L2 Writing Process, Proficiency, and Evaluation : Focusing on Keystroke Logging Approach

Kawaguchi, Yusaku

(Citation)
Learner Corpus Studies in Asia and the World,5:25-42

(Issue Date)
2020-12-21

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81012487

(URL)
https://hdl.handle.net/20.500.14094/81012487

Learner Corpus Studies in Asia and the World, vol. 5 (Papers from LCSAW2020) ISSN: 2435-2632
Published by School of Language & Communication, Kobe University, Japan

L2 Writing Process, Proficiency, and Evaluation —Focusing on Keystroke Logging Approach—

Yusaku KAWAGUCHI
Aichi Gakuin University

Abstract

The writing process, which seems to relate to writers' writing proficiency and quality of written products, has been one of the most important concepts in second language (L2) writing research. This article deals with the relationship between the L2 writing process and writing proficiency/evaluation, focusing on the keystroke logging approach. The paper reviews previous studies regarding the relationship between process and evaluation and summarizes the features of high-proficiency writers' writing process that these studies have pointed out. Also, it discusses the advantages of two kinds of method frequently employed in L2 writing process research: self-reported methods (i.e., verbal protocol, questionnaire) and behavioral methods (i.e., keystroke logging) used in these studies. It introduces an application for recording writers' writing process using the keystroke logging technique. Then, it presents studies on the relationship between writing process and proficiency/evaluation using that application and suggest the possibility of using the patterns of change in numbers of words over time. Through the discussion, some directions and implications for future writing process research are elucidated, including revisiting indices concerning the writing process. Finally, the paper discusses how to design a writing process corpus, considering task conditions referring to the previous attempt to build a writing process corpus.

Keywords

writing process, writing strategy, keystroke logging, statistical model, corpus building

1. Introduction

Many studies of second (L2) or foreign language (FL) writing have focused on writing process. This is because the writing process seems to be a factor that relates to writer's

proficiency and the quality of the written product. The writing process is also thought to provide a lot of information about the writer, such as revision activities (e.g., Stevenson et al., 2006), formulation activities (e.g., Roca de Larios et al., 2008), pausing behaviors (e.g., Spelman Miller et al., 2008), and strategy use (e.g., Sasaki, 2000; Yamanishi, 2009). We cannot obtain that information if we look only at the final product. How these behaviors relate to the evaluation of the final product or the proficiency of the writer is one of the most frequently discussed topics in writing process research.

The purpose of the present article is to examine the relationship between the writers' writing process and their proficiency and quality of the final product and the usefulness of using keystroke logging approach. First, we will review the studies on the writing process and its relation to writers' proficiency and the rating of essays. We will also discuss the benefits of using the keystroke logging technique for writing process research from the perspective on its convenience.

2. Background

2.1 Sub-processes and Evaluation of L2 Writing

Hayes and Flower (1980), one of the most frequently mentioned studies in L2 writing process research, suggested a writing process model including three sub-processes: (a) planning, (b) translating (or formulation), and (c) reviewing. Many studies on the writing process in L2 and FL have also focused on these three sub-processes.

Previous studies, such as Victori (1999), have emphasized the importance of metacognitive ability and strategic knowledge to demonstrate clearer sub-processes and successfully control the allocation of time for the sub-processes during tasks. Victori (1999) also pointed out that these abilities or knowledge relate to the writers' proficiency. Previous studies on L2 writing process have focused on the relationship between these abilities or strategic knowledge and L2 writing proficiency and writers of highly-rated essays (Hirose & Sasaki, 1994; Roca de Larios, et al., 2001; Roca de Larios et al., 2008; Sasaki, 2000; Sasaki & Hirose, 1996; Xu & Qi, 2017; Victori, 1999; Yamanishi, 2009).

The features of highly proficient writers' processes, which were pointed out in these previous studies, can be summarized in the following three points. First, they spend a lot of time on paragraph-level global planning before they start writing (Hirose & Sasaki, 1994; Sasaki, 2000; Sasaki & Hirose, 1996; Xu & Qi, 2017; Yamanishi, 2009). Second, once they plan what to write before they start writing, word-/sentence-level local planning strategy is not used frequently during the task (Sasaki, 2000; Yamanishi, 2009). Finally, therefore, they spend less time on formulating than low-proficiency

writers (e.g., Roca de Larios et al., 2001; Roca de Larios et al., 2008)

2.2 Methods to Collect Writing Process Data

We can roughly classify L2 writing process data into two categories: self-reported data and behavioral. Self-reported data refers to data collected by retrospective or self-evaluation methods, such as verbal protocol methods (e.g., think-aloud protocol, stimulated recall) and questionnaire. In the verbal protocol methods, participants say what they are thinking during a task or what they thought after the task. Also, the questionnaire asks participants to report and evaluate what strategies they use and how frequently or consciously. These methods are better to observe learners' cognitive activities that we cannot access if we observe just the learners' behavior. Many studies on the L2 writing process have employed this method to examine the writers' cognitive aspect (e.g., Sasaki, 2000; Roca de Larios et al., 2008). Moreover, questionnaires with ordinary scales (e.g., Yamanishi, 2009) are labor-saving because researchers do not need any coding procedures and they are compatible with quantitative analyses.

