

## **Increase in Syntactic Knowledge Using English News and its Correlation with Development of Reading and Writing Skills**

Kazumi Tsutada  
Kyoto Sangyo University

### **ABSTRACT**

On the assumption that grammar acquisition is a crucial part of English education in Japan, this paper proposed and conducted a weekly grammar learning exercise using authentic English news with the intention of learning grammar in a real-life discourse. The participants were 45 university students with their TOEIC scores ranging from 550 to 700, and the training was conducted throughout a single semester. Results showed noticeable increase in grammatical knowledge. Then with a view to fostering practical English users, this study examined whether acquired grammatical knowledge could be successfully transferred to practical reading and writing skills (in terms of complexity, accuracy, fluency, and criticality). Practical skills largely showed improvement with medium to large effect sizes except for writing accuracy. Lastly correlation between rise in grammatical knowledge and that in each factor of reading and writing skills was examined. It showed, although below the expectations, moderately linear development among all factors. It is hoped this study will contribute to further consideration of the importance of grammar for Japanese learners of English, its effects on practical skills, and a more detailed correlation among all the factors.

**Keywords:** syntactic knowledge, English news, reading and writing skills, correlation

## **I. INTRODUCTION**

### **1. Background**

With further complications in international relations and increasing ethnic diversity in the Japanese domestic situation, the need for English skill has become far more emphasized for Japanese people. However, there exists a persistent notion that English is a high barrier, difficult to overcome. It has been argued how English pedagogy in Japan should be changed, with no fruitful outcome so far. Many teaching methods have been introduced from overseas amid the general tendency that the transition to communicative approaches introduced from Western nations will lead to the development of communicative skills. For Japanese learners of English, however, it has been proven otherwise so far, which is mainly due to the farthest language distance between English and Japanese claimed by Elder and Davies (1998). In this regard the Japanese should develop their own unique methods and approaches to compensate for language difference. Now is the time to consider specifically what measures should be taken and how they be put into practice, so that the Japanese can play a key role as English users in the international arena.

In addition it is worth investigating how English education at school has been reflected on social and global activities in a society, that is how industry-university liaison has been currently actualized. With increasing ethnic diversity in Japanese companies, there is more need for use of English. Among others, according to the industrial survey, reading and writing are considered the most frequently used skills (Tsuji, 2014), and outputting by writing is significantly important (Kurosaki, 2014). Matsuda (2010) states the advent of the Internet and various new technologies for written communication has greatly increased opportunities and needs for international written communication. In the company survey above on what should be done at universities, significance of authentic English education using current news (Tsuji, 2014), and of written communication skill which successfully conveys one's thought (Kurosaki, 2014) has been brought to the fore. Incidentally in a framework of Basic Interpersonal Communication Skills (BICS) or conversational fluency, and Cognitive Academic Language Proficiency (CALP) for demanding tasks beyond carrying out daily conversation (Cummins, 1979), Krashen and Brown (2007) claim that "... many second and foreign language programs have recognized the importance of academic language proficiency and consider it to be a central goal of language teaching program" (p. 1).

## **2. Research objectives**

First, with a view to the need of locally appropriate English teaching method, this study considers the acquisition of grammar the indispensable aspect of English acquisition for Japanese learners of English, and intends to design and practice a unique grammatical training targeted at university students. Although grammatical knowledge contains many factors that are deeply interconnected, grammar in this study refers to the structure of a language (syntax). Therefore, 'grammatical' refers to 'syntactic' in this paper.

Another focus of this study is using English news as grammar teaching materials. Conventional grammar textbooks often contain example sentences unnatural and far from authentic communication, and their meanings are processed on a sentence-level, not a discourse-level. As English is to be learned as a tool of actual communication, even grammatical items need to be studied in real discourse. Based on these assumptions, this study, with the intent of contextualization of grammar learning materials, uses English news in which real-world communication is interwoven with what is actually happening in the world. This is anticipated to not only heighten participants' motivation to learn grammar, but raise their global awareness as global citizens. While not a few researchers have introduced teaching methods using "spoken" English news, this study explores the efficacy of "written" English news, through which specific grammatical items could be learnt with due attention.

Third objective is to examine how the grammatical knowledge (as declarative knowledge) learned through the grammatical training is transferred to practical reading and writing skills (as procedural knowledge), and a correlation of development between grammatical knowledge and reading and writing skills. For this purpose this study conducts pre- and post- grammatical tests, pre- and post- reading and writing tasks, examines the improvement in each area, and then explores

the correlation between the development of syntactic knowledge and that of reading and writing skills. It is the author's interest to find whether the more linguistically knowledgeable a learner becomes, the more proficient he/she becomes in reading and writing. So far research has been made on correlation between specific English skills as cross-sectional study. The present study, however, intends to bear significance in exploring the correlation between the development of syntactic knowledge and that of practical reading and writing skills on a longitudinal basis.

Research questions presented in this study are (1) Does grammatical training using English news enhance grammatical knowledge? and (2) What is the correlation between rise in grammatical knowledge and development of reading and writing skills?

The ultimate goal of this study is to facilitate the equipping of university students with English as a tool, and send them into international society with adequate power and confidence as a global citizen.

## **II. LITERATURE REVIEW**

### **1. Grammar teaching in SLA**

Amid the currently prevalent assumption that English ability equates to English speaking ability, it is true that grammar teaching seems gradually receding from the central scene. In addition there is a tendency to view grammar teaching as a factor conflicting with the fostering of communicative competence. However, many researchers address the importance of grammar teaching in SLA. Ur (1999) claims that grammar teaching is beneficial, "provided it is taught consistently as a means to improving mastery of the language, not as an end in itself" (pp. 77-78). It is also significant to lend our ears to Wilkins, one of principal advocates of communicative language teaching, who states that "grammar is the means through which linguistic creativity is ultimately achieved and an inadequate knowledge of the grammar would lead to a serious limitation on the capacity for communication" (Wilkins, 1976, p. 66).

Based on all this research, this study intends to design and study a grammar teaching method that hopefully goes beyond the traditional prescriptive rule-teaching.

### **2. Effects of grammar on reading and writing**

Grammar, which is described as a 'force of expression' (Knapp and Watkins, 2005), is assumed to play a key role both in reading writing, especially for Japanese learners of English. Grabe (1991) claims syntactic knowledge is a critical component and has a significant effect in reading comprehension. Meanwhile, Matsuda (2010) refers to the relationship between grammar and writing fluency and accuracy. He says fluency requires some level of grammatical competence, and accuracy without some degree of fluency is of little use (p. 16). These research could justify the grammar teaching targeted for fostering better readers and writers.

One thing to note is that writing should not be a means of grammar exercises, in which only sentence-level structural skills are examined. Matsuda (1999) claims 'second-language writing should be seen as an integral part of both composition studies and second-language studies' (p. 789).

