

Self-Assessment of Topic Development in Written Production among High School Students

Yoko Suganuma Oi
Waseda University

Oi, Y. S. (2014). Self-assessment of topic development in written production among high school students. *Journal of Pan-Pacific Association of Applied Linguistics*, 18(1), 145-170.

The present study mainly focuses on the topic development in student written production through the consistency between student self-assessment and teacher assessment. In the present study, topic development means “cohesion”, “overall organization”, and “coherence”. It proposes the next hypotheses: (i) Students could assess cohesive devices such as discourse markers as well as teachers do. (ii) Students could assess the overall organization pattern of their written production as well as teachers do. (iii) Students could assess the coherence of their written production as well as teachers do. 169 Japanese high school students whose ages ranged from 16 to 18 years old were asked to write an essay with about 150 words in 30 minutes. One American English teacher and one Japanese English teacher also participated in the study. Teachers and students used the same self-assessment sheet written in both English and Japanese. And then 82 students’ written production was chosen as the subject based on the number of the words*. They were divided into three groups referring to teachers’ assessment. The results show that the writing patterns in terms of topic development are divided into eight patterns. The self-assessment of the students is not different from teacher assessment in the aspect of overall organization. However, the consistency between teacher assessment and student self-assessment was not found in the assessment of other cohesive devices and overall coherence.

Key Words: self-assessment, written production, topic development

1 Introduction

The objective to teach written production has been recognized to further enhance students’ abilities to evaluate facts and opinions from multiple perspectives and communicate through reasoning and a range of expressions. So the necessity to teach how to develop a topic has been more strongly focused after the enforcement of the new national government guidelines in

* The participants ethically approved of the cooperation with the study because the participation of the study was also beneficial for the students.

Japan (Designated for Course of Study for Senior High Schools, 2011). Besides, a style of writing by oriental students in English has been compared to the “approach by indirection, turning in a widening gyre” (Kaplan, 1966, p. 10). B.D. Harder and H. Harder also stated that the Japanese essays written in English seemed to be disorganized and illogical, filled with nonrelevant material, developed incoherently with statements that remain unsupported (1982). Though the shortage of concrete data in both studies has been indicated, it is a fact that Japanese students have had difficulties mastering rhetorical patterns at the level of discourse, not the sentences (Oi, K., 1986).

It has also been discussed that Japanese writers tend to hesitate to express their initial positions, so the readers must wait for the final paragraph to understand the writers’ conclusion. Hinds (1990) characterized Japanese organizational pattern as “quasi-inductive”. Kubota (1998) presented that a characteristic of Japanese writing is “induction”.

On the other hand, the reliability of the assessment of topic development has been discussed because it might be subjective or ambiguous. For topic development comprises multi-components such as cohesion, coherence, and structure. Furthermore, these components are not independent and related each other. The accuracy of language use in a text might be also influential when evaluators assess written production. Therefore the present study tries to define topic development and find the useful assessment criteria of topic development for Japanese senior high school students.

The reason why I focus on students’ self-assessment is that self-assessment is effective to motivate learners and nature self-reliance, and independence. If it were conducted for several times, students develop learner awareness and monitor while assessing their own written production (Oi, S. Y., 2012). Finally it will help students to revise their written production, so it is needed to establish a self-assessment system especially for topic development, because topic development is one of the key factors to make the written production closer to the readers. As a result of it, students would enjoy self-attainment through writing in English.

Luoma and Tarnanen support this idea that self-assessment is recognized to support learning through raising the learners’ awareness of their learning goals and their achievements through feedback and/or assessment discussion (2003, p. 441).

Thus, it is important for Japanese students to learn how to develop a topic in English composition and to make their idea more persuasive through being aware of what part is insufficient in their written production in terms of topic development. For the researcher believes that self-awareness and self-reflection work to develop learner autonomy and encourage students to revise their written production by themselves (Oi, S. Y., 2013).

The self-assessment criteria for the topic development help students to organize structure and see their written production more objectively, because self-assessment bridges the gap between their present English proficiency and

ideal English proficiency (Oi, S. Y., 2013).

1.1 Self-assessment

Le Blanc and Painchaud (1985) investigated the usefulness of self-evaluation as a second language placement instrument (pp. 673-87). Bachman & Palmer also prove that self-ratings can be reliable and valid measures of communicative language abilities (1981, pp. 67-86). Other researchers state benefits and suggestions on self-assessment. For instance, Oskarsson presented six advantages of using self-assessment: 1) promotion of learning, 2) raising level of awareness, 3) improving goal-orientation, 4) expansion of range assessment, 5) sharing assessment burden, 6) beneficial post-course effects (Oskarsson, 1989, pp. 1-13). Cheng also showed three implications of self-assessment. The first one was the development of self-reflection on their performance and learning process. The second was the aid to train students to become better raters and learners. The final implication was the students' psychological factors. The students in Cheng's study tended to evaluate themselves lower and it was caused by their affective factors (Cheng, 2008, pp. 254-255).

Though some researchers doubt that learners could evaluate their own proficiency correctly and might overstate or downgrade their achievement (Boud & Falchikov, 1989; Kent, 1980), Little insisted that self-assessment is fundamental to effective European Language Portfolio (ELP) use, showing the usage of "can-do" statements (Little, 2002).

The reliability of student self-assessment has been researched, focusing on the agreement between teacher assessment and student self-assessment. For instance, Crocker and Cheeseman (1988) presented that even 141 young children who are in three infant schools could rank themselves with a high degree of agreement with teacher assessment.

On the other hand, Peirce, et al. (1993) investigated the self-assessment of French proficiency made by 500 Grade 8 students and indicated that (1) self-assessment of language proficiency correlated only weakly with objective measures of language proficiency; (2) self-assessment measured on specific tasks were more highly correlated with tested proficiency than were global self-assessment measures; (3) irrespective program, students agreed on the relative difficulty of oral and literacy tasks in French under specific conditions of reception and production.

The researcher also examined the self-assessment of about one-minute oral speeches in English conducted by 92 Japanese senior high school students. It was found that there was a consistency between teacher assessment and student self-assessment only in the evaluation component, Language Use, not other components such as Topic Development and Delivery (Oi, Y. S., 2012). However, it was conducted on one-minute oral

speeches, so it was difficult for the students to develop a topic. Therefore, the researcher decided to investigate the self-assessment about written production because it seemed to be easier for students to develop a topic in written production, compared to oral production.