On the contrary, many studies on the L2 writing process have also employed behavioral data (e.g., Kawaguchi, in press; Kawaguchi et al., 2016; Xu & Qi, 2017). Behavioral data refers to observed, real-time behaviors during writing tasks (i.e. formulating, pausing, deleting, etc.), collected by video recording, computer screen recording, or key-stroke logging technologies. Researchers can obtain many kinds of writing process data, such as the change in numbers of words/sentences over time, revision history, and pausing duration/frequency/location. These behavioral data could be employed to reveal writers' writing process features.

2.3 Keystroke Logging Techniques in L2 Writing Process Research

In many kinds of methods to collect behavioral data, keystroke logging is the most convenient way to obtain precise data on writing activities. Keystroke logging data show the changes in numbers of words/sentences over time, history of formulation, revision, and deletion, and pausing behavior (duration, frequency, and location). Moreover, keystroke logging techniques do not require researchers coding process and have less risk of human errors than methods requiring transcribing, like verbal protocol methods. Also, the data collection procedure does not interfere with the writers' thinking and behaviors, which alleviates concerns over ecological validity issues. Inputlog (Leijten & Van Waes, 2013) is one of the most frequently used applications in L2 writing process research. Its recording module logs data (i.e., keystroke, mouse

operation, switching programs) in Microsoft Word with a timestamp. In addition, its analysis module provides various kinds of information such as process/product information, revision information (insertions/deletions/breaks), and pause information (number, duration, and location).

WritingMaetriX (WMX, Kusanagi et al., 2015), a keystroke logging application for recording writers' writing process, is used in the present study. The recording module collects all information regarding real-time writing processes in a writing task. It works with its own user interface, independently from other application such as Word. Every time a key is pressed, the pressed key, the input texts at the point, and the timestamp are recorded in the computer's memory, and all the information is outputted to a text file format with its own filename extension, *.klg* when the save button is pressed (Figure 1). It also outputs a copy of the final product as a separate file.

Figure 1

An example of WritingMaetriX's writing process data

```
*^  9:58:23:353^637300^ Nowad*↵
*^  9:58:23:456^637403^ Nowada*↵
*^  9:58:23:888^637835^ Nowaday*↵
*^  9:58:23:969^637916^ Nowadays*↵
*^  9:58:24:105^638052^ Nowadays*↵
*^  9:58:29:25^ 642972^ Nowadays*↵
*^  9:58:29:145^643092^ Nowadays T*↵
*^  9:58:29:369^643316^ Nowadays Te*↵
@^  9:58:30:537^644484^ Nowadays T@↵
@^  9:58:30:681^644628^ Nowadays@↵
*^  9:58:31:401^645348^ Nowadays t*↵
*^  9:58:31:505^645452^ Nowadays te*↵
*^  9:58:31:809^645756^ Nowadays tec*↵
*^  9:58:32:881^646828^ Nowadays tech*↵
*^  9:58:32:937^646884^ Nowadays techo*↵
@^  9:58:34:513^648460^ Nowadays tech@↵
*^  9:58:35:249^649196^ Nowadays techn*↵
*^  9:58:35:329^649276^ Nowadays techno*↵
*^  9:58:35:505^649452^ Nowadays technol*↵
*^  9:58:35:697^649644^ Nowadays technolo*↵
*^  9:58:36:185^650132^ Nowadays technolog*↵
*^  9:58:36:385^650332^ Nowadays technology*↵
*^  9:58:36:657^650604^ Nowadays technology*↵
```

Ishii et al. (2015) made an initial attempt to construct a writing process corpus using WMX's process data. They suggested that research methods employed in corpus linguistics (i.e., searching, sorting, automatic evaluation) and educational data mining methodologies are compatible with writing process data collected with keystroke

logging techniques. Ishii (2016) also suggested that time series analysis or sequential pattern mining can be applied to writing process research.