Accordingly, this study aims to develop students' original paragraph-level as well as sentence-level writing skills based on structural skills to be enhanced throughout the semester.

### **3. Focus on form**

The traditional focus on forms instruction has been a primary focus in English pedagogy in Japan. This approach has been proved to effectively lead to higher scores in discrete-point language test. However, Norris and Ortega (2000) found the effectiveness of such instruction to be considerably low when it comes to learners' ability to use the language in actual communication. Furthermore, they state that leaving learners in the contrived rule-governed learning environment could easily lead to demotivation. Based on these claims, it may be necessary to learn grammar in a contextualized context focusing both form and meaning. Doughty & Williams (1998) notes that "primary concern of the teacher should always be the question of how to integrate attention to form and meaning" (p. 261). This study intends to make an attempt to focus the learners' attention to forms in news articles as meaningful contexts.

### **4. Explicitness**

Whether or not the teaching of explicit knowledge has value continues to be a controversial issue. Krashen (1981) in the noninterface position claims explicit and implicit knowledge are completely separate, and explicit knowledge can never be converted into implicit knowledge, hindering rapid and fluent communication. Meanwhile, Ellis (2006) claims that "explicit knowledge can convert into implicit knowledge if the learner is ready to acquire the targeted feature" (p. 97).

Acknowledging that the ultimate goal of language learning is to use it unconsciously and automatically using implicit knowledge, this study holds the position that first priority for Japanese English learners is on building up foundations of grammatical skills through accumulation of explicit knowledge. Swan (2006) asserts "explicit teaching and practice can help students along the rocky road towards reasonably correct production" (p. 13). In the present experiment, all the relevant grammatical items were explained explicitly in Japanese whenever it's necessary while discussing grammatical points after each grammatical test conducted every week.

### **5. News as authentic materials**

Recent years have seen a growing interest in the use of authentic teaching materials, which has been partly advocated by Newspaper In Education (NIE). Yashima (2011) has claimed that having interest in international matters in general should lead to higher motivation to learn and use English, which she calls "international posture". She states that one with a developed international posture is generally more conscious of his or her existence in the world, which is likely to encourage willingness to communicate in English under rapidly globalized circumstances (*ibid*). It was anticipated that students would heighten their motivation to learn English through reading a variety of overseas and domestic news, and constructing their views in writing about them.

In addition to the studies above, it can be said that news is an everyday example of authentic

output providing a variety of English styles, which contrasts with conventionally designed textbooks whose “repertoire is thin and flat” (Hwang, 2005, p. 3). Hwang states “when learners truly enjoy authentic materials ... they could be gradually ‘hypnotized into’ the rhythm and pattern of the target language” (ibid, p. 8).

Furthermore it naturally reflects the variability in the use of English as the language is constantly changing. Berardo (2006) states that “[news provides] real language exposure with language change being reflected ... textbooks tend to become outdated very quickly” (p. 6).

It also seems adequately significant to support the use of English news from the aspect of English for General Purposes (EGP) and English for Specific Purposes (ESP), which have been dichotomously debated in the history of Japanese English education, and for which favor continues to swing alternately like a pendulum. ESP has drawn more attention owing to arguments for its ability to meet the immediate needs of learners and potential to generate higher motivation for learning. Many university students, however, find it difficult to decide their future paths, and survey results collected in 2010 by the Ministry of Health, Labor and Welfare<sup>1</sup> show a high turnover rate (31%) of university graduates across industries (within three years after graduation). Considering this situation, English news could be a sensible choice as teaching materials, because news is presumably relevant to everybody and furnishes them with empirical knowledge necessary as social citizens regardless of their professions. This is another strong reason this study uses English news as teaching materials.

When it comes to previous empirical studies on the use of English news, most have focused on “spoken” language. Among the most notable empirical studies on the use of audio English news is Hino’s Integrated Practice in Teaching English as an International Language (IPTEIL) (Hino, 2012a), which has successfully improved students’ practical English skills through authentic use of real time English. The present study, however, deals with “written” language, with a foundation placed on reading. It is postulated that learners will deepen their understanding of sentence structure within meaningful context, which is rather difficult to attain through listening and speaking activities. Amid accelerated alienation of the younger generation from written language, the role of reading and writing remains critical, particularly for retrieving information from Internet and communicating via e-mails. In addition intensive reading has customarily been the primary approach in English teaching in Japan, and as a result tends to be the most comfortable way of learning grammar for Japanese students.

Sufficiently supported by all the above-mentioned theories and proven methods, this study aims at efficient and practical use of English news as grammar teaching materials.

### **III. METHODS**

#### **1. Participants**

The experiment was conducted in three compulsory English classes at a university, which were mainly aimed to improve TOEIC scores as a part of acquiring communication skills. The data was compiled from a total of 45 first-year male and female students whose majors vary both in the

humanities and the science courses. Their English level ranged between 550 to 700 points for the TOEIC. Permission to use the data collected was given from all of the students.

## **2. News as teaching material**

### **2.1 Topic familiarity**

As not to distract their attention from the relevant grammatical points, their familiarity with the content should be a top priority. Although news selection was ultimately the author's subjective decision, attention was paid to select from the students' viewpoints.

In addition to this familiarity, the articles were chosen from those released within two weeks before each class, so that the students could deal with them while still having a conceivably fresh memory of the news stories. In terms of suitability as classroom materials, excessively controversial political issues were excluded on the basis of equal appropriateness for all learners.

### **2.2 Readability**

To avoid disturbing the developmental path of their grammatical knowledge by constant fluctuation of difficulty level of the weekly articles, this study employed the Flesch Reading Ease (FRE)<sup>2</sup> scale to help align the difficulty of each article. Selection criteria was Flesch scale 45 – 55, which are evaluated between “Fairly Difficult” and “Difficult”, which refer to English native high school students and university students respectively. Accordingly those within this range seemed a little challenging for the present intermediate-level participants; however, based on Krashen's “*i + 1*” theory which claims that we acquire linguistic competence when input is a little beyond the current level (Krashen, 1981), they seemed acceptable as appropriate as teaching materials for the current participants. Krashen claims that language is understood “with the aid of extra-linguistic context” (Krashen, 1981, p. 103).

### **2.3 Grammatical items**

Since there is no grammatical syllabus or guidelines for instruction, it is wholly the teacher's responsibility to organize grammar instruction including selection of useful grammatical items to be explained explicitly and the amount of time to spend on explanation of each item. In this study not more than three items were explained in a single day to avoid possible confusion or discouragement among learners due to excessive loads.

