

JALT2017 • LANGUAGE TEACHING IN A GLOBAL AGE

NOVEMBER 17-20, 2017 • IBARAKI, JAPAN

Impact of Writing Models on Accuracy, Complexity, and Overall Performance

John Peloghitis International Christian University

Reference Data:

Peloghitis, J. (2018). The impact of writing models on accuracy, complexity, and performance. In P. Clements, A. Krause, & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world.* Tokyo: JALT.

Model texts are effective tools in the writing process, particularly when implemented in the feedback cycle. This paper is a report on a study that investigated the effect model texts have in developing accuracy, complexity, and overall performance in essay writing. 38 lower intermediate Japanese learners of English were divided into a control group, which primarily received peer feedback, and an experimental group, which incorporated model essays and genre awareness activities into the feedback process. The study spanned 15 weeks and included 4 essays accompanied by in-class feedback sessions. Independent sample *t* tests and paired sample *t* tests were used to assess the effect of the model essays on accuracy, complexity, and overall performance between pretest scores and posttest scores. Results indicate that both groups made progress but the experimental group made greater improvement in their overall writing performance. The findings suggest that model texts should play a greater role in writing instruction.

"Model text" (ライティング見本) はライティングの過程、特に定期的フィードバックの効果的な手段である。本論は、小論文 を書く上で"Model text"が文章の正確さ及び複雑さ、そして総合的な評価においてもたらす効果を調べたものである。38人の 中級以下の英語学習者??を、主にビア・フィードバックを行う制御グループと、Model essay (見本となる文) とジャンル意識活 動をフィードバック過程に取り入れた実験グループに分けた。15週間、フィードバックセッションを受けた4名のレポート (小論 文)を研究対象とした。初めと最後のテストスコアで"Model text"が文章の正確さおよび複雑さそして総合評価における効果 を評価するにあたりインディペンデントサンプルTテストとペアサンプルTテストを用いた。その結果、両グループともにスコア の伸びが見受けられたが、実験グループは総合評価がさらに向上していることが判明した。従って、"Model text"はライティン グ指導において取り入れるべき意義深い手段であると示唆している。 The lack of exposure in learning academic writing in Japanese high schools has left students ill-prepared to meet the challenges of university writing programs (MEXT, 2016). Another obstacle is the overemphasis on grammar translation methods, which has led to the common belief that proficiency in writing is largely reached by producing accurate and complex sentence structures. To be successful, however, learners need to understand how to organize texts and meet the specific demands that a written genre requires (Kern, 2000).

An important element in introducing 1st-year university students to academic writing is to familiarize them with the multiple-draft feedback cycle. There are several types of approaches for giving feedback, such as teacher's essay correction, reformulation, and peer feedback. One approach that is often overlooked is the use of model texts and promotion of genre-awareness (Qi & Lapkin, 2001). Some scholars have advocated this approach and even claim that it can be more beneficial than teacher error correction and peer editing (Hanaoka, 2007; Qi & Lapkin 2001). It is necessary for writing instructors to tie the formal and functional properties of a language together to facilitate learners' recognition of how and why linguistic conventions are employed for particular rhetorical effects (Bhatia, 2004). If learners have opportunities to analyze the rhetorical structure of content, common patterns can be identified in each genre.

Review of the Literature

The use of model texts has its origins in the genre-based approach, the main goal of of which is to cultivate an awareness of the structure and purpose of different text types so that students can eventually replicate them in their own writing. The role of the writing instructor is to generate genre awareness through explanation and activities that assist students in noticing and analyzing key rhetorical features of the genre being learned (Abbuhl, 2011). Bagheri and Zare (2009) maintained that writing exercises to help students manipulate relevant language forms are vital in raising awareness in how grammatical features are used in authentic discourse contexts.



Two characteristics of genre-based writing activities are particularly salient: the concept of modeling and the explicit instruction of the genre. Modeling typically consists of three stages: the deconstruction of texts, the joint construction of texts, and independent construction of texts. The concept of modeling is based on the idea that successful language learning requires social interaction, shared context, and scaffolding (Derewianka & Jones, 2013). Hyland (2004) claimed that the modeling stage and explicit instruction helps students to explore the genre and understand its rhetorical structures or frames and formulaic sequences.