3. Research 1: Kawaguchi et al. (2016)

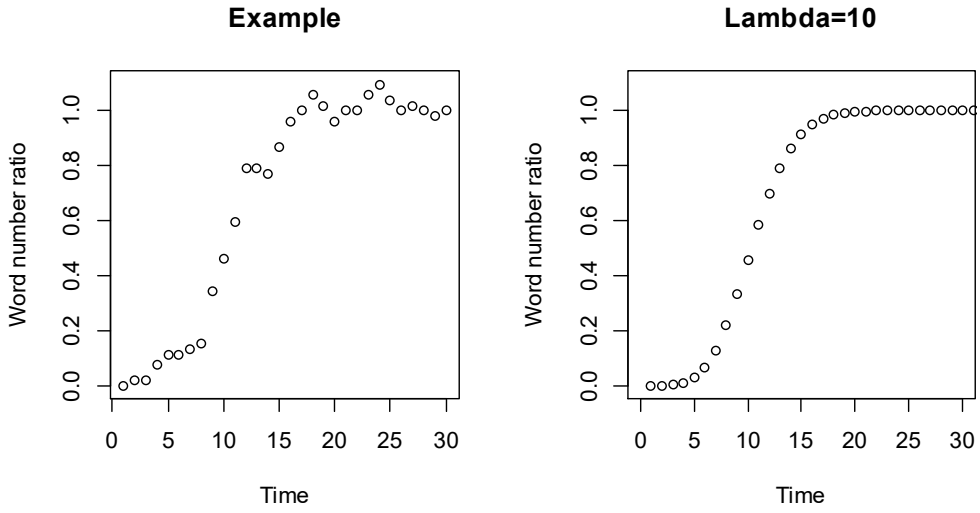
3.1 Background and Research Questions

Based on the features of highly proficient writers shown above in the background section, we made assumptions regarding how the number of words of highly proficient writers increases at each stage of a task. In the early interval of a task, the number of words does not increase heavily because the highly proficient writers do global planning regarding the overall organization of their essays. Then, in the middle interval, the number of words does not increase heavily, because the highly proficient writers formulate based on global planning they have done before they started writing. Finally, in the end, the number of words does not increase heavily, because the highly proficient writers revise and review the contents and expressions they have written so far.

From the above assumptions, Kawaguchi et al. (2016) created two statistical models of the change in numbers of words over time in a writing task: a Poisson distribution model and a linear regression model. The Poisson distribution model represents the writing process of highly proficient writers. When comparing the shapes of the cumulative distribution function of the Poisson distribution and the highly proficient writers' change in numbers of words over time, they are supposed to show a similar shape to each other (see Figure 2). The time series data of the increasing number of words are fit to the Poisson distribution by a test of goodness-of-fit, in which three indices will be computed: χ^2 value, which indicates the goodness of fit to the Poisson (the lower the number, the better the data fit to the model), and the parameter λ (lambda), which defines the shape of the distribution. If the value of λ is half of the task time (15 if the task time is 30 minutes), then the number of words has increased in the middle interval of the task.

Figure 2

An example of highly proficient writers' writing process and the cumulative distribution function of the Poisson (λ = 10)



Meanwhile, the linear regression model represents the writing process of less-proficient writers. That is the process that writers keep formulating from the beginning to the end of a writing task, without pauses in the initial and final intervals of the task. It means that there is little prewriting planning before formulation and little revision/reviewing after formulation. To fit the pattern of the change in the number of words into the linear regression model, a simple linear regression analysis is performed with the ratio (the number of words at each minute/ the final number of words) as the independent variable, and time (minutes) as the dependent variable. Then, the slope, intercept, and the coefficient of determination, R^2, which represents the goodness of fit, will be computed.

From the above discussion, the research questions are the following:

(a) Which fits the change in numbers of words over time better, the Poisson distribution model or the linear regression model?

(b) Do goodness-of-fit indices of these two models correlate to the rating of essay writing?

(c) Does λ of the Poisson distribution correlate to the rating of essay writing?

3.2. Method

Students at a Japanese national university who studied English in the required classes (N= 35) participated in the study, but one was excluded from analysis because of his/her unnatural patterns of changes in the number of words. The participants' mean

self-reported TOEFL ITP score was 487.32 (*SD* = 45.07), so they had intermediate proficiency of English in Japan. Almost half of them had received instruction in paragraph writing.

Each participant was engaged in a 30-min argumentative essay writing task on WMX on Windows PC. The topic was "Do you agree or disagree with the following statement? *Technology has made the world a better place to live.* Use specific reasons and examples to support your opinion," retrieved from a TOEFL writing section. During their tasks, they were not allowed to use copy and paste features in order to collect precise patterns of the number of words. All the participants finished their tasks within 30 minutes.