### **2.4 Simplification**

Due to general difficulty in using authentic materials, many researchers propose reduction and replacement of words, paraphrasing, or simplification of sentence structures. However, as Nuttall (1996) argues “however good a simplification is, something is always lost” (p. 178), such simplification could potentially lead to misunderstanding or distortion of meaning. Accordingly in the current study no simplification was made, and instead necessary scaffolding and pre-reading strategies were applied wherever necessary, for instance by referring to unusual structure that

sometimes emerges in the articles, to help students concentrate on targeted grammatical aspects.

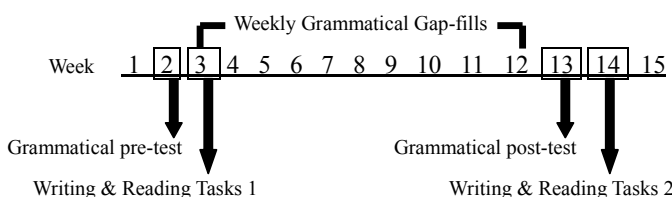
### 3 Procedures and analytic approach

The study was conducted over 15-week period according to Figure 1.

#### 3.1 Grammatical gap-fills

The purpose of this task is to build up syntactic knowledge through English news. It was carried out weekly from week 3 through week 12 using NHK World website<sup>3</sup>. Two printed articles were prepared with ten blanks and four choices to choose from to fill the blanks. Careful attention was taken to blank out those words having a crucial role in terms of syntax according to the objective of this research (See Appendix 1). Specifically targeted grammatical items are subject-verb structure of a principal clause, subject-verb structure of a dependent clause, parallel construction, relatives, conjunctions, participles, and infinitives.

Once students filled in the blanks, they were asked to discuss their answers in groups and prepare to explain the grammatical reason for each answer to avoid overlooking lucky guess. Then the teacher elicited students' answers, and asked them the reason for their choices, regardless of accuracy of their answer and appropriateness of their explanation. This process is meaningful since it is important to have them realize they lack specific grammatical knowledge even when their answers are correct. Finally the teacher gave explicit explanations whenever necessary. When using this method, a teacher should make efficient preparations deciding what information must be explained explicitly and planning a simple way to provide crucial information, especially as no syllabus or guidelines are provided. The weekly transition of scores of gap-fills was illustrated in a table.



**Figure 1. Course schedule**

#### 3.2 Pre-test and post-test

In order to measure the extent of improvement of syntactic knowledge the same grammatical test consisting of thirty questions was conducted (in weeks 2 and 13) as pre- and post-tests as shown in Figure 1. With reference to a few commercially available grammar exercise drills, the author formulated the test to focus exclusively on assessing structural knowledge such as those described in the previous section. Examples are shown in Appedix 2. Same test was used as a pre- and post-tests as it was hardly possible to make two tests with the identical mean score and reliability. Accordingly, full explanation was provided only after the post-test. At the end both test scores were analyzed to see if significant differences could be found.

### 3.3 Reading

This task aims to see if the declarative grammatical knowledge was effectively transferred into reading skill. Through weekly grammatical gap-fills students were continuously encouraged to read by chunks or blocks while paying attention to sentence structures. It was posited that better understanding of sentence structure facilitated chunk reading, which should lead to effective speed reading.

The reading task was conducted in weeks 3 and 14, as shown in Figure 1, with different articles as pre- and post-reading tasks. The teacher prepared four latest articles with 200 to 300 words in length with a total of ten comprehension questions and four alternative answers to choose from. Selected topics were of the latest domestic and overseas news such as Japanese government's education policies, epoch-making technological innovation in the world, Japan's financial contribution to countries in need, and so on. Since this exercise was meant to measure practical and effective reading skill, a time limit was imposed based on 200 wpm. According to Nation (2009), a reasonable goal for second language learners is around 250 wpm for reading materials containing no unknown words or grammar. Based on this theory, 200 wpm seems an appropriate goal for the current participants for reading news articles. In addition, Eskey & Grabe (1988) claim that accurate comprehension is hindered with the reading speed less than 200 wpm. Results of pre- and post-tasks were analyzed, and the correlation between increase in reading scores and rise in syntactic knowledge was examined and discussed.

### 3.4 Writing

This task was also intended to see if the transition of declarative grammatical knowledge to writing skill as procedural knowledge was successful. It is generally assumed that writing largely depends on competence of building up sentences based on correct syntactic knowledge. Hwang (2005) says, "output-based tasks can activate the input stored in receptive memory and thus transform knowledge into skills" (p. 4). Shirai (2012) also claims that output activities need to be incorporated in English learning so that input knowledge can efficiently lead to acquisition. From these viewpoints, this seemed to be a meaningful task, and examination of how receptive skill could be activated into productive skill was of much interest.

The writing task was carried out in weeks 3 and 14 as shown in Figure 1 alongside the reading task. On those two days students were given six articles including two for weekly grammatical gap-fills and four for reading tasks. From those six articles they were told to choose one article, and assigned to write on it for twenty minutes, focusing on expressing their opinions on the article, instead of just summarizing the content, as well as grammatical accuracy. Use of dictionaries and words in the article was allowed, so that students could properly concentrate on syntactical aspect while writing.

With regard to writing evaluation, the specific evaluation approach was used focusing on structural aspects according to the objective of the present study. Three commonly acknowledged measurement factors were used; complexity, accuracy, and fluency with the underlying assumption

that these three aspects of language development progress in tandem, which were specifically assessed by objective indexes shown in Table 1 respectively (Wolfe-Quintero, Inagaki, & Kim, 1998). A T-unit used in this paper is based on Ishikawa (2006) who defined it as “a main clause plus any subordinating clauses” (p. 202). It must also be necessary to define “error” to count as it should vary depending on what is to be measured. In the present research errors in spelling, articles, prepositions, word uses, capitalization and punctuation were not counted as they are not directly related to sentence structure focused on this research. In addition it should be noted that any sentence left unfinished due to time constraint at the end of the writing was disregarded.

**Table 1. Writing assessment index by factor**

Factor	Index
Complexity	Clauses / T-unit
Accuracy	Error-free T-units / T-units
Fluency	Words / T-unit

Another factor used was criticality. It was added because production of clear, persuasive, and critical sentences requires adequate structural skill to be employed. It was the author’s assumption that the more proficient the structural skill is, the more critical writers could be either affirmatively or negatively. As this is subjective evaluation, it was conducted by two raters, a native speaker of English who has worked as a professional proofreader in Japan and the present author, based on the 5-scale original rubrics made by the present author (Appendix 3). Students were informed in advance that their writing was to be evaluated based on these criteria. At the end the results were analyzed, and the correlation between the rise in syntactic knowledge and the development of these four writing assessment factors (complexity, accuracy, fluency, and criticality) was examined.

