies with significant research outcomes.
e. 教育的寄与 Educational contribution	示された成果が教育において有用か。教育効果向上が期待できるか。 The manuscript offers valuable pedagogical suggestions.
f. 将来的発展性 Future expandability and application	得られた知見・手法等が，教育・研究の新たな方向性を提案し，その実現に寄与する可能性があるか。 The manuscript is likely to suggest and enable new directions in research and education through new insights and methodology.
g. 完結性 Completeness	内容に一貫性があり，論文全体を通して，独立した1つの研究であると評価できるか。 The manuscript is a comprehensive study.

(3) *JACET-KANTO Journal* が扱う主な専門分野 (Areas of Interest)

1. Corpus
2. Curriculum (Pedagogy)
3. ESP/EAP
4. Grammar (Syntax)
5. ICT/CALL
6. Language Policy
7. Learner Development (Motivation, Strategies, Autonomy, Learner Belief)
8. Linguistics
9. Listening
10. Psycholinguistics
11. Reading
12. SLA (Child Language Acquisition, Bilingualism, Immersion, Interlanguage)
13. Sociolinguistics (Pragmatics)
14. Speaking (Phonetics, Phonology, Pronunciation, Conversation Analysis)
15. Teacher Education
16. Testing and Assessment
17. Vocabulary
18. Writing
19. Language Variation and Change (World Englishes, English as a Lingua Franca)

(4) スケジュール (諸事情により変更の可能性があります)

- 応募原稿の締切：2026年8月31日23時59分
- 一次査読の結果通知：2026年11月頃
- 発行：2027年3月31日

**【JACET 関東支部紀要第 13 号 編集委員会・査読者一覧】**

**JACET-KANTO Journal Volume 13**

**Editorial Board and Reviewers**

---

**EDITORIAL BOARD**

**CHIEF EDITOR**

SUZUKI, Kentaro / 鈴木 健太郎 (北海道教育大学)

**ASSOCIATE EDITOR**

NAKATAKE, Maiko / 中竹 真依子 (学習院大学)

**EDITORS**

INADA, Takako / 稲田 貴子 (日本保健医療大学)

KAMIMURA, Kozo / 神村 幸蔵 (筑波技術大学)

KOYA, Taeko / 小屋 多恵子 (法政大学)

OGISO, Tomoko / 小木曾 智子 (富山大学)

**REVIEWERS**

ASAOKA, Chitose / 浅岡 千利世 (獨協大学)

COLLINS, Peter J. / コリンズ ピーター J (東海大学)

FUJIEDA, Yutaka / 藤枝 豊 (大阪経済大学)

ITO, Yasuko / 伊藤 泰子 (神田外語大学)

IZUMI, Shinichi / 和泉 伸一 (上智大学)

MORI, Yoshinobu / 森 好紳 (白鷗大学)

NISHIKAWA, Megumi / 西川 恵 (東海大学)

SHIMOYAMA, Yukinari / 下山 幸成 (東洋学園大学)

SUGITA, Yoshihito / 杉田 由仁 (明治学院大学)

YAMAGUCHI, Takane / 山口 高領 (秀明大学)

## 編集後記

今年度は計5本の応募があり、10名の査読委員による厳正な審査を経て、研究論文2本、研究ノート2本の計4本の原稿を掲載することとなりました。査読者の皆さまには、短い期間の中で多くの的確かつ建設的なコメントをいただきました。この場を借りて、心より御礼申し上げます。

複数回の査読と修正を経て原稿がよりよいものへと磨き上げられていく過程は、本誌の質を支える重要な営みであると改めて感じております。また、幅広いテーマや多様な研究手法が扱われており、英語教育という分野の広さと奥深さを改めて実感いたしました。

今年度からは、中竹先生（学習院大学）に副委員長をご担当いただき、頼りない委員長を力強く支えていただきました。編集委員による校正作業においても、表現や形式面を含めて丁寧に確認を重ね、読者にとってわかりやすく、読みやすい誌面づくりを心がけてまいりました。

2026年度発行予定の第14号につきましても、多くの皆さまからのご投稿を心よりお待ちしております。

紀要編集委員長 鈴木健太郎（北海道教育大学）

2026年3月31日

JACET 関東支部・紀要編集委員会

委員長： 鈴木 健太郎 / 副委員長： 中竹 真依子

委員： 稲田 貴子, 小木曾 智子  
神村 幸蔵, 小屋 多恵子

一般社団法人大学英語教育学会 (JACET)  
2025 年度 関東支部紀要 第 13 号  
(*JACET-KANTO Journal* Vol. 13)

---

発行日 2026 年 3 月 31 日  
発行者 一般社団法人大学英語教育学会  
関東支部 (紀要編集委員会)  
〒162-0831 東京都新宿区横寺町 55  
TEL. (03)3268-9686

---

ISSN 2436-1993