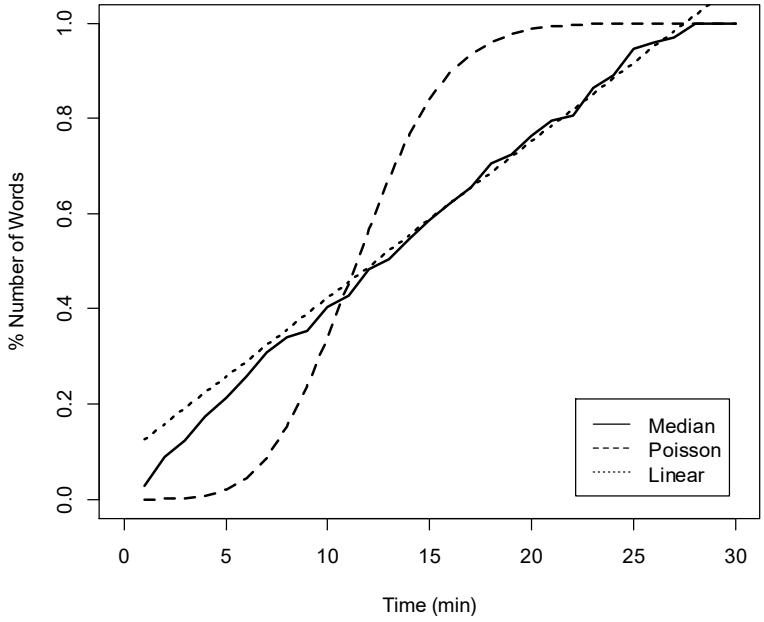
The analysis was conducted using R (R Core Team, 2019), a statistical computing environment. First, all the writing process data were transformed into the data of the change of the number of words over time. Then, they were fitted to the Poisson distribution model and the linear regression model to obtain parameters of each model as features of the writing process. Finally, a multivariate correlation analysis among all the variables was conducted to investigate the relationship between essays' ratings, the total number of words, lexical richness index (Guiraud Index, GI), and writing process features.

3.3 Results and Discussion

For the first RQ, let us look at Figure 3, which shows overall patterns of the changes in numbers of words of all the participants. The solid line shows the time series data for the participants' median of number of words at each minute; the dashed line shows the cumulative distribution function of the Poisson distribution at all participants' mean λ; and the dotted line shows all participants' median of the estimated number of words when fitted to the linear regression model. Comparing the forms of the three lines, most of the writers' patterns of change in the number of words over time fit the linear regression model better than the Poisson distribution model. This indicates that the participants of the study tended to keep formulating from the beginning to the end of the writing task, without pauses for prewriting planning or revising/reviewing.

Figure 3

Overall patterns of the changes in numbers of words: The actual data and the data based on all participants' parameters of two models



Also, for the second and third RQs, we conducted multivariate correlation analyses among the indices representing the participants' writing process and the ratings of the products (Table 1). This revealed that the coefficient of determination (R^2), as a goodness of fit index for linear-regression model, correlated to the ratings, the total number of words, and index of the lexical richness of the essays. It indicates that writers who kept formulating from the beginning to the end could write highly rated essays. This result is contrary to the prediction that highly proficient writers who write highly rated essays stop for prewriting planning in the initial interval and revision/reviewing in the final interval of the tasks.

Table 1

Correlation Matrix of All Variables

	λ	χ^2	Intercept	Slope	R^2	Words	GI	Rating
λ		.57	-.84	.28	.63	.27	.30	.40
χ^2	.74		-.25	-.21	.23	.48	.46	.43
Intercept	-.90	.53		-.69	-.86	-.42	-.39	-.49
Slope	-.38	-.17	-.56		.67	.46	.32	.41
R^2	-.35	.18	-.58	-.11		.53	.53	.52
Words	-.47	.62	-.29	.23	-.05		.74	.77
GI	-.04	.12	.03	-.03	.25	.30		.61
Rating	.15	-.11	.05	.00	.00	.52	.06	

Note. Upper diagonal values represent correlation coefficients, and lower diagonal values represent partial correlation coefficients.

There are some possible reasons why the participants who wrote highly rated essays in the study did not show the same patterns as the good writers in the previous studies. The first is the time allocated to the task was too short for the participants to do prewriting planning and revision/reviewing. Under the time restriction, they perhaps attempted to do formulation and planning/revision/reviewing simultaneously. Next, the familiarity with the topic was so high that the participants did not need so much time to plan the contents. These suggest that the task conditions need to be reconsidered.