### 3.5 Final questionnaire

At the end of the course a final questionnaire was collected to directly listen to students’ voices, in which students were asked the following three questions:

- Do you think English grammar is important for practical communication?
- What is your motivation toward further grammar learning?
- Which do you prefer as grammar learning materials, conventional grammar textbooks or English news?

In addition a group interview was arranged with five students after the last class, where the author could hear their honest opinion through face-to-face conversation, difficult to collect with formally constructed questionnaire.

## IV. RESULTS AND DISCUSSION

### 1. Grammatical gap-fills

The results of weekly grammatical gap-fills, which had played a key role in building up syntactic knowledge in this study, are shown in Table 2.

**Table 2. Scores of grammatical gap-fills ( $N = 45$ )**

Week	3	4	5	6	7	8	9	10	11	12
<i>Mean</i>	4.64	5.82	6.58	6.8	6.36	5.2	6	6.62	7.93	7.8
<i>SD</i>	1.29	1.31	1.22	1.37	1.21	1.4	1.39	1.57	1.01	0.97

*Note.* With a full score of 10

The transition of scores deviated from the original expectations of steady improvement to the course end, with the increase of only 0.4 from week 3 to week 8. Meanwhile a rather stable rise in scores was seen in the last four weeks, resulting in the final increase of 3.16 throughout the course. Three main reasons seem to attribute for the unstable rise and fall for the earlier weeks. First, as the students were not used to reading English news, their affective filter was high, making them feel nervous and uneasy in dealing with grammatical points in the news. Second, sentences in the articles issued are generally longer than those they usually deal with in grammar exercise textbooks. It is natural to imagine that they needed certain time to get accustomed to analyzing such long sentence structures in the news. Third, even relatively advanced students obviously lacked basic grammatical knowledge due to inadequate study of grammar in high school. Some students even stated that they had learned nearly no grammar as their English classes mostly focused on speaking activities. Other students admitted they lost their grammatical knowledge after the entrance examinations, thus it seemed to take a while to recall what they learned previously.

However, the general increase in scores throughout the course could be attributed to the effects of intensive reading focused on structural analysis and accumulation of explicit grammatical knowledge given by the teacher. What was most obvious was changes of students' attitude toward this task. They more enjoyed discussing grammar in real-life discourse and showed enhanced willingness to solve their grammatical problems they encountered in the news articles. This was considered the desirable outcome of enhanced international posture and of focus on form instruction.

## **2. Pre- and post-grammatical tests**

Pre- and post-grammatical tests were conducted to examine the development of syntactic knowledge over the course duration. Table 3 shows descriptive statistics, which indicates a rise in mean of 4.34.

Table 4 shows the results of dependent *t*-test. Cohen's *d* in Table 4 is the standardized mean difference between the two tests. According to Mizumoto and Takeuchi (2008), Cohen's *ds* of .20, .50, and .80 are considered small, medium, and large, respectively. With an effect size of *d* = .90, the results shown demonstrate that there was a significant difference between the two tests with a large effect size. This result demonstrates actual development of syntactic knowledge. For *p* value  $\alpha$  level was adjusted to .008 (.05/6) by Bonferroni correction as *t*-tests were used for six factors in this study.

Increase in Syntactic Knowledge Using English News  
and its Correlation with Development of Reading and Writing Skills

**Table 3. Descriptive statistics of pre-test and post-test**

	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>Skewness</i>	<i>Kurtosis</i>	<i>α</i>
Pre	45	15.64	5.24	6	27	.2	.93	.81
Post	45	19.98	3.99	8	29	-.34	.47	.71

*Note.*  $\alpha$  is Cronbach's coefficient alpha.

**Table 4. Results of paired *t*-test for pre-test and post-test**

<i>Mean difference</i>	<i>t</i>	<i>df</i>	<sup>a</sup> <i>p</i>	<sup>b</sup> <i>d</i>
-4.33	-7.01	44	<.001	.90

*Note.* <sup>a</sup>  $\alpha$  level was adjusted to .008 by Bonferroni correction. <sup>b</sup> *d* (Cohen's *d*)

It should be noted that the enhancement of syntactic knowledge is also attributed to many other causes and circumstances such as learning in other classes or self-study. However the positive results of post-test show that execution of weekly grammatical gap-fills more or less contributed to the development of their syntactic knowledge. Stated another way, learning grammar through English news likely had a significant positive effect on the development of syntactic knowledge.

### 3. Reading

The purpose of this task was to compare how well students were able to grasp the content of articles by speed reading, before and after grammatical training. Descriptive statistics shown in Table 5 show that there was a large difference in reading comprehension of ten points among the participants in the pre-task, which was narrowed to five points in the post-task, raising the mean by 3.45 points. Figures in Table 6 indicate there is significant difference with a large effect size, which shows that in general the participants became more skilled in speed reading with improved accuracy.

It is worth considering why their speed reading skill improved. First, it is assumed that the improvement of reading skill is a corollary effect of acquisition of grammatical knowledge throughout the course. This is because it is generally admitted that effective speed reading could be facilitated by slash or chunk reading, as well as other specific reading techniques such as no regression or picking up content words. In this sense, specific exercises to analyze sentence structure done by weekly grammatical gap-fills contributed to nurturing able readers. It could be possible to say acquired grammatical knowledge was put into practical use in reading news for some successful readers. This is compatible with Grabe's claim grammar plays a key role in reading (Grabe, 1991). On the other hand it should be noted that there were a few who failed to improve their reading scores. In fact some students were unable to read the articles to the end both in the pre- and post-tests because they spent too much time analyzing each sentence with their limited syntactic knowledge, resulting in lower reading scores. It could be explained by their lack of motivation toward reading or possible confusion in understanding sentence structure.

**Table 5. Descriptive statistics of pre- and post-reading**

	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>Skewness</i>	<i>Kurtosis</i>	$\alpha$
Pre	45	4.73	2.47	0	10	.12	-1.07	.68
Post	45	8.18	1.17	5	10	-.17	-.33	.72

Note.  $\alpha$  is Cronbach's coefficient alpha.

**Table 6. Results of paired *t*-test for pre- and post-reading**

<i>Mean difference</i>	<i>t</i>	<i>df</i>	<sup>a</sup> <i>p</i>	<sup>b</sup> <i>d</i>
-3.44	-10.31	44	<.001	1.65

Note. <sup>a</sup>  $\alpha$  level was adjusted to .008 by Bonferroni correction. <sup>b</sup> *d* (Cohen's *d*)

#### 4. Writing

It was of much interest to examine how the participants' writing skill was enhanced owing to the accumulation of explicit syntactic knowledge. Writing skill development according to each specific factor previously defined in this study is shown in Table 7, which shows more or less certain improvement in all those factors.