1.2 Criteria of topic development defined

Witte (1983) discusses that the development of the discourse topic is comprised of a succession of hierarchically ordered subtopics, each of which contributes to the discourse topic, and that it should be treated as a sequence of ideas because most sentences must be related to the same subtopic from a sequence.

In the present study, the researcher investigates Topic Development in terms of micro-structure and macro-structure. Micro-structure is defined as cohesive device, and macro-structure was defined as overall organization. According to van Dijk (1980), macro structures achieve global coherence because they are higher level semantic or conceptual structures that organize the “local” micro-structures of discourse, interaction, and their cognitive processing.

Micro-structures are local and cohesive structures which establish relationships between the topics of successive sentences (Hobbes & Evans, 1979; van Dijk, 1980). van Dijk indicates that cohesive ties create only “local” coherence (cohesion), and are unable to create discourse-level or “global” coherence. Discourse markers play an important role to connect information ties together and make coherence. So, it is analyzed in the present study whether discourse markers such as conjunctions, adverbs, idioms and tense are used adequately in written production.

Macro-structure means “theme”, “topic plan”, “purpose”, and “intention of writers” (van Dijk, 1980). It establishes a relationship between sentence topics and discourse topic (Hobbes & Evans, 1979). As overall organization on the level of macro-structure is investigated in the present study based on the self-assessment sheet consisted of “introduction”, “body”, and “conclusion”. In English expository writing, it is common to have a topic sentence in the beginning part of the text (Oi, K., 1986, p. 26). So an ideal “introduction” should have a topic sentence which presents “situation”, “problem”, “solution” and “outcome (comments)” (Nakano, 2001). “Body” develops the main idea, presenting examples and reasons. “Conclusion” restates a main idea.

2 The Present Study

2.1 Participants

Self-Assessment of Topic Development in Written Production among High School Students

The data of the present study for the analysis came from the written production of 169 (79 boys and 90 girls) Japanese high school students. Their ages ranged from 16 to 18 years old. 62 percent of them belonged to public high schools and 38 percent of them belonged to private high schools. Ten of those students have lived in English speaking countries. One American English teacher and one Japanese English teacher also participated in the study.

They were also required to take the Measure of English Grammar (MEG) Test (Nakano, et al., 2003) to assess their English grammar level before the experiment (Table 1). The types and tokens of their written production were also calculated to analyze their vocabulary. The mean number of types was 489.34 and that of tokens was 2355.83. They were also asked to answer questionnaires related to the self-assessment after the experiment.

Table 1. The SD and Mean of the Results of Measure of English Grammar (MEG) Test

N	MAX	MIN	MEAN	SD
169	32	2	14.9	6.8

Note. The maximum score is 35 points; MAX means the highest score; MIN means the lowest score; SD means standard deviation.

2.2 Hypotheses

The researcher proposed the following hypotheses to seek efficacy of self-assessment in terms of topic development of written production:

1. Students could assess cohesive devices such as discourse markers as well as teachers do.
2. Students could assess the overall organization of their written production as well as teachers do.
3. Students could assess the coherence of their written production as well as teachers do.

3 Method

Before the study, every student received the feedback of writing patterns and comments about their compositions which were written before the present study. Teachers also gave students a lecture on overall organization, cohesive devices such as discourse markers, and coherence.

And then, students were asked to write an essay on the guided writing topic, "A foreign visitor has only one day to spend in your country. Where should this visitor go on that day? Why? Use specific reasons and details to

support your choice”. Students were given for thirty minutes, without using dictionaries during writing. Students were also asked to write at least fifteen lines, because the average number of words in one line words was about ten, for it was expected to have 150 words. Further, it was counted by students, since they had to write it by hand in classroom.

Students were not informed of the topic before writing. And then the students assessed their written production after writing, using the same assessment sheet, as teacher’s sheet was written in both English and Japanese. After the study, students received both the focused-on form corrective feedback and commentative feedback in English in one week.

The written production of 82 students was chosen as the subject of the present study, because they could write more than ten lines. And then 82 students were divided into three groups based on the total score of five assessment components; high-scored group (16–20 points), middle-scored group (15–11 points), and low-scored group (10–2 points). Table 3 and 4 present the data of their written production.

Table 3. The Mean Number of Tokens

N	MAX	MIN	MEAN	SD
82	274	115	166.85	38.46

Table 4. The Mean Number of Sentences

N	MAX	MIN	MEAN	SD
82	42	8	17.2	5.63

3.1 Assessment criteria

The self-assessment sheet was developed by the researcher based on TOEFL Test Integrated Writing Rubrics (2008) and Common European Framework of Reference for Languages (2001). It consisted of six components: “Introduction”, “Body”, “Conclusion”, “Discourse Markers”, and “Coherence”, and the total score of all of the five components. Each component was composed of 3 scores, and general evaluation was 5 scores. This assessment sheet was used to analyze the inter-rater reliability between two teachers, and to find the difference between students’ self-assessment and teachers’ assessment (see Appendix A).

3.2 Methods of analysis

Four types of analysis were conducted to examine research questions. At first, Kendall’s tau was used to see the inter-rater reliability between the assessments of two teachers. Second, the Wilcoxon signed-rank test was conducted to look for differences between student self-assessment and

Self-Assessment of Topic Development in Written Production among High School Students

teacher assessment. Thirdly, two teachers categorized students' written production based on Topic Development criteria into eight patterns. Fourthly, cohesive devices such as conjunctions were analyzed based on Halliday and Hassan's categories (Halliday & Hasan, 1976; Oi, K., 1986).

Finally, two measures were conducted in order to assess coherence. At first, the most frequently used word was identified as a key concept in compositions. And then, the organization of the "theme (topic)" (2014, Halliday), i.e. the topic was classified taking into consideration the presence or absence of "theme".

82 students were divided into three groups based on the results of the total score of the assessment sheet. Table 5 shows the ratio of each group.