Some research studies have found that L2 writers utilize models (Angelova & Riazantseva, 1999; Tardy, 2006) and that L2 writers use them to address their lexicogrammatical errors during the feedback process (Qi & Lapkin, 2001). However, only Henry and Roseberry's study (1998) compared the use of models with other instructional techniques. The experimental study used tourist information texts to examine the impact of explicit instruction in the modeling stage within a genre-based approach. Participants were divided into two groups: one provided with explicit instruction of the model texts and another that was not. The genre group read several examples of the target texts, analyzed the organizational and lexicogrammatical features, and revised flawed versions of the model texts. In contrast, the nongenre group completed traditional grammar exercises relevant to the written task. After 3 weeks, differences in the gain in scores on pretests and posttests revealed that the genre-based groups made considerable improvement in the organization, cohesion, and persuasiveness of their texts. In contrast, the nongenre group's scores did not significantly improve. Henry and Roseberry's study (1998) reaffirmed the claim that explicit instruction is an essential supplement when models are introduced in the writing process.

A qualitative study by Macbeth (2010) found that the modeling stage provides lower proficiency L2 writers with the support needed to generate salient features commonly found in writing such as thesis statements, topic sentences, and supporting sentences. Macbeth claimed that models facilitated instruction and offered learners a visible roadmap on how to write their essays. Lastly, Ferretti, Andrew-Weckerly, and Lewis (2007) examined the impact of the genre-based approach on learners writing argumentative texts with the participation of six 1st-year ESL students in a North American university. Despite a short 2-week period of instruction, the findings showed that participants improved their argumentative writing at the level of organization, argumentation, and thesis construction. Models have been found to be a valuable resource for students with lower levels of proficiency because they offer support for writers to notice the "discrete elements of language" (Abbulh, 2011, p. 2). Moreover, model texts also allow students to become more flexible in their thinking and eventually realize how authors organize their writing. Nevertheless, the number of studies investigating models is surprisingly small. Little research is available on the specific effects of models on the writing process or on the effects of various kinds of models. The general purpose of this study was to examine how effective using a model-based approach to writing is compared with a more common approach that incorporates peer editing and teacher feedback. The study specifically targeted accuracy and complexity because these areas are generally perceived by students to be instrumental in gaining writing proficiency and often a major concern in the feedback cycle.

Research Questions

- RQ1. Does the use of model texts and genre-awareness activities have a greater impact on accuracy in writing than peer editing and teacher feedback?
- RQ2. Does the use of model texts and genre-awareness activities have a greater impact on complexity in writing than peer editing and teacher feedback?
- RQ3. Does the use of model texts and genre-awareness activities have a greater impact on overall performance in writing than peer editing and teacher feedback?

Methodology

Participants and Instructional Context

The participants in this study were 38 first-year Japanese students enrolled in the second semester of a yearlong compulsory reading and writing course. All of the participants were Liberal Arts majors attending a private university located outside Tokyo, Japan. The participants placed in the lower intermediate to intermediate range based on their scores from the Computerized Assessment System for English Communication (CASEC), the university's placement test administered to all 1st-year students shortly before the start of the first term. The participants all reported little or no experience in writing academic texts prior to entering university. The primary goal of the reading and writing course is to develop reading fluency and to familiarize students with a process-approach and basic academic writing.

Participants were from two writing classes taught by the same instructor. Both classes met for two consecutive 90-minute classes once a week over a 15-week span. The control group consisted of 18 students and the experimental group had 20 students. Both groups were required to complete four multiple-draft writing assignments. The



writing assignments were all argumentative in nature but in different formats such as description, compare and contrast, and stating a preference. Students were encouraged to write more than 300 words and use a five-paragraph organizational structure. The feedback cycle was managed by the instructor, and two English language instructors at the same university evaluated the essays independently to ensure objectivity and check reliability. The raters both had more than 8 years of experience in writing instruction and in using the rubric. Written consent was obtained from each participant at the outset of the study.