4. Research 2: Kawaguchi (in press)

4.1 Background and Research Questions

As mentioned earlier, Sasaki and Hirose (1996) and Yamanishi (2009) used a questionnaire to capture the writers' preferences for writing strategies. The benefits of this method are that we can access the writers' thoughts and recognition before, during, and after the writing task, and it does not interfere with the writers' thoughts and behaviors during the task. Besides, quantitative questionnaires, such as those used in Yamanishi (2009), do not require transcribing or coding, and they are relatively easy to obtain data without any concerns for human errors during transcribing and disagreements among coders during coding processes.

On the other hand, the questionnaire cannot determine whether or not the writing

strategies that the writers responded that they use (or have used) are actually used in the task. In other words, the data obtained by the questionnaire are the writers' self-reports of their preference toward the writing strategies. Therefore, if the writers' proficiency is too low to correctly metacognize their own strategy-use, they will not be able to discriminate between the strategies that they used and the strategies that they indicated in the response. The writing strategies used by the respondents may not be consistent with the actual writing process. On the contrary, the more proficient writers with the high metacognitive ability and the ability to use writing strategies consciously, the more likely their self-reported strategy-use and their response to the questionnaire are consistent.

From the above discussion, the research questions are stated as follows:

(a) Do the writers' preferences scores of writing strategies correlate to their observed behaviors (i.e., writing process indices)?

(b) Are there any differences in the relationship between the two of them for high-and intermediate-proficiency writers' groups?

4.2 Method

The sample in the study (N = 215) consisted of a group of writers (n = 21) who participated in the experiment the author conducted and a group of writers (n = 194) from a corpus built with writing process data, named WritingMaetriX Corpus (Ishii et al., 2015).

The writers in those two groups had different proficiencies. The former group included relatively highly proficient undergraduate/graduate students who studied English as FL/L2 in Japanese universities. Their mean TOEIC score was 756.25 (SD = 178.18), and half of them had an experience of studying abroad and majored in the fields relating to English, such as applied linguistics. On the other hand, the latter group included Japanese undergraduate students. The mean TOEFL-ITP scores was 467.23 (SD = 31.82), so they were relatively less proficient writers than the former group: they were intermediate writers.

The conditions of the argumentative essay writing tasks were different from each other. First, task time was different. While the participants of the highly proficient group were given 30 minutes to finish the task, those of the intermediate proficient group could take 20-60 minutes to finish the task. Next, the topics of the task were different. The topics assigned for highly proficient group are (a) Do you agree or disagree with the following statement? Technology has made the world a better place to

live. Use specific reasons and examples to support your opinion, and (b) Do you agree or disagree with the following statement? Parents are the best teachers. Use specific reasons and examples to support your answer. Both topics were retrieved from a TOEFL writing section. On the other hand, the topics assigned for the intermediate group (i.e., data from WMX corpus) were "It is important for college students to have a part-time job," "Smoking should be completely banned at all the restaurants in the country," retrieved from the International Corpus Network of Asian Learners of English (ICNALE; Ishikawa, 2013), and "school education" retrieved from Nagoya Interlanguage Corpus of English (NICE; Sugiura, 2011).

To examine the participants' writing strategy preference, Yamanishi (2009)'s writing strategy questionnaire was employed. The participants answered the questionnaire. It consists of four sub-scale: global planning (k = 10), local planning (k = 7), revision/reviewing (k = 8), and avoidance strategies (k = 8).