Table 5. The Total Sum of Two Teachers' Assessment about Topic Development

N	MAX	MIN	MEAN	SD
82	20	2	13.20	4.02

Table 6. Three Groups Based on the Total Score of Topic Development

	High-scored group	Middle-scored group	Low-scored group
Scores	20-16 points	15-11	10-0
Number of students	27	27	30

3.3 Analysis of cohesive devices

According to Halliday and Hasan (1976, p. 18), "cohesion" is defined as the set of possibilities that exist in the language for making text hang together: the potential that the speaker or writer has at his/her disposal.

The data of cohesive devices were considered in terms of the following: 1. pronominal: exophoric, anaphoric; 2. demonstrative; 3. comparative; 4. substitution; 5. ellipsis: regular, pronominal; 6. conjunctions: additive, adversative, causal, temporal, others; 7. lexical repetitiveness.

About categories 1 through 6, and frequency of use was analyzed for each written production.

Concerning category 7, repeated use of the same word which indicated the main topic in its written production, was counted. The data were analyzed in terms of the number of repetition sets (those words occurring more than two times).

3.4 Analysis of the overall organization pattern

The written production of 82 students was classified into eight groups, referring to the result of the previous study (see Appendix B).

3.5 Analysis of organization of “theme (topic)”

Each written production was classified, according to the presence, absence and location of the theme (topic) and comments.

The location of the theme (topic) was determined by identifying a sentence that serves to answer the question, “A foreign visitor has only one day to spend in your country. Where should this visitor go on that day? Why? Use specific reasons and details to support your choice.” The sentence that contains “the place” where the visitor should go is identified as the theme (topic) in written production. Moreover, the sentences that include reasons, examples, comparison and description are categorized as comments. All of the sentences were numbered and labeled as “theme”, “comment”, or “undetermined”. The following text is a representative example.

In the following example, waved lines indicate theme (topic), and (C) means comments. The number means the number of comments, and hyphenated numbers indicate related comments. In this case, the most frequently used word is “Asakusa”, therefore “Asakusa” is understood as the key concepts in this composition.

Example 1.

1. I think a foreign should visit Asakusa. (T)
2. There are 4 reasons. (C)
3. So I will explain them.
4. First, there is "Kaminari-mon" in Asakusa. (C1)(T)
5. It is a big and beautiful Japanese gate. (C1-1)
6. The way to "kaminari-mon" from station of Asakusa. (C1-2)(T)
7. there are a lot of shops. (C1-3)
8. They sell Japanese snacks, subenior and so on. (C1-4)
9. So a foreign can enjoy Japanese culture. (C1-5)
10. Second, a foreign can see "Sky-tree" in Asakusa. (C2)(T)
11. Sky-tree is the highest building in Japan. (C2-1)
12. That of height is 634 meter!! (C2-2)
13. If a foreign want to go sky-tree, he will watch the beautiful sight of Tokyo. (C2-3)
14. Third, there is "sora machi" in Asakusa. (C3)(T)
15. "Sora-machi" is a shopping mall. (C3-1)
16. There are a lot of delicious food and Japane traditional goods. (C3-2)
17. For example, ラメン , takoyaki, furoshiki, sensu and so on. (C3-3)
18. So a foreign will have a good time in sora-machi. (C3-4)
19. Fianally, a foreign can feel tour seasons of Japanese in Asakusa. (C4-1)(T)
20. In spring, there are a lot of cherry blossom. (C4-2)
21. In summer and autumen, the festivals are hold on kaminari-mon. (C4-3)

Self-Assessment of Topic Development in Written Production among High School Students

- 22. The countries that have four seasons are rare. (C4-4)
- 23. So So if a foreign visit Asakusa, he will be surprise at the change of seasons. (C4-5)(T)
- 24. In conclusion, I think a foreign should visit Asakusa. (T)
- 25. Whenever I visit Asakusa, I enjoy very much. (T)
- 26. If a foreign visit Asakusa, he will feel Japanese traditional very much. (T)

Note. “C” means “supporting comment”. ; “T” indicates “theme (topic)”.

The next Figure 1 shows moves between concepts through the relationship between theme and comments. There were 28 moves in the composition.

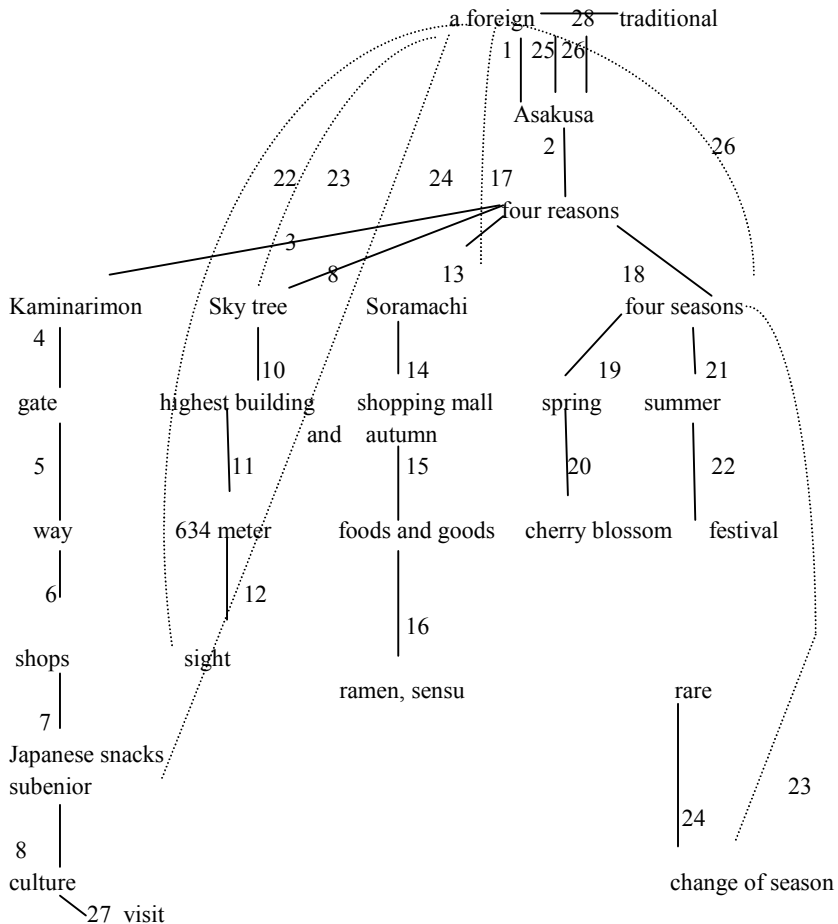


Figure 1. The hierarchy of moves between theme (topic) and comments of Example 1

Note. Solid lines indicate sequential moves, and the dotted lines indicate extended moves.