**Table 7. Descriptive statistics of pre- and post- writing by factor**

		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>Skewness</i>	<i>Kurtosis</i>
Complexity	Pre	45	1.39	.18	1.11	1.83	.71	-.25
	Post	45	1.68	.20	1.18	2.17	-.09	-.27
Accuracy	Pre	45	.78	.12	.5	1	-.55	-.32
	Post	45	.81	.11	.56	1	-.1	-.65
Fluency	Pre	45	10.81	1.46	7.85	13.90	.09	-.65
	Post	45	12.62	1.72	9.12	16.75	.26	-.3
Criticality	Pre	45	3.14	.69	2	4.5	-.21	-.91
	Post	45	3.51	.55	2	4.5	-.45	-.31

Note. Criticality in the table refers to the mean of assessment by raters 1 and 2.

All data was analyzed by paired *t*-test with the results shown in Table 8, which indicates varied results according to the factors. First, the improvement of complexity could be explained by the fact that the participants became more interested in incorporating dependent clauses at the post-writing task, as teaching of syntactic knowledge regarding complex sentences had been constantly focused throughout weekly grammatical gap-fills, whereas there were less conscious efforts to construct complex sentences at the pre-task. It was obvious students better understood how to utilize complex sentences effectively in their post-task compositions. In particular, "if" for conditional clauses and subjunctive clauses, and "as" and "because" to describe reasoning, were used more often. In addition, "that-clauses" such as "I think (that)...", "they say (that) ...", or "... made a proposal (that) ..." were employed frequently, which contributed to a rise in complexity ratio. Furthermore advanced students successfully incorporated the secondary dependent clause in the primary dependent clause. It can be considered one of the positive effects of studying with English news as such complex sentences are commonly used in the articles. For less skilled learners it still seemed difficult to use dependent clauses with adequate success. It was often the case that

they started with subordinate conjunction, but ended with no main clause. Despite their attempt to use their acquired knowledge about subordinate conjunctions, they were not adequately equipped with a sufficient level of performance to put them into practical use in proper manner. For those students most of the sentences were compound sentences, which were often connected by “and” and “but”, instead of complex sentences.

With regard to accuracy Table 8 indicates that no significant difference in the accuracy development was found after Bonferroni correction with small effect size, which was contrary to the expectations of linear progress in all writing assessment factors. It is worth examining students’ writings for both pre- and post-tasks. As described earlier many of the students were less conscious of building up complex sentences at the onset of the course, therefore many of the sentences were compound sentences connected by coordinate conjunctions or with no conjunctions, which resulted in fewer mistakes in syntactic aspect. Meanwhile, post-task writings revealed many attempts to use complex sentences in compositions, but some were not properly executed. Some specific errors were due to the paucity of full understanding of difference between subordinate and coordinate conjunctions. From the author’s teaching experience the inability to distinguish between these two types of conjunctions generally tend to remain until learners reach the fairly high level of proficiency. It held true with the present study. Most frequent misusages were those among “but”, “however”, and “although”. It was supposed that even when they understood the difference fairly well as linguistic knowledge, they often were unable to transfer this knowledge into productive skill. Another common mistake was found in the usage of conjunctive adverbs such as “moreover”, “therefore”, and “yet”. Most of the time students realized their mistakes instantly when given feedback from the teacher afterwards. This suggests that they merely failed to utilize learned grammar correctly in spite of their understanding of the syntactic rules. It is assumed this phenomenon of lowering accuracy could also be explained by the trade-off hypothesis that committing attention to one area could cause lower performance in others. That is, writers’ consciousness toward writing more complex and long sentences, which are evaluated by complexity and fluency respectively, presumably resulted in lowering their accuracy.

Meanwhile it is also worth referring to some of the specific grammatical aspects in the post-task which seemed to have been relatively successfully put into practical use after being learnt. One is the use of relative clauses, which were used appropriately in many post-writings. Another aspect was the subject-verb agreement. Much fewer mistakes were found with the agreement rules in the post-task even when verbs are located rather far from subjects. This cautiousness was considered to be one of the constructive effects of weekly grammatical gap-fills, which had been constantly focused on identifying a subject and verb in longer sentences.

As for fluency it is natural that the number of words per T-unit increases as more dependent clauses are used in general. This applied to skilled writers in particular who used more subordinate clauses in the post-task, increasing the number of words in a T-unit. Less skilled writers on the other hand who were unable to construct complex sentences successfully often produced compound structure even in the post-task instead of complex structure, reducing the number of words in each

T-unit. However, as a whole, significant difference was found with a large effect size.

Finally, as an additional criterion, criticality was assessed by two raters. With the inter-rater reliability of 0.83 and 0.81 for the pre- and the post-tasks respectively it was considered appropriate to use the mean value as criticality assessment. Table 8 shows there was a significant difference between the results of pre- and post-tasks with a medium effect size.

Writing critically, which is regarded one of the crucial elements in terms of media literacy, is a skill considered difficult to attain without secure footing to write well-organized and persuasive sentences based on correct syntactic knowledge. It is the author's assumption that properly structured sentences are more appealing to readers, and it is also easier for writers to convey their thoughts if they are capable of writing well-structured sentences. The fact that this study found certain development in criticality, beyond sentence-level composition skill, is a positive indication that they effectively utilized syntactic knowledge in paragraph-level writing conveying their ideas more clearly and logically. According to Pinker (1994), knowing language is knowing how to express ideas and thought in words. Accordingly criticality can be said one of the meaningful criteria to measure how linguistic knowledge is transferred into productive skill. Meanwhile, it should be added that the rise in writing criticality should also be partially attributable to students' critical thinking skill cultivated through dealing with English news over the entire course, in addition to the development in their performance in writing complex sentences.

**Table 8. Results of paired *t*-test for pre- and post- writings by factor**

	<i>Mean difference</i>	<i>t</i>	<i>df</i>	<sup>a</sup> <i>p</i>	<sup>b</sup> <i>d</i>
Complexity	-.29	-8.69	44	< .001	1.51
Accuracy	-.03	-2.60	44	.01	.27
Fluency	-1.81	-8.16	44	< .001	1.12
Criticality	-.37	-5.24	44	< .001	.57

*Note.*<sup>a</sup>  $\alpha$  level was adjusted to 0.008 by Bonferroni correction. <sup>b</sup>*d* (Cohen's *d*). Criticality in the table refers to the mean of assessment by raters 1 and 2.

Lastly attention must be paid to the increase in the total number of words in their compositions, apart from all the factors mentioned above. One possibility might be students' increased motivation to write. Particularly as they were often advised to be critical thinkers by the teacher, it seemed they were more willing to express their opinions on real-life stories. This tendency is relevant with "international posture" claimed by Yashima (2011), stating that interest in international affairs and consciousness of being a part of the world heightens willingness to communicate either by writing or speaking.