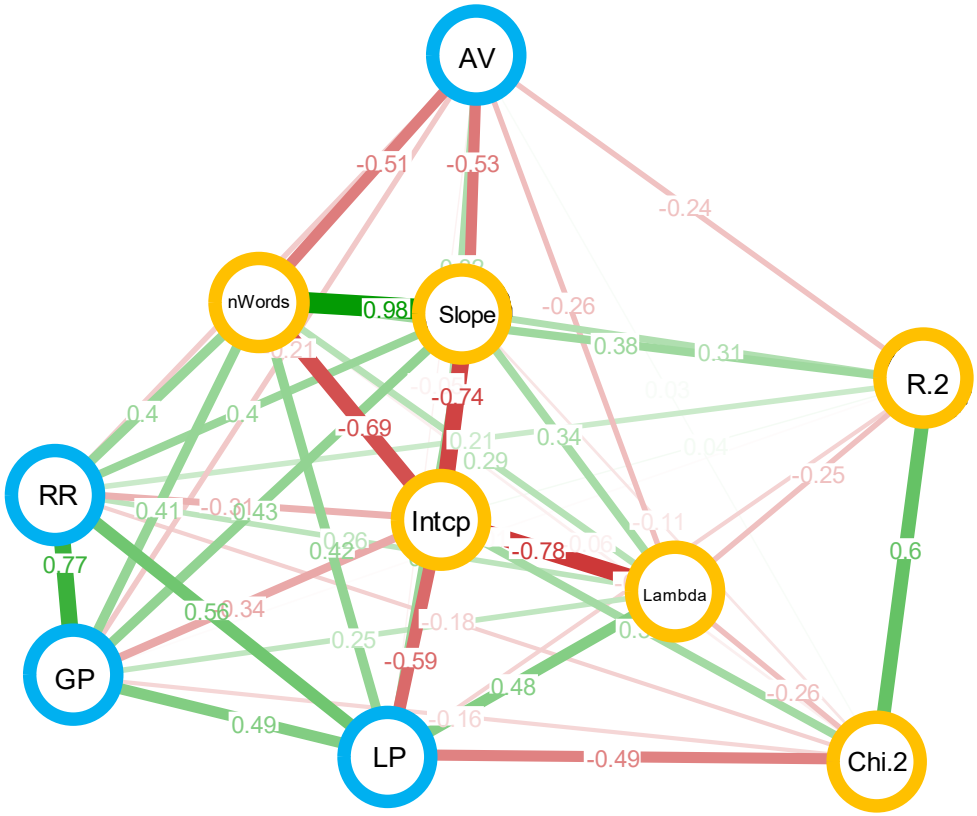
As well as Kawaguchi et al. (2016), all the writing process data were transformed into the data of the change of the number of words over time. Subsequently, the data were fitted to the Poisson model and the linear regression model, and we obtained five parameters (e.g., λ and χ^2 in the Poisson distribution model, and slope, intercept, and R^2 in the linear regression model) as writing process indices. Subsequently, correlation coefficients between all the variables was computed, and the network of writing process indices and writing strategy preference were described for both highly and intermediate proficient group. These procedures above were all conducted using R (R Core Team, 2019) and the *qgraph* package (Epskamp et al., 2012).

4.3 Results and Discussion

The network structure based on correlation coefficients among all the variables is shown in Figure 4. Each node represents the variable, and the thickness of each edge shows the strongness of correlation between the variables. In the highly proficient group's network (Figure 4), it can be seen that all the variables are densely clustered, indicating a very strong correlation among all the variables. Although the relationship between each of the indices is quite complicated, we can conclude that the writing strategies, that highly proficient writers preferred, tend to be reflected in the behavioral writing process.

Figure 4

The network of correlation between all the variables of the advanced writer groups

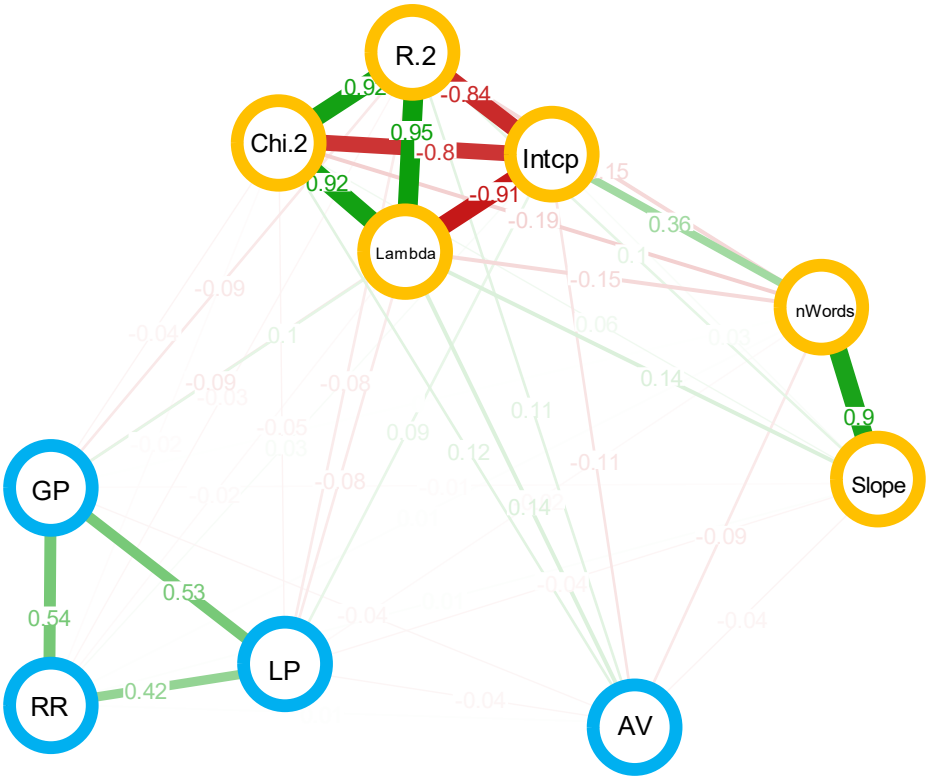


Note: Yellow circles are indices of writing process, and blue circles are writing strategy preference scores. Green edge shows a positive correlation, and red ones show negative. nWords = number of words; Chi.2 = χ^2; Intcp = intercept; R.2 = coefficient of determination, R^2; GP = global planning; LP = local planning; AV = avoidance; RR = revision/reviewing