4 Results

4.1 Inter-rater reliability between the assessment of two teachers

Before the analysis, the inter-rater reliability between the assessments by two teachers was calculated. It was found that there was a high correlation between the assessments by two teachers as Table 7 shows. Therefore, the consistency between student self-assessment and teacher assessment could be analyzed based on the results of two teachers.

Table 7. Inter-rater Reliability between Two Teachers' Assessment

Components	Introduction	Body	Conclusion	Discourse Markers	Coherence	Total Score
Kendall's tau coefficient	.970 **	.939**	.867**	.965**	.888**	.922**

* $p < .05$ ** $p < .01$

4.2 Hypothesis 1: Students could assess cohesive devices such as discourse markers as well as teachers do.

According to the results of the Wilcoxon signed-rank test between the assessment by two teachers and student self-assessment, high and middle scored students could not assess cohesive devices as well as teachers do. On the other hand, the results of students who had low scores showed the consistency only with native English teacher's assessment. It is assumed that high and middle scored students could not judge their usage of discourse markers compared to low scored students.

Table 8. The Wilcoxon Signed-rank Test between Student Self-assessment and Native English Teacher Assessment of Discourse Markers

Discourse markers of 82 students		Discourse markers of 27 high-scored students		Discourse markers of 27 middle-scored students		Discourse markers of 30 low-scored students	
T	z	T	z	T	z	T	z
9	-4.502	9.50	-3.989	19.00	-3.300	33.00	.000

$p < .05$ T (the sum of ranks of the smaller of the two ranks); z (Wilcoxon signed-rank test z score); < 1.96

Self-Assessment of Topic Development in
Written Production among High School Students

Table 9. The Wilcoxon Signed-rank Test between Student Self-assessment and Japanese English Teacher Assessment of Discourse Markers

Discourse markers of 82 students		Discourse markers of 27 high scored students		Discourse markers of 27 middle scored students		Discourse markers of 30 low scored students	
T	z	T	z	T	z	T	z
10	-4.253	9.50	-3.989	18.00	-3.153	36.00	-.247

$p < .05$ T (the sum of ranks of the smaller of the two ranks); z (Wilcoxon signed-rank test z score); < 1.96

The reason why there was not any consistency between student assessment and teacher assessment except low scored students was assumed to be related to the ratio and the way how to use cohesive devices. So the study tries to find the characteristics of cohesive devices which were used by each group's students. The following data were collected by dividing the number of occurrences of cohesive devices by cohesive devices by the total number of the sentences in each group.

Table 10. Pronominal

82 students	Exophoric 10/1416=0.0070621	Anaphoric 423/1416=0.2987	Total 433/1416=0.3057909
High scored students	Exophoric 2/479=0.0041753	Anaphoric 117/479=0.2442588	Total 119/479=0.2484342
Middle scored students	Exophoric 4/481=.008316	Anaphoric 136/481=.2827442	Total 140/481=.2910602
Low scored students	Exophoric 4/488=.0081967	Anaphoric 170/488=.3483606	Total 174/481=.3617463

As the above table presents, anaphoric devices are more frequently used in written production compared to exophoric devices in every group of students. And it is also found that high-scored students used the least pronominal devices of the three groups.

On the other hand, the other devices show different results. Table 11, 12, and 13 present the ratio of occurrences of "substitute" and "ellipsis". The students did not frequently use substitution and ellipsis. Compared to the usage of pronominal, there was not a big difference between the groups. Any example of substitution and regular ellipsis was not found in the group of low

scored students. On the other hand, high-scored students used pronominal and ellipsis most frequently of all three groups.

Table 11. Substitution

82 students	High-scored students	Middle-scored students	Low-scored students
11/1416=0.0077683	5/479=.0104384	6/481=.0124	0/488=0

Table 12. Regular Ellipsis

82 students	High-scored students	Middle-scored students	Low-scored students
3/1416=.0021186	1/479=.00208	2/481=.00415	0/488=0

Table 13. Pronominal Ellipsis

82 students	High-scored students	Middle-scored students	Low-scored students
4/1416=.0028248	2/479=.004175	1/481=.002079	1/488=.0020491

Table 14 shows the list of conjunctions found in the data. In the second, third, and fourth columns, the number of occurrences of a given conjunction is presented. Furthermore, all the conjunctions were counted up according to the five categories. The percentages were calculated by dividing all the occurrences by the total number of the sentences.

Table 14. Conjunctions and Adverbs

	82 students	High-scored	Middle-scored	Low-scored
Additive				
and	118/1416=0.0833	37/479=.0772	33/481=.0686	48/488=.0983
also	22/1416=0.0156	10/479=.020	6/481=.012	6/488=.0124
additionally	0/1416=0	0/479=0	0/481=0	0/488=0
in addition	6/1416=0.0042	4/479=.008	1/481=.002	1/488=.0020
besides	0/1416=0	0/479=0	0/481=0	0/488=0
furthermore	0/1416=0	0/479=0	0/481=0	0/488=0
moreover	1/1416=0.0007	0/479=0	0/481=0	1/488=.0124
that is to say	0/1416=0	0/479=0	0/481=0	0/488=0
like this	0/1416=0	0/479=0	0/481=0	0/488=0
or	1/1416=0.0007	1/479=.00208	0/481=0	0/488=0

Self-Assessment of Topic Development in
Written Production among High School Students