## 5. Correlation

Lastly and most importantly, all the results were examined for correlation of the development using Pearson product-moment correlation coefficient, the principal purpose being to measure the effect of increase in syntactic knowledge (as declarative knowledge) on reading and writing skills

Increase in Syntactic Knowledge Using English News  
and its Correlation with Development of Reading and Writing Skills

(as procedural knowledge). Many researchers have examined the correlation among various aspects of English skill on cross-sectional basis, however little has been done for correlation of longitudinal progress of those aspects. This is possibly another significance of the present study.

The first line in Table 9 indicates the correlation between development of syntactic knowledge and of all the aspects of practical skill, which is directly relevant to the objective of this study.

**Table 9. Correlation among development by assessment factor**

	Syntactic knowledge	Reading	Complexity	Accuracy	Fluency	Criticality
Syntactic knowledge	---	.56*	.76*	.31*	.42*	.56*
Reading		---	.39*	.23	-.01	.16
Complexity			---	.29	.34*	.23
Accuracy				---	.13	.09
Fluency					---	.11
Criticality						---

*Note.* Pearson product-moment correlation coefficient. \*Correlation is significant at the .05 level (2-tailed).

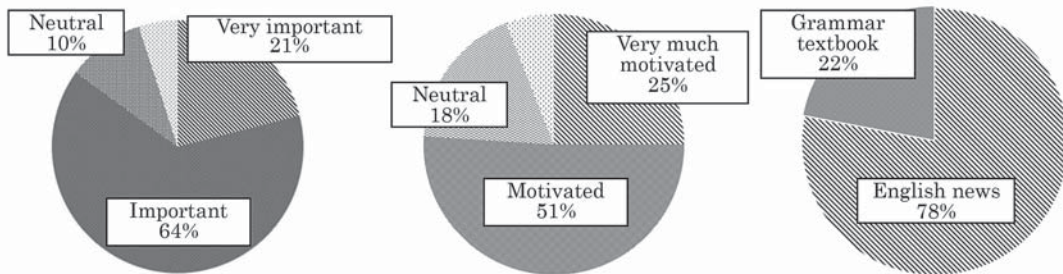
First, with respect to the correlation between rise in syntactic knowledge and development of reading skill, the correlation was 0.56, which was below the author's expectations. It seemed difficult for many of the readers to improve their reading skills proportionally with noticeable rise in syntactic knowledge. Still this medium correlation is a favorable sign that the rise in syntactic knowledge could contribute to the development of reading skill. Concerning writing complexity, the high correlation with improvement in syntactic knowledge could be largely attributable to the fact that most attention had been directed to analysis and construction of complex sentences during grammar instruction in the classroom. It seemed students' enhanced interest in incorporating complex sentences in their post-writing tasks also raised the complexity. Accuracy improvement, which didn't show significant difference between pre- and post-tasks after Bonferroni adjustment, was proved to have low-level correlation with rise in syntactic knowledge. This shows, despite certain improvement in accuracy, the extent of its development did not overtake the notable improvement in syntactic knowledge. It can be said it was difficult to assimilate linguistic knowledge and transfer it into writing accurately over a single semester. This might also be explained by trade-off effect as described earlier, which could cause possible failure of heightening the level of writing performance for all the factors (complexity, accuracy, and fluency) simultaneously. Further empirical study will be necessary to verify the relationship between grammar and writing accuracy. With regard to fluency, the medium correlation was also below the original expectations. This could be partially due to the fact that less skilled writers unintentionally produced compound sentences instead of complex sentences even in the post-writing task, which possibly hindered the proportional development of fluency with improvement of syntactic knowledge. Result of correlation in the area of criticality was actually of much interest, as it seemed an unusual attempt to see if becoming more knowledgeable in grammar could lead to becoming more critical writers. The medium correlation of 0.56 could justify my attempt to measure this

relationship that criticality improves along with a rise in syntactic knowledge.

## 6. Final questionnaire

First pie chart in Figure 2 shows 85% of students thought English grammar is “Very important” or “Important”. The ratio might have been much lower if the same survey had been done at the beginning of the course, as many students initially commented that they felt the communicative approach was far more important. However, in a group interview at the end of the course not a few students said convincingly that communication is not fully effective if they are unable to construct proper sentences, provided that they intend to use English as a social citizen. Others acquiesced a need to know grammar to write well as e-mail writing now plays a vital role in a society. One student who marked “Not very important” frankly stated that trying to construct grammatically correct sentences often makes him feel awkward when using English, and too much attention to grammar might hinder smooth communication. His comment indicates difficulty in making sentences on the spot at this stage. Second, the motivation survey result shown in the second chart in Figure 2 indicated 76% are “Very much motivated” or “Motivated”, which are surprisingly encouraging for the author. Comments from motivated students were “I’d like to continue studying grammar as grammatical knowledge considerably helps improve my comprehensive English skills including reading and writing”, “It was interesting to see grammar rules activated and kept alive in the news”, “Grammar used to be a subject ‘I have to do’, but it is now a subject ‘I want to do’, and others. Lastly it was of particular interest whether learners themselves had found learning grammar with English news useful and significant. The last chart in Figure 2 illustrates 78% chose English news. This result also far exceeded the original expectations. Affirmative comments collected in the group interview include “It was exciting to learn grammar in the real context”, “Learning grammar with English news made me feel confident that my practical English skill was improving”, “It was fun to find the grammar rules I learned at school actually used in more complicated manner.” On the contrary those who preferred grammar textbooks commented “English news often carries too many unknown word, which discourages me from reading it” and “I would prefer English news if my present English level were much higher”. However, those comments, while being rather negative about studying with news, show their desire at the same time to learn living English from the news when their English level reaches a certain level.

# Increase in Syntactic Knowledge Using English News and its Correlation with Development of Reading and Writing Skills



**Figure 2. Results of student survey on importance of grammar (left), motivation toward learning grammar (middle), and text preference (right)**

## **IV. CONCLUSIONS**

On the assumption that the English teaching method should be locally appropriate, this study first focused on acquiring syntactic knowledge, which seems highly important for Japanese learners of English. A unique grammatical training method was designed using English news and implemented with intermediate-level university students over ten weeks. Grammatical gap-fills using news reinforced by explicit explanation of individual grammatical items resulted in a noticeable rise in syntactic knowledge, as shown by the results of pre- and post- grammatical tests. This affirms the first research question. Furthermore, it was found upon completion of the entire course 85% of the participants acknowledged the importance of grammar, and 76% showed motivation in learning grammar. It was also encouraging for the present author that nearly 80% voiced the preference of English news as grammar learning material over conventional grammar textbooks.