On the other hand, the network structure of the intermediate proficient group (Figure 5) clearly shows a different pattern from that of the highly proficient group. The writing strategy preference scores and the writing process indices are located far apart, indicating a very weak correlation between the two. Among the writing strategies, the global planning (GP), the local planning (LP), and the revision/reviewing strategy (RR) were strongly correlated with each other as well as with the highly proficient group, but they were not correlated with the writing process indices. It can be assumed that intermediate-proficiency writers' strategies did not affect their writing processes.

Figure 5

The network of correlation between all the variables of the intermediate writer groups



Note: Yellow circles are indices of writing process, and blue circles are writing strategy preference scores. Green edge shows a positive correlation, and red one shows a negative one. nWords = number of words; Chi.2 = χ^2; Intcp = intercept; R.2 = coefficient of determination, R^2; GP = global planning; LP = local planning; AV = avoidance; RR = revision/reviewing

These results can be summarized by stating that the highly proficient writers' writing strategies were reflected in their patterns of changes in the number of words, while the intermediate-proficiency writers' strategies were not. This is because the highly proficient writers had a higher degree of metacognitive ability related to the strategy use than the intermediate writers. However, these results might have been influenced by the differences of task conditions such as the task time, the limit of the number of words, and topic. Also, there is a large difference in sample size between the two groups, and yet the sample size of highly proficient group was very small. No matter how robust the analysis is, sampling error is likely to occur due to the small sample size. Thus, future studies should be conducted using large data sets collected in writing tasks with

controlled conditions.

5. Discussion

5.1 General Findings

The current study introduced the previous studies and the author's studies on the relationship between L2 writing process and proficiency or evaluation with keystroke logging approach. Although there is some limitations to consider, the present study reveals following three points: (a) the pattern of changes in the number of words over time is associated with essay ratings, (b) there is a strong association between writing strategy preferences and indices of the changes in the number of words over time as writing process indices for the high proficient writers, (c) there is a weak association between writing strategy preferences and indices of the changes in the number of words over time as writing process indices for the intermediate proficient writers. These findings suggested that there is a relationship between writing process and proficiency or evaluation.

5.2 Future Direction for L2 Writing Process Research

Previous research has not fully examined the interrelationships among writing process indices and their relationship between these indices and features related to evaluation of essay and writers' proficiency. Examining these relationships and organizing the writing process indices are important from the perspective of integrating research findings. As part of the current research project, we are going to start research project on this point by adding essay ratings to the WritingMaetriX corpus' data by e-rater or human raters. Also, indices that contribute to the automatic assessment of the writing process itself should be explored.

Research practices with a mind to an application in the pedagogical context also should be carried out. For instance, the writing process indices, which include those that we presented in the article (e.g., λ and χ^2 in the Poisson distribution model, and slope, intercept, and R^2 in the linear regression model), will bring us a probability that EFL teacher give their students feedback on their writing process using a single index representing process features.

5.3 Implication for Building a Writing Process Corpus

As Ishii et al (2015) and Ishii (2016) stated, future research topics for writing process corpus study include the following points. First, we should investigate the relationship

between errors/mistakes and revision activities, word choice and rewriting, cognitive process in writing and individual differences. Also, the possibility of automated scoring based on writing process should be examined. Moreover, we should attempt to create models of the writing process considering writers' first language, proficiency, and task condition (e.g., genre, topic, task time, permission of reference). Finally, it is possible to analyze the time series of changes in linguistic indices (e.g., complexity, accuracy, and fluency). For instance, we can find whether revision or reviewing strategies are actually used by investigating improvements in complexity and accuracy indices at the end of the writing task.

6. Conclusion

This article provided an overview of the relationship between the L2 writing process and writing evaluation and showed the potential of the keystroke logging approach. Also, it presented two examples of studies on this relationship using key-stroke logging technology and suggested that the change in numbers of words over time as an observable index of writing process is related to essays' evaluation, writers' proficiency, and writers' self-reported preference of strategies. As keystroke-logging technology was employed, and it allowed use of many writing process indices. In future studies, to integrate the research findings and compare the results of the studies, we should revisit the relationship among these indices to decide on common measurement variables. Also, researchers on L2 writing and learner corpus studies should continue to make attempts to build a writing process corpus, considering conditions of essay writing tasks.