Research Design and Procedure

In order to investigate the research questions, a control group–experimental group design was adopted. The control and experimental groups were separated by class, thus the grouping was based on convenience. Independent samples *t* tests were employed on the pretest data to check for differences between the two groups at the onset of the study. The same statistical analysis was used to examine if differences existed between the groups in the posttest. Changes in accuracy, complexity and overall writing performance were measured using paired-samples *t* tests on scores collected from the pretest and posttest.

Training Sessions and the Feedback Cycle

Three 90-minute training sessions were provided to familiarize the participants with the composing and revision process. At the outset of the training session, both groups were introduced to the general features of academic/argumentative essay writing. However, the activities and objectives of the training sessions varied considerably between the groups. The control group learned the basic structure of a five-paragraph essay and the function of thesis statements, topic sentences, and supporting statements. Students reviewed a sample essay and engaged in composition exercises such as creating outlines with thesis statements and main ideas and later drafting a paragraph. Students met in a computer room for the last part of the training session where they reviewed formatting guidelines (i.e., indentation) and the rubric chosen for assessing the essays. Lastly, they conducted a peer-review and revision activity. In contrast, the experimental group was introduced to several model essays and completed a series of activities aimed at deconstructing the texts and providing metalanguage. Students examined rhetorical and organizational elements and their functions by engaging in collaborative exercises that manipulated the structures and style used in the model texts such as reorganizing strips of paper with sentences on them to create a new model text and a labeling

exercise that examined the basic structure of an argument. Students then manipulated a model essay to create a slightly revised version on a different topic. Students in the experimental group also met in a computer room for their last training session to work on formatting guidelines followed by a short practice activity where students needed to find ways to improve a sample text. Similar activities were used throughout the course to teach writing to each group. For example, the control group conducted composition exercises in preparing for the second essay (compare & contrast; see Appendix for details). Activities included practicing supporting details, signaling comparison and contrast, and using subordination effectively. For introducing the second essay, the experimental group deconstructed and analyzed discourse in two model compare and contrast essays, then jointly constructed a similar version of the text.

The first drafts for each essay were assigned as homework and submitted after a feedback session in a computer room; the control group engaged in peer editing whereas the experimental group reviewed models and collaborated on searching for structures that they could incorporate into their papers. These drafts were returned the following week with teacher comments and a holistic score. Although the comments targeted specific problems in the essays and included suggestions for improvement, the nature of the comments varied between the groups. The control group received indirect corrective feedback on grammar and vocabulary as well as comments that addressed problems with organization, development, and coherence. However, comments for the experimental group mostly referred back to the model text to confirm if their draft incorporated all the relevant organizational, rhetorical, and lexicogrammatical features. The second draft was then completed in a computer room after the feedback for the first drafts was given. The same procedure was followed for all writing tasks.

Instruments and Analysis

A pretest writing task was administered in the second week of the course to check if both groups had similar proficiency levels. The posttest was administered in class in week 14. Students were given a full class period to complete the essay. Accuracy was measured by calculating the number of the grammatical errors per total number of T-units. Form errors included plurals, sentence structure problems, tenses, prepositions, and comparatives and superlatives. Lexicon errors in lexis (word choice) were included only when a word obscured meaning. Complexity was measured by calculating the proportion of clauses to T-units. This analysis has been shown to be a reliable measure, correlating well with other measures of complexity (Foster & Skehan, 1996). Two raters scored the pretest and posttest holistically on a scale of 0 to 5 (with 0 representing poor



performance and 5 excellent performance) based on the TOEFL independent writing task rubric. All of the topics assigned for the pretest, the four writing assignments, and the posttest were chosen from a list of test prompts from previous TOEFL tests. The topics for all the tests and writing assignments can be seen in the Appendix.

The alpha for achieving statistical significance was set at .05. Additionally, effect sizes using Cohen's *d* were calculated on the *t* tests to evaluate the stability and strength of significance. A value of .2 is generally considered a small size effect, .5 a moderate size effect, and .8 a large effect. Because two raters were used throughout the study, interrater reliability was measured. The overall computed Pearson correlation coefficient was significant for the pretest (r = .763, p < .01) and the posttest (r = .788, p < .01), which indicates that a significantly strong relationship was found between the scores assigned by the raters on both essays. The data sets were all at acceptable levels of skewness and kurtosis to ensure normality of the distribution of data.