Total	148/1416=0.1045	52/479=.108	40/481=.083	56/488=.114
Adversative				
though	1/1416=0.000706	0/479=0	1/481=.00207	0/488=0
but	59/1416=0.04166	23/479=.0480	14/481=.0291	22/488=.0450
however	6/1416=0.004	2/479=.00417	3/481=.00623	1/488=.00204
instead	0/1416=0	0/479=0	0/481=0	0/488=0
on the other hand	2/1416=0.001	1/479=.00208	0/481=0	1/488=.00204
although	1/1416=0.000706	1/479=.00208	0/481=0	0/488=0
on the contrary	0/1416=0	0/479=0	0/481=0	0/488=0
by contrast	0/1416=0	1/479=.00208	0/481=0	0/488=0
Total	69/1416=0.0487288	28/479=.0584	18/481=.0374	25/488=.0512
Causal				
then	9/1416=0.0063559	3/479=.00626	0/481=.0374	6/488=.0122
therefore	16/1416=0.0112994	8/479=.0167	7/481=.0145	1/488=.00204
in this sense	0/1416=0	0/479=0	0/481=0	0/488=0
so	99/1416=0.0699152	37/479=.0077	23/481=.0047	39/488=.0079
because	52/1416=0.0367231	14/479=.0292	15/481=.0311	23/488=.0471
on that condition	0/1416=0	0/479=0	0/481=0	0/488=0
as a result	0/1416=0	0/479=0	0/481=0	0/488=0
accordingly	0/1416=0	0/479=0	0/481=0	0/488=0
that is because	5/1416=0.00353	3/479=.00626	2/481=.00415	0/488=0
from the above-mentioned	0/1416=0	0/479=0	0/481=0	0/488=0
considering this	0/1416=0	0/479=0	0/481=0	0/488=0
in light of this	0/1416=0	0/479=0	0/481=0	0/488=0
if so	0/1416=0	0/479=0	0/481=0	0/488=0
since	2/1416=0.001412	2/479=.00417	0/481=0	0/488=0
the reason	1/1416=0.000706	0/479=0	1/481=.00207	0/488=0
by these reasons	1/1416=0.000706	0/479=0	1/481=.00207	0/488=0
as	3/1416=0.002	0/479=0	1/481=.00207	2/488=.00409

	1186			
That is why	4/1416=0.002 8248	1/479=.00208	1/481=.00207	2/488=.00409
These are the reasons why for	2/1416=0.001 412	2/479=.00417	0/481=0	0/488=0
	1/1416=0.000 706	0/479=0	1/481=.00207	0/488=0
Total	195/1416=0.1 377118	70/479=.146	52/481=.1081	73/488=.1495
Temporal				
first	45/1416=0.03 17796	16/479=.0334	17/481=.0353	12/488=.0245
firstly	1/1416=0.000 706	1/479=.00208	0/481=0	0/488=0
secondly	2/1416=0.001 412	1/479=.00208	1/481=.00207	0/488=0
then	1/1416=0.000 706	1/479=.00208	0/481=0	0/488=0
in short	0/1416=0	0/479=0	0/481=0	0/488=0
at first	0/1416=0	0/479=0	0/481=0	0/488=0
second	32/1416=0.02 25988	14/479=.0292	14/481=.0291	4/488=.00819
third	20/1416=0.01 41242	11/479=.0229	7/481=.0145	2/488=.00409
fourth	0/1416=0	0/479=0	0/481=0	0/488=0
last	3/1416=0.002 1186	0/479=0	3/481=.00623 7	0/488=0
in conclusion	12/1416=0.00 84745	9/479=.0187	1/481=.00207	2/488=.00409
from now on	0/1416=0	0/479=0	0/481=0	0/488=0
first of all	0/1416=0	0/479=0	0/481=0	0/488=0
next	13/1416=0.00 91807	2/479=.00417	5/481=.0103	6/488=.01229 5
finally	8/1416=0.005 6497	4/479=.00835	2/481=.00415	2/488=.00409
last	1/1416=0.000 706	0/479=0	0/481=0	1/488=.00204 9
one	1/1416=0.000 706	0/479=0	0/481=0	1/488=.00204 9
two	1/1416=0.000 706	1/479=.00208	0/481=0	0/488=0
the other	1/1416=0.000 706	0/479=0	0/481=0	1/488=.00204 9
another	2/1416=0.001	0/479=0	2/481=.0103	0/488=0

Self-Assessment of Topic Development in
Written Production among High School Students

	412			
Total	143/1416=0.1009887	60/479=.1252	52/481=.1081	31/488=.0635245
<hr/>				
Others				
for example of course	57/1416=0.040706	22/479=.0459	19/481=.0395	16/488=.0327
after all in fact	0/1416=0 1/1416=0.000706	0/479=0 1/479=.00208	0/481=0 0/481=0	0/488=0 0/488=0
anyway surely	0/1416=0 1/1416=0.000706	0/479=0 0/479=0	0/481=0 0/481=0	0/488=0 1/488=.002049
now such as	3/1416=0.002531	1/479=.00208	2/481=.00415	0/488=0 3/488=.0061475
like	2/1416=0.001412	0/479=0	1/481=.00207	1/488=.002049
actually	1/1416=0.000706	0/479=0	0/481=0	1/488=.002049
definitely	1/1416=0.000706	0/479=0	1/481=.00207	0/488=0
such as	1/1416=0.000706	1/479=.00208	0/481=0	0/488=0
like	1/1416=0.000706	1/479=.00208	0/481=0	0/488=0
Total	74/1416=0.052706	27/479=.0563	25/481=.0519	22/488=.0450
Total of conjunctions and adverbs	629/1416=0.4449	237/479=0.49	187/481=.388	207/488=.424

High and low scored students showed the similar results in the frequency of Additive, Adversative, Causal. In these cases, middle-scored students showed the least frequency of all three groups. On the other hand, only Temporal and Others showed the highest frequency of high-scored students, and middle-scored students presented the second highest. It is assumed that the usage of temporal cohesive devices is related to the score of topic development. Temporal cohesive devices play a key role to organize written production. It is natural that high-scored students manage them well.

Table 15 presents lexical cohesion. The average number of words repeated more than twice was calculated by adding the number of repeated words and then dividing by the number of total sentences. As the most frequently used words symbolize the theme of the students' written

production symbolizes with the theme (topic) of their written production. There was a tendency that high-scored students had high ratio of lexical cohesion while low-scored students did not repeat the words which describe as key concepts in their written production.