References

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*, 1–18. http://www.jstatsoft.org/v48/i04/

Ishii, Y. (2016). Eigogakushuusha no raityingu purosesu no kaimei: Kii nyuuryoku kiroku wo mochiita purosesu no kashika [Investigating learners' writing processes: Keystroke logging approach]. *Eigo Kyouiku* [The English Teacher's Magazine], *65* (3), 66–67.

Ishii, Y., Ishii, T., Kawaguchi, Y., Abe, D., Nishimura, Y., & Kusanagi, K. (2015) WritingMaetriX wo mochiita gengoshigen no kouchiku to eigogakushuusha no raityingu-purosesu no kaimei [Constructing language resources based on WritingMaetriX and investigating learners' writing processes]. *Gaikokugo kyouiku*

medyia gakkai dai 55 kai zenkoku kenkyuu taikai happyou youkoushuu [The Proceedings of the 55th Annual Conference of Japan Society for Language Education and Technology], 190–193.

Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In S. Ishikawa (Ed.), *Learner corpus studies in Asia and the world*, *1* (pp. 91–118). Kobe University.

Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Lawrence Erlbaum Associates.

Hirose, K., & Sasaki, M. (1994). Explanatory variables for Japanese students' expository writing in English: An exploratory study. *Journal of Second Language Writing*, *3*, 203–229. https://doi.org/10.1016/1060-3743(94)90017-5

Kawaguchi, Y. (in press) Kakite no raityingu-houryaku ha gosuu no huekata ni hanei sarerunoka [Are writers' writing strategies reflected on the change in numbers of words over time?] In N. Tsuji, Y. Isobe, Y. Kawaguchi, S. Aota, & M. Ueno (Eds.) *Eigokyouiku no rekishi ni manabi, genjou wo toi, mirai wo hiraku: Erikawa Haruo sensei taishoku kinen ronshu* [Past, present, and future of English education: A festschrift for Professor Erikawa Haruo on his retirement].

Kawaguchi, Y., Murota, D., Goto, A., & Kusanagi, K. (2016). Writing processes and ratings of products: A keystroke logging and model-fitting approach. *Language Education & Technology*, *52*, 319–343.

Kusanagi, K., Abe, D., Fukuta, J., & Kawaguchi, Y. (2015). Developing a software to record, visualize, and analyze learner's writing process: WritingMaetriX. *LET Journal of Central Japan*, *26*, 23–34. https://doi.org/10.20656/letcj.26.0_23

Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, *30*, 358–392. https://doi.org/10.1177/0741088313491692

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Roca de Larios, J., Marin, J., & Murpjy, L. (2001). A temporal analysis of formulation processes in L1 and L2 writing. *Language Learning*, *51*, 497–538. https://doi.org/10.1111/0023-8333.00163

Roca de Larios, J., Manchón, R., Murphy, L., & Marin, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*, 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, *9*, 259–291. https://doi.org/10.1016/S1060-3743(00)00028-X

Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, *46*, 137–168. https://doi.org/10.1016/S1060-3743(00)00028-X

Spelman Miller, K., Lindgren, E., & Sullivan, K. P. H. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, *42*, 433–454. https://doi.org/10.1002/j.1545-7249.2008.tb00140.x

Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revision in L1 and FL. *Journal of Second Language Writing*, *15*, 201–233. https://doi.org/10.1016/j.jslw.2006.06.002

Sugiura, M. (2011). Gengoshuutokukenkyuu no tameno gakushuushakopasu [Learner corpora for second language acquisition studies]. In I. Fujimura, & N. Takizawa (Eds.), *Gengokenkyuu no gihou: Deta no shuushuu to bunseki* [Techniques for language studies: Data collection and analysis] (pp. 123–140). Hitsuji Shobou.

Victori, M. (1999). An analysis of writing knowledge in EFL composing: A case study of two effective and two less effective writers. *System*, *27*, 537–555. https://doi.org/10.1016/S0346-251X(99)00049-4

Xu, C., & Qi, Y. (2017). Analyzing pauses in computer-assisted EFL writing: A computer-keystroke-log perspective. *Journal of Educational Technology & Society*, *20*, 24–34. https://www.jstor.org/stable/26229202

Yamanishi, H. (2009). Japanese EFL learners' use of writing strategies: A questionnaire survey. *The Bulletin of the Writing Research Group, JACET Kansai Chapter*, *8*, 53–64.