Results

Descriptive statistics and independent sample *t* tests for the pretest and posttest can be seen in the data presented in Table 1 and Table 2. The data indicates that both groups had similar mean scores in all three areas being measured on the pretest. Similar means were seen in the posttest; however, there was a much greater difference in the holistic scores on the posttest between the two groups. It is important to clarify that the descriptive statistics show lower scores in accuracy on the posttest in both groups, which indicates improvement because the ratio of errors found in sentences decreased over time.

Table 1. Results of Independent T tests and Descriptive Statistics on the Pretest by Group

Pretest	Group						95% Cl		
	Control			Experimental			for M		
	М	SD	п	М	SD	п	difference	t	df
Accuracy	1.41	0.15	18	1.43	0.13	20	-0.18, 0.07	-0.53	36
Complexity	1.45	0.15	18	1.35	0.18	20	-0.01, 0.21	1.81	36
Holistic	1.97	0.36	18	2.05	0.35	20	-0.31, 0.16	-0.66	36

* *p* < .05

Table 2. Results of Independent *T* tests and Descriptive Statistics on the Posttest by Group

Pretest			Gr	oup		95% Cl			
	Control			Experimental			for M		
	М	SD	п	М	SD	п	difference	t	df
Accuracy	1.20	0.19	18	1.24	0.19	20	-0.17, 0.08	-0.71	36
Complexity	1.55	0.23	18	1.53	0.19	20	-0.11, 0.16	0.37	36
Holistic	2.44	0.48	18	2.65	0.49	20	-0.52, 0.11	1.30*	36

* p < .05

The independent samples *t* tests were performed to check if significant differences exist in accuracy, complexity, and overall writing performance in the pretest scores between the two groups. The results for the pretest can be seen in Table 1, which illustrates there were no significant differences in accuracy (t(36) = -0.53, p = .59), complexity (t(36) = -1.81, p = .078), and overall writing performance (t(36) = -0.66, p = .51) between the two groups. Independent *t* tests were performed to measure if there were significant changes between the two groups in the posttest, which is shown in Table 2. Although the independent *t* tests on the posttest did not indicate any significant differences in complexity (t(36) = 0.37, p = .71) and accuracy (t(36) = -0.71, p = .48), the experimental group did score significantly better than the control group in overall writing performance (t(36) = -1.30, p = .03) and a Cohen's *d* calculation (r = .43) revealed a moderate effect size.

Paired-samples *t* tests were performed to examine if significance was found in the scores between pretest and posttest in each group. Table 3 reveals the data from the paired *t* tests for the control group. Significant differences are found in accuracy (t(17) = 5.14, p = .00) and overall writing performance (t(17) = -5.00, p = 0.00) between the two writing tasks. The effect sizes for accuracy (r = .77) and holistic scores (r = .78) are high. No statistically significant differences were found in complexity, t(17) = -2.03, p = .06.



Table 3. Paired-Samples *T* test on the Control Group's Pretest and Posttest Scores

	М	SD	SE M	95% Cl for <i>M</i> difference	t	df	Sig (2-tailed)
Pair of accuracy scores	0.20	0.17	0.04	0.12, 0.29	5.14	17	.000*
Pair of complexity scores	-0.11	0.22	0.05	-0.21, 0.00	-2.03	17	.059
Pair of holistic scores	-0.47	0.40	0.09	-0.67, -0.27	-4.99	17	.000*
*							

* *p* < .05

For the experimental group, significant differences in all three measurements were realized between the pretest and posttest scores in accuracy (t(19) = 5.18, p = .00), complexity (t(19) = -3.99, p = 0.01), and overall performance (t(19) = -8.30, p = .00). Cohen's *d* reveals a large size effect on holistic scores (r = .81) and accuracy (r = .76) and a moderate effect size for complexity (r = .68).