Another objective was to explore the effect of syntactic knowledge (as declarative knowledge) on reading and writing skills (as procedural knowledge). First, the results of pre- and post-reading and writing tasks were analyzed by dependent *t*-test, by which significant difference with medium to large effect sizes was found in all factors (reading, and writing complexity, fluency, and criticality) except writing accuracy in this study. Relatively small improvement in accuracy with no significant difference after Bonferroni correction showed the difficulty of correctly transferring acquired linguistic knowledge into writing accurately. This was also partially due to the trade-off effect, which makes it difficult to attain development in the area of complexity, accuracy, and fluency simultaneously. Next, the correlation between rise in syntactic knowledge and development of all the reading and writing skills was examined. Results showed the correlation was high for writing complexity, medium for reading, writing fluency and criticality, and low for writing accuracy. These results were lower than expected as it was originally assumed that all these factors should improve largely proportionally with a rise in syntactic knowledge. However, taking into consideration the complicated factors required to enhance each area of skills and a limited span of a single semester, the results showed a moderate tendency of linear development between syntactic knowledge and reading and writing skills, which answers the second research question.

It is generally considered that acquisition of syntactic knowledge is a lengthy process, and it might not be fully appropriate to evaluate its effects only after a single semester of a university. However, in this kind of empirical study, it might be sensible to think the significance lies in the accumulation of practical teaching, giving birth to better English teaching method. For future study, it is the author's intention to explore the use of other writing assessment indexes as well as reading tasks, and examine more deeply how the increase in syntactic knowledge affects the development of reading and writing skills. It is hoped this research would contribute to further consideration of importance of grammar for Japanese learners of English, its relation with development of practical reading and writing skills, thus helping more Japanese university students be equipped with English as a communication tool to act as a global talent with confidence and assurance.

## ACKNOWLEDGEMENTS

I would like to thank anonymous reviewers for their valuable comments and helpful suggestions.

## NOTES

<sup>1</sup> <http://www.mhlw.go.jp/topics/2010/01/tp0127-2/12.html>

<sup>2</sup> Flesch Reading Ease was developed by Rudolph Flesch and is considered as one of the oldest and most accurate readability formulas.

<sup>3</sup> <http://www3.nhk.or.jp/nhkworld/>

## REFERENCES

- Berardo, S. A. (2006). The Use of Authentic Materials in the Teaching of Reading. *The Reading Matrix*, 6(2), 60-69.
- Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 121-129.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. William (Eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 197-261). Cambridge: Cambridge University Press.
- Elder, C., & Davies, A. (1998). Performance on ESL Examinations: Is There a Language Distance Effect?. *Language and Education*, 12(1), 1-17.
- Ellis, R. (2006). Current Issues in the Teaching of Grammar: An SLA Perspective. *TESOL QUARTERLY*, 40(1), 83-107.
- Eskey, E., & Grabe, W. (1988). Interactive models for second language reading: perspectives on instruction. In P. Carrell, J. Devine, & E. Eskey (Eds.), *Interactive Approaches to Second Language* (pp. 223-237). Reading: Cambridge University Press.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL quarterly*, 25(3), 375-406.
- Hino, N. (2012a). Participating in the community of EIL users through real-time news: Integrated Practice in Teaching English as an International Language (IPTEIL). In A. Matsuda (Ed.),

- Principles and practices of teaching English as an international language* (pp. 183-200). Bristol: Multilingual Matters.
- Hwang, C. C. (2005). Effective EFL Education Through Popular Authentic Materials. *Asian EFL Journal*, 7(1), 1-12.
- Ishikawa, T. (2006). The effects of task complexity and language proficiency on task-based language performance. *The Journal of Asia TEFL*, 3, 193-225.
- Knapp, P., & Watkins, M. (2005). *Genre, text, grammar: Technologies for teaching and assessing writing*. UNSW Press.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.
- Krashen, S., & Brown, C. L. (2007). What is academic language proficiency. *STETS Language & Communication Review*, 6(1), 1-5.
- Kurosaki, S. (2014). *kigyo niokeru eigoshiyo nitsuiteno ichibunseki – ryutsu kaihatsu gyokai no kesu* [Analysis of English Use at Corporations – in Distribution and Development Industries]. *Jissen Women's University Faculty of Humanities and Social Science kiyo* [Journal of Jissen Women's University Faculty of Humanities and Social Science]. 10, 103-122.
- Matsuda, P. K. (1999). Composition studies and ESL writing: A disciplinary division of labor. *College Composition and Communication*, 699-721.
- Matsuda, P. K. (2010). English writing in Japan: Toward integration. *JACETJournal*, 50, 15-20.
- Mizumoto, A., & Takeuchi, O. (2008). Basics and Considerations for Reporting Effect Sizes in Research Papers. *Eigo kyoiku kenkyu*, 31, 57-66.
- Nation, P. (2009). Reading Faster. *International Journal of English Studies*, 9(2), 131-144.
- Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2<sup>nd</sup> ed.). Oxford: Heinemann
- Pinker, S. (1994). *The Language Instinct*. New York: William Morrow and Company.
- Schmidt, R.W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics*, 11 (2), 129-158.
- Shirai, Y. (2012). *Eigo kyoshi no tameno daini gengo shutokuron nyumon* [Second Language Acquisition: An Introduction for English Language Teachers]. Tokyo: Taishukan.
- Swan, M. (2006). Teaching Grammar – Does Grammar Teaching Work?. *Modern English Teacher*, 15(2), 5-13.
- Tsuji, K. (2014). Research on English Education Needed for the Optimization of English Communication at Corporations. Grant-in-Aid for Scientific Research (C) Number 23520779.
- Ur, P. (1999). *A Course in Language Teaching: Practice and Theory*. Cambridge: Cambridge University Press.
- Wilkins, D. A. (1976). *Notional Syllabuses*. Oxford: OUP.
- Worfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity*. Honolulu: Second Language Teaching &

Curriculum Center, University of Hawaii.

Yashima, T. (2011). *Tabunka wo ikiru chikara to eigo comyuniqueishon* [Living in a multicultural society and English communication]. *Kokusai rikai kyoiku kenkyu senta nenpo* [Annual of research center of international understanding education], 6, 1-14.

## Appendix 1 Grammatical Gap-fills (Example)

### ***External English tests urged for college entrance***

... Makoto Yamaguchi (pseudonym), a panel member and teacher at a cram school, says entrance exams for English \_\_\_\_ changed for Japanese people to keep up with global society.

(A) is (B) to be (C) being (D) must be

Entrance exams in English at Japanese universities \_\_\_\_ criticized for having too much emphasis on reading and writing..

(A) is frequently (B) to be frequently (C) have frequently been (D) which are frequently

Experts say two other skills -- listening and speaking --- \_\_\_\_ tested equally. ..