Table 15. Total of Lexical Cohesion

82 students	High-scored students	Middle-scored students	Low-scored students
494/1416=0.34887	187/479=0.390	166/481=0.345	146/488=0.299

Note. Appearance of repetitive words/total words

4.3 Hypothesis 2: Students could assess the overall organization pattern of their written production as well as teachers do.

The Wilcoxon signed-rank test was calculated to see the difference between student self-assessment and teacher assessment of overall organization. The results present that students could assess introduction as well as native English teachers, but they could not assess “body” and “conclusion”. This is the same result of all of the levels. However, students could assess the overall organization, as the results of the Wilcoxon signed-rank test of two each teacher and student self-assessment showed. There was the consistency in both relationship between two teachers’ assessment and student assessment.

Table 16. The Wilcoxon Signed-rank Test between 82 students’ Self-assessment and Native English Teacher Assessment of Overall Organization

Introduction		Body		Conclusion		Total Score	
T	<i>z</i>	T	<i>z</i>	T	<i>z</i>	T	<i>z</i>
39	-.222	29	-2.820	26	-3.111	64	-1.250

p<.05 T (the sum of ranks of the smaller of the two ranks); *z* (Wilcoxon signed-rank test *z* score); < 1.96

Table 17. The Wilcoxon Signed-rank Test between 82 students’ self-assessment and Japanese English Teacher Assessment of Overall Organization

Introduction		Body		Conclusion		Total Score	
T	<i>z</i>	T	<i>z</i>	T	<i>z</i>	T	<i>z</i>
40	-.327	35	-1.617	27	-4.023	73	-.245

p<.05 T (the sum of ranks of the smaller of the two ranks); *z* (Wilcoxon signed-rank test *z* score); < 1.96

Based on the analysis of students’ writing organization, students’ written production was categorized into 8 Patterns (see Appendix C). 82.9% students presented Pattern 8.

Table 18. Topic Development Patterns

N	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5	Pattern 6	Pattern 7	Pattern 8
82	0	0	0	1	1	1	11	68

4.4 Hypothesis 3: Students could assess the coherence of their written production as well as teachers do.

Students were asked whether their written production had coherence or not about their written production. In order to analyze the reliability of student self-assessment about coherence, the next two things were conducted. Firstly, the coherence was analyzed in terms of moves between themes (topics), and then theme (topic) development and comments was analyzed from the viewpoint of the relationship between theme (topic) and comments.

Secondly, teachers assessed the coherence of students' composition based on the hierarchy of moves between concepts (see Figure 1). And then, they judged whether each composition was coherent or not. If teachers evaluated the coherence of the composition was incomplete, it was independently scored 2 points (of full mark 3 points) by each teacher. As the correlation coefficient between two teachers was significant, we can say that it is reliable (see Table 7).

Table 19 shows the patterns of the hierarchy of theme (topic). Concerning theme (topic) development, the types of the topic development were categorized into three types by two teachers; 1: Coherent, 2: Not so coherent, 3: Not coherent, and teachers assessed written production based on the hierarchy map of topic development.

Table 19. The Hierarchy of Theme (Topic)

Scores	Category	Content
3	Coherent	Specific or supporting comments are well arranged and show a key concept.
2	Not so coherent	Difficult to find theme and comments because specific or supporting comments are not arranged well.
1	Not coherent	Very difficult to find theme and comments because specific or supporting comments are not arranged well.

For instance, the composition of Example 1 presented coherence. In the center of theme (topic), four main comments are located as Figure 2 presents. The key concept, "Asakusa" has four comments and each comment also has

comments which are linked with each other. Moreover, the key concept is restated in the conclusion.

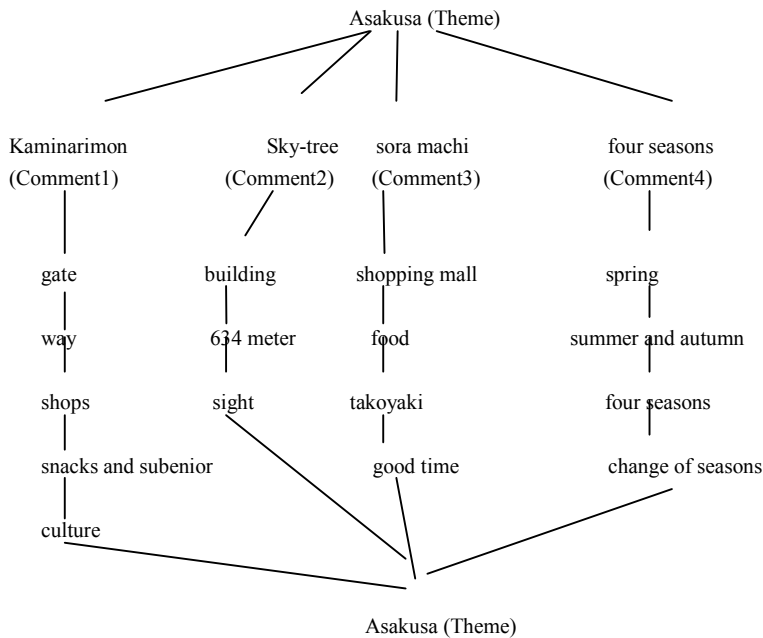


Figure 2. The relationship of theme and comments of Example 1

On the other hand, the next example (Figure 3) shows “not so coherent composition” (see Appendix C). It is difficult to find a key concept and also linkage between comments, the writer presents the main idea just before the conclusion. However, the reader is demanded to have an imagination to connect each comment together.

Table 20 presents the ratio of the relationship of coherent or not coherent compositions. 48.7 percent of all of the 82 students were evaluated as coherent compositions. The compositions which were written by high scored students were more coherent. On the other hand, low-scored groups wrote many incoherent compositions. Further, the quality of each comment demanded the readers’ imagination to understand the relationship between theme and comments.