Table 4. Paired-Samples T test on the Experimental Group's Pretestand Posttest Scores

	14	an	07.14			10	<i>a</i> :
	M	SD	SEM	95% CI for	t	df	Sig
				<i>M</i> difference			(2-tailed)
Pair of accuracy scores	0.18	0.16	0.03	0.11, 0.26	5.17	19	.000*
Pair of complexity scores	-0.18	0.20	0.04	-0.27, -0.08	-3.99	19	.000*
Pair of holistic scores	-0.60	0.45	0.10	-0.81, -0.39	-6.00	19	.001*
* 05							

* p < .05

Discussion

The present study examined the effect of incorporating model texts and activities that promote genre awareness into the feedback process. The research questions explored three important elements of writing proficiency: accuracy, complexity, and overall

performance. Both groups improved significantly over the semester, but two findings seem particularly relevant. First, the independent sample *t* test on the posttest shows that the experimental group had significantly higher holistic scores than the control group. One explanation is that more time was given for the experimental group to reflect on composing their texts. Gibbons (2009) maintained that reflection can make language more visible and accessible and can help learners to internalize the skills and processes needed to successfully complete their essays. Implementing metacognitive exercises in the writing process is also an important part in creating autonomous learning. In principle, process-based approaches do encourage metacognitive strategy use but more conscious-raising activities can be integrated into the process. The descriptive statistics do indicate that both groups made gains in complexity, accuracy, and overall performance over the course but this result is not surprising because all the participants wrote numerous drafts and revisions during that period.

Second, the paired-samples *t* tests indicated that both groups saw significant gains in accuracy and overall performance; however, only the experimental group made significant gains in complexity. For the control group, improvements in accuracy are likely due to the large amount of feedback and editing exercises that were completed. For instance, many of the revisions from the teacher and from peers targeted grammatical structures such as verb tenses and word forms on the second drafts. For the experimental group, collaboration during the revision process likely contributed to the gains in accuracy because learners worked together in reexamining the models in searching for alternative grammatical patterns to improve their essays. Gains in overall performance in both groups can be attributed to many different factors. The amount of writing completed over the course is certainly a chief factor. The materials and lessons were carefully scaffolded and the feedback process combined various types and sources of feedback that support the recursive nature of revision. Complexity was the only area in which the control group did not reach significance. Content targeting the use of subordination, coordination, and transitional phrases was integrated into the writing process to help the group write more sophisticated sentences. However, the results on the posttest revealed that students may have avoided making mistakes with these phrases and structures. Instead, more effort was placed on creating simple, clear, and grammatical sentences. The experimental group made considerable gains in complexity by completing exercises that focused on finding transitions and other rhetorical devices and examining their role in the sentence or paragraph. When students jointly constructed a text and independently constructed a text repeatedly throughout the course, they were asked to replicate these phrases and conventions. Many of these phrases and structures were memorized and replicated during the posttest. Because



the group did apply these patterns and structures to a new context, the models also appear to help students see patterns in paragraphs and how various rhetorical and lexicogrammatical features contribute to their development. These results suggest that model texts can serve as powerful instructional tools in the feedback process. When model texts are implemented properly in a series of stages, they offer a strong alternative to other forms of written feedback. They encourage developing metacognition and working collaboratively to establish a sense of authorship in another language.

Limitations and Conclusion

The most obvious limitation of this study is the small number of participants. More complex analyses could not be performed on the data. Furthermore, the scope of the study was confined to data collected in the pretest and posttest. Little knowledge can be gained in how these methods facilitated writing and revision in the essays completed over the course. A more robust investigation could have been achieved by including a qualitative analysis to examine classroom activities and revision practices. The same is true for the scoring of the essays. Due to the lack of analysis in examining the feedback process, it is not certain if feedback was the primary factor that contributed in the gains revealed from the posttest. However, the findings do suggest what other studies have claimed, that is, model essays and fostering awareness and noticing skills can be strong pedagogical tools in helping second language learners build proficiency in academic writing (Macbeth, 2010; Qi & Lapkin, 2001). Providing a modeling stage allows students to explore the genre and understand its rhetorical structures or frames and formulaic sequences. As teachers, more effort is needed to seek and develop instructional practices that utilize this valuable learning resource.