Table 20. Distribution of the Hierarchy of Theme (topic)

	Scores	3 points	2	1
82 students	82	40/82=.487	33/82=.402	11/82=.134
High-scored	27	19/27=.703	8/27=.296	0/27=.000

Self-Assessment of Topic Development in
Written Production among High School Students

Middle-scored	27	12/27=.444	11/27=.40	4/27=.148
Low-scored	30	9/30=.300	14/30=.466	7/30=.233

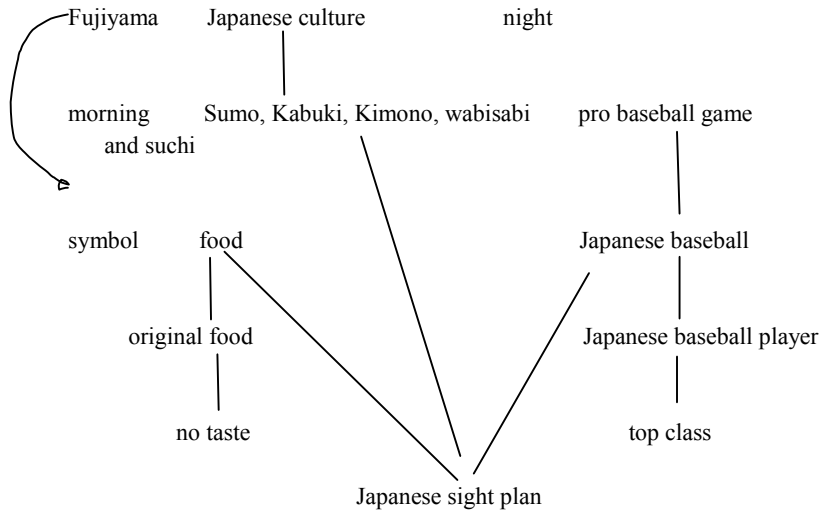


Figure 3. The relationship of theme and comments of Example 2

Table 21 shows the result of the Wilcoxon signed rank test of coherence between teacher assessment and student self-assessment. It was found that students assessed coherence differently from teachers.

Table 21. The Wilcoxon Signed-rank Test between 82 Students' Self-assessment and Teacher Assessment of Coherence

Student self-assessment and native English teacher		Student self-assessment and Japanese English teacher	
T	z	T	z
104.00	-4.318	127.00	-4.121

$p < .05$ T (the sum of ranks of the smaller of the two ranks); z (Wilcoxon signed-rank test z score); < 1.96

4.5 Discussion

4.5.1 Hypothesis 1: Students could assess cohesive devices such as discourse markers as well as teachers do.

The result of the present study says that students showed the consistency of self-assessment and teacher assessment about cohesive devices except low scored students. Besides, the ratio of frequency of cohesive devices was not proportional with the level of scores. For instance, the group of middle scored students showed the least frequency of cohesive devices except in the categories of “Temporal” and “Others”. The category of temporal is related to the overall organization. On the other hand, the category of “Others” plays a role of presenting examples and specific facts. As there are various kinds of cohesive devices, it is assumed that some students could not manage to use them. Further, the usage of cohesive devices is closely connected with the knowledge of language use. So it is also assumed that the students who were short of the vocabulary felt difficulty in using cohesive devices. Besides, the present study did not focus on the accuracy of cohesive devices. There is a possibility to be influenced by cohesive devices. If students used cohesive devices more accurately, the topic development could be different. Therefore, there is a need to survey the accurate usage of cohesive devices, because some students did not accurately use conjunctions and adverbs. The frequency of the other cohesive devices was not proportional to the scores of students, because discourse markers were closely related to the knowledge of vocabulary. In this study, the accuracy of cohesive devices was not examined, so it is difficult to say that the ratio of cohesive devices was related to teacher assessment. Therefore the accuracy of cohesive devices should be examined as a further study.

4.5.2 Hypothesis 2: Students could assess the overall organization pattern of their written production as well as teachers do.

Students could assess overall organization as well as teachers did, however, the consistency of native English teacher and students, and Japanese English teacher and students showed a different result in each part. Therefore, we can say that students could assess the whole organization but are not good at assessing specific components. It seemed to be comparably easier for students to write the Introduction, because students have already understood the role of Introduction which shows readers a topic. Concerning overall organization, most of the students also understood the ideal structure of written production as Pattern 8 has the biggest ratio of all the patterns (see Table 18). However, it seemed to be more difficult for students to organize Body and Conclusion, since these components supply supporting details to broaden a topic which is presented in Introduction. It is also needed to devise and use appropriately cohesive devices to make their written production persuasive and impressive. The size of vocabulary and techniques of its usage of high school students are limited to organize written production, so students needed to manage them adequately within their sphere. Students could not assess Body and Conclusion because they did not have confidence in the

accuracy and usage of cohesive devices. Therefore, the consistency between teacher assessment and student self-assessment was not found in Body and Conclusion.

4.5.3 Hypothesis 3: Students could assess the coherence of their written production as well as teachers do.

Students could not assess the coherence of their written production as well as teachers did. It was expected to give students a chance to write the relationship between topic development. For students are expected to observe their writing objectively. If they could view the flow of theme (topic) at the first glance, it would encourage students to revise their written production. The self-revision needs high ability, but it might encourage students to assess themselves. If students are trained, they might be connected with the self-devised correction. It finally leads to self management. Low-scored students' writing is simpler than others. So it might be easiest for the students to revise and assess their written production.

5 Conclusion

It was basically difficult for students to assess Topic Development except overall organization. Therefore, a study should be devised how to assess other components. Students should also be aware that the assessment sheet could be used as a revision sheet. Temporary cohesive devices might be linked with overall organization because high-scored students used temporary cohesive devices more frequently. So it is a good idea for low-scored students to learn how to use temporary cohesive devices. The researcher's previous study says that the predictor of teachers' evaluation about oral production is the ratio of accurate T-units, so the accuracy of cohesive devices should be researched as a further study. In conclusion, writers are expected to be aware of the presence of readers. So, written production should be persuasive to let readers understand the writer's idea clearly. It is inevitable for us to communicate with people who have different backgrounds influenced by the globalization. Therefore, writers should be trained to be better writers to communicate with people from different cultures. In that sense, self-assessment helps writers to make them reflect themselves and revise the written production. It is efficient for students to improve their written production by self-assessment.

References

Bachman, L. F., & Palmer, A. S. (1981). The construct validity of the FSI oral interview. *Language Learning*, 31, 67-86.

- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education, 18*, 529-549.
- Cheng, Y. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12*(2), 254-255.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Crocker, A. C., & Cheeseman, R. G. (1988). The ability of young children to rank themselves for academic ability. *Educational Studies, 14*(1), 105-110.
- Educational Testing Service. (2008). TOEFL test integrated writing rubrics. Retrieved June 18, 2013, from the World Wide Web: <http://www.ets.org/>
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Edinburgh: Longman.
- Halliday, M. A. K. (2014). *Halliday's introduction to functional grammar*. London: Routledge.
- Harder, B. D., & Katz-Harder, H. (1982). Cultural interface and teaching English composition in Japan. *The English teacher's magazine, 7*, XXXI, 4 (pp. 19-23). Tokyo: Taishu-kan.
- Hinds, J. (1990). Inductive, deductive, quasi-inductive: Expository writing in Japanese, Korean, Chinese and Thai. In U. Connor & A. M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 87-109). Alexandria, VA: Teachers of English to Speakers of other Languages (TESOL).
- Hobbes, J. R., & Evans, D. (1979). Conversation as planned behavior. Technical Report 203, Artificial Intelligence Center, SRI International, Menlo Park, CA.
- Kaplan, R. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*, 1-20.
- Kent, D. (1980). Self-grading versus instructor grading. *Journal of Educational Research, 73*(4), 207-211.
- Kubota, R. (1998). An investigation of Japanese and English L1 essay organization: Differences and similarities. *The Canadian Modern Language Review, 54*, 475-507.
- LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly, 19*, 673-87.
- Little, D. (2002). The European language portfolio: Structure, origins, implementation and challenges. *Language Teaching, 35*(3), 182-189.

Self-Assessment of Topic Development in
Written Production among High School Students

- Luoma, S., & Tarnane, M. (2003). Creating a self-rating instrument for second language writing: from idea to implementation. *Language Testing*, 20, 441.
- Nakano, M. et al. (2003). Measure of English grammar (MEG). *SEISAKU KAGAKU*, 10(3), 59.
- Nakano, M. (2001). *A textbook for English writing*. KENKYUSHA.
- Oi, K. (1986). Cross-cultural differences in rhetorical patterning: A study of Japanese and English. *The Japan Association of College English Teachers* 17, 23-48.
- Oi, S. Y. (2012). *A study of student evaluation and teacher evaluation*. Unpublished master's thesis. Waseda University, Tokyo.
- Oi, S. Y. (2013). A pilot study of self-evaluation and peer evaluation. *Selected Papers of the 17th Conference of Pan-Pacific Association of Applied Linguistics*, 1-11.
- Oskarsson, M. (1989). Self-assessment of language proficiency: Rational and applications. *Language Testing*, 6(1), 1-13.
- Peirce, B., Swain, M., & Hart, D. (1993). Self-Assessment, French immersion, and locus of control. *Applied Linguistics*, 14(1), 25-42.
- The Ministry of Education, Culture, Sports, Science and Technology. (2011). Designated for course of study for senior high schools. Retrieved June 27, 2013, from the World Wide Web: http://www.mext.go.jp/a_menu/shotou/new-cs/youryou/eiyaku/1298353.htm
- van Dijk, T. (1980). *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum.
- Witte, S. (1983). Topical structure and revision. *An Exploratory Study*, 34, 313-341.

Yoko Suganuma Oi

Waseda University, Graduate School

1-6-1, Nishi-Waseda, Shinjuku-ku, Tokyo, 169-8050, Japan

Tel: +81 3 3203 4141

E-mail: yokosuganuma@suou.waseda.jp

Received: April 8, 2014

Revised: July 18, 2014

Accepted: July 20, 2014

Appendices

Appendix A

Components	Scores				
Introduction (present a topic) (導入) トピックの提示があるか	3	2	1		
	Good	So so	Bad		
Body (develop a main idea) (本論) 主要な考えがまとめられているか	3	2	1		
	Good	So so	Bad		
Conclusion (restate a main idea) (結論)	3	2	1		
	Good	So so	Bad		
Discourse Markers (usage of conjunctions and other expressions to develop a topic) 話題の発展や話題のつながりのために接続詞や副詞などが使われているか	3	2	1		
	Good	So so	Bad		
Coherence 英作文全体にまとまりがあるか	3	2	1		
	Good	So so	Bad		
General Evaluation	5	4	3	2	1

Appendix B

Pattern NO	The characteristics of Topic Development
1	Too few sentences. There are fewer than three lines.
2	It does not follow the task or direction of written production.
3	There is a topic sentence in introduction. The body and conclusion do not discuss the topic. There is no coherence.
4	There is a topic sentence in the introduction. Topic development is somewhat found but discourse makers such as conjunctions and adverbs are not adequately used or are not found. The main idea is not restated in the conclusion.
5	There is a topic sentence in the introduction. The composition shows topic Development but discourse markers such as conjunctions and adverbs are not adequately used or are not found. The main idea is not restated in the conclusion.
6	There is a topic sentence in the introduction. The composition shows topic development is found but discourse markers such as conjunctions and adverbs are not adequately used or are not found. The main idea is restated in the conclusion.
7	There is a topic sentence in the introduction. The composition shows topic development. Discourse markers such as conjunctions and adverbs are adequately used. The main idea is not restated in the conclusion.
8	There is a topic sentence in the introduction. The composition shows topic development. Discourse markers such as conjunctions and adverbs are adequately used to develop the topic. The main idea is restated in the conclusion.

Self-Assessment of Topic Development in Written Production among High School Students

Appendix C

Example 2

1. Fujiyama is most favorite place in Japan.
2. So, I will go there with a foreign visitor at morning.
3. Fujiyama is Japanese symbol, and admired world records last month.
4. After the times, My plan is a foreign visitor touch the Japanese cultures.
5. For example, Sumou, Kabuki, Kimono, wabisabi and Sushi.
6. Sushi is famous food in the world.
7. Supecially, "Wasabi. is Japanese original food.
8. It has no taste in the world.
9. After the night, I taked pro baseball games.
10. Japanese baseball level is very hire.
11. because Team Japan was vitory in the World Baseball Clasis.
12. Japanese baseball is the most speedy and techic in the world.
13. Famous Japanese baseball player, Ichiro, Matsui, and Darubish.
14. They has a supecial techicat top class.
15. They are my Japanese sight plan.
16. A foreign visitors will be happy.