Bio Data

John Peloghitis currently resides in western Tokyo and is presently teaching as an instructor in the English Liberal Arts Program at International Christian University in Japan. He teaches academic reading and writing, debate, and research writing. He is an active member of Japan Association of Language Teachers and Japan Association of College English Teachers and is interested in second language writing, metacognitive strategies, syllabus design, and e-learning. <jpeloghitis@hotmail.com>

References

- Abbuhl, R. (2011). Using models in writing instruction: A comparison with native and non-native speakers of English. *Sage Open, 1*(3), 1-12. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/2158244011426295
- Angelova, M., & Riazantseva, A. (1999). "If you don't tell me, how can I know?": A case study of four international students learning to write the U.S. way. *Written Communication*, *16*, 491-525. https://doi.org/10.1177/0741088399016004004
- Bagheri, M. S., & Zare, M. (2009). The role of using IELTS model essays in improving learners' writing and their awareness of writing features. *JELS*, *1*(1), 115-130. Retrieved from http://jels. iauctb.ac.ir/article_510817_79051025a237cbbdefd5e9d625790a39.pdf
- Bhatia, V. K. (2004). Worlds of written discourse. New York, NY: Continuum.
- Derewianka, B., & Jones, P. (2013). *Teaching language in context*. Melbourne, Australia: Oxford University Press.
- Ferretti, R. P., Andrews-Weckerly, S., & Lewis, W. E. (2007). Improving the argumentative writing of students with learning disabilities: Descriptive and normative considerations. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 23*(3), 267-285. https://doi.org/10.1080/10573560701277740
- Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, *18*(3), 299-324. https://doi.org/10.1017/S0272263100015047
- Gibbons, P. (2009). *English learners academic literacy and thinking: Learning in the challenge zone.* Portsmouth, NH: Heinemann.
- Hanaoka, O. (2007). Output, noticing and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, *11*(4), 459-479. https://doi.org/10.1177/1362168807080963
- Henry, A., & Roseberry, R. L. (1998). An evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Quarterly*, *32*, 147-156. https://doi.org/10.2307/3587913
- Hyland, K. (2004). Genre and second language writing. Ann Arbor: University of Michigan Press.
- Kern, R. (2000). Literacy and language teaching. Oxford, England: Oxford University Press.
- Macbeth, K. (2010). Deliberate false provisions: The use and usefulness of models in learning academic writing. *Journal of Second Language Writing*, *19*, 33-48. https://doi.org/10.1016/j.jslw.2009.08.002
- MEXT. (2016). *Heisei 26 nenndo eigoryoku chousa kekka (koukou 3 nensei) no gaiyou (shousaiban)* [Detailed outline of the results for the 2015 survey of English ability (3rd year high school students)]. Retrieved from http://www.mext.go.jp/component/a_menu/education/detail/______icsFiles/afieldfile/2015/07/03/1358071_02.pdf



- Qi, D., & Lapkin, S. (2001). Exploring the role of noticing in a three-stage second language writing task. *Journal of Second Language Writing*, *10*(4), 277-303. https://doi.org/10.1016/S1060-3743(01)00046-7
- Tardy, C. M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of Second Language Writing*, *15*, 79-101. Retrieved from https://doi.org/10.1016/j.jslw.2006.04.003

Appendix The Prompts for Pretest, Posttests, and the Essay Assignments

Pretest Prompt

Do you agree or disagree with the following statement? People should not be allowed to use mobile phones on trains? Give reasons and examples to support your opinion. (Agree or Disagree)

Essay Topic 1

Some people think that people should save money for the future while other people think they should spend it on vacations and trips? Which do you think is more important? Give reasons and examples to support your choice. (Preference)

Essay Topic 2

Some young children spend a great amount of their time practicing sports. Discuss the advantages and disadvantages of this. Use specific reasons and examples to support your answer. (Compare & Contrast)

Essay Topic 3

If you were asked to send one thing representing your country to an international exhibition, what would you choose? Why? Use specific reasons and details to explain your choice. (Description)

Essay Topic 4

What is a very important skill a person should learn in order to be successful in the world today? Choose one skill and use specific reasons and examples to support your choice. (Description / Explanation)

Posttest Prompt

Do you agree with the following statement? Boys and girls should attend separate schools? Give specific reasons and examples to support your opinion. (Agree or Disagree)