(A) is (B) to be (C) which are (D) should be

Retrieved Jul. 4, 2014, from <http://www3.nhk.or.jp/nhkworld/>

## Appendix 2 Pre- and Post-test (Examples)

1. Mr. Anderson feels that \_\_\_\_ is certainly the best way to resolve any problems among co-workers. (A) talks (B) has talked (C) talking (D) will talk
2. Five cents of every dolar \_\_\_\_ on goods in the ABC Store will go toward local charities that help children. (A) spending (B) spend (C) spends (D) spent
3. The software sold with the Mat Computer \_\_\_\_ the user to perform basic word processing and spreadsheet work. (A) enables (B) that enable (C) enabling (D) to enable

## Appendix 3: Rubrics for Criticality Assessment

1. The writing lacks understanding of or knowledge about the topic.  
No ideas or arguments are contained.
2. The writing is based on the topic but often strays from the main point. No clear ideas are presented and irrelevant arguments are often seen.
3. Neutral
4. General ideas are presented and the articles are summarized relatively well, but the writing does not display much systematic reasoning. Although written based on the topic, the writer's own views are not clearly expressed.
5. Ideas and opinions are expressed clearly and the writer's views are intelligible. The writing is interesting to read, appealing to readers, and well-organized with smooth (proper) story development

## **The Effects of Chunk Combination Practice on a Picture Description Task in a Second Language**

Koichi Yamaoka  
Graduate School, Kansai University

### **ABSTRACT**

Speaking with balanced fluency and accuracy is desirable for successful oral communication. However, it is not easy to achieve roughly parallel gains in both these capacities in a second language, since they are in a trade-off relationship due to limited cognitive resources. One possible way to improve both is to ease the related cognitive burden by utilizing chunks of speech, which are processed holistically, contribute to fluency, and may also improve accuracy if they are memorized correctly. However, even with correct chunks, errors still occur at the junctions between them. This study investigated the effects of chunk combination practice to reduce these errors without harming fluency. The participants were Japanese first-year university students, who remained at beginner's level. They engaged in a picture description task. The data of 15 participants who completed the task confirmed a trade-off relationship between some fluency and accuracy measures. Then the participants were divided into two groups for a reading-aloud activity: One group used a random-sentence list and the other used a patterned-sentence list presenting chunk combinations. The results indicated that the patterned-sentence list promoted more accurate verb and more fluency improvement than the random list, indicating an effect of chunks improving fluency and accuracy.

**Keywords:** fluency, accuracy, chunk, picture description task

### **I. INTRODUCTION**

Vallette (1993) emphasized the importance of balanced fluency and accuracy in second language learning by briefly comparing traditional accuracy-based language instruction and proficiency-based language instruction. However, fluency and accuracy are in a trade-off relationship attributed to the limited human capacity to bring cognitive resources to bear on a given task (Ahmadian & Tavakoli, 2011). When cognitive resources are used primarily for analytic processing to attain grammatical accuracy, fluency suffers; when cognitive resources are primarily allocated to processing of meaning to achieve fluency, accuracy suffers. This can lead to performance that exhibits imbalanced fluency and accuracy. However, if the cognitive burden is eased for fluency and accuracy, both of them may improve. One possible way of easing cognitive burden is utilization of chunks (Lewis, 1993).

This study explores the effects of chunk usage on fluency and accuracy. To this end, performance on picture-description monologue tasks was evaluated for fluency and accuracy in order to identify any effects of the reading aloud activity, which was expected to enhance implicit

knowledge of chunk combination.

## II. LITERATURE REVIEW

Chunks are grouped or organized input sequence units (Miller, 1956), including formulaic sequences or formulae that are “lexical chunks which result from memorizing the sequence of frequent collocations” (Ellis, 2003, p.68). Chunks are processed holistically with little cognitive burden (Skehan, 1998). Further, as long as chunks are memorized and deployed correctly, their junctions are the only possible places for grammatical errors to sneak in:

formulaic sequences (at least those that are ‘correctly’ committed to memory) may help speakers reach a degree of linguistic accuracy, because these pre-fabricated chunks constitute ‘zones of safety’ and appropriate use of them may thus confine the risk of ‘erring’ to the spaces in between the formulaic sequences in one’s discourse. (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006, p. 247)

In reality, however, the literature is yet to confirm any improvement of accuracy as a result of chunk instructions. For example, Boers et al. (2006) reported that the evidence for improvement of learners’ perceived accuracy was not so convincing even after eight months of instructions to direct the learners’ attention to formulaic sequences. One possible reason for this is that simply emphasizing chunks for memorization is not enough to reduce errors, which sneak in at the junctions between chunks as pointed out by Boers et al. (2006). Therefore, the key to higher accuracy is how to combine chunks correctly. At the same time, this needs to be achieved without harming fluency at chunk junctions. One possible answer to how to achieve this goal is utilization of implicit knowledge, which is “available for automatic use” (Ellis, 2004, p. 231). If implicit knowledge can be utilized to combine correct chunks without errors, accuracy can be expected to improve without harming fluency.

## III. METHOD

### 1. Purpose

The purpose of this study was to investigate the effects of implicit knowledge of chunk combinations on spoken language accuracy as well as fluency. This study hypothesizes that more saliently presented patterns of chunk combinations lead to higher accuracy in spoken performance in a second language along with at least the same level of fluency compared to less saliently presented chunk combinations. Here, saliency means the structural patterns are noticeable and expected to be picked up by learners more easily than by random presentation.

## 2. Definitions

In this study, *fluency*, *accuracy*, and *chunk* are defined as follows.

- (1) Fluency is rapid and smooth real-time language use.
- (2) Accuracy is real-time language use without errors related to target grammar items
- (3) A chunk is a group of word(s) that can be processed as a single sequence that corresponds to meaning.

Segalowitz (2010) categorized fluency into three domains: *cognitive fluency*, *utterance fluency*, and *perceived fluency*. Cognitive fluency is fluency of the cognitive processing that takes place in the brain; utterance fluency is fluency of the actual utterance, measurable with temporal measures; and perceived fluency is fluency as perceived by listeners. This study focuses on utterance fluency, because the interest here is the effects of chunks on actual language performance. The temporal measures are further categorized into three subcategories (Tavakoli & Skehan, 2005): *breakdown fluency*, *speed fluency*, and *repair fluency*. Breakdown fluency is related to pausing; for example, lower pause frequency and shorter pauses mean higher breakdown fluency. Speed fluency is based on the speed of the utterance. Measures of speed fluency include speech rate, articulation rate, phonation–time ratio, and mean length of run: Higher values for these measures indicate higher fluency. Finally, repair fluency is concerned with repair phenomena such as reformulation, replacement, false starts, and repetitions: Fewer repairs mean higher fluency. This study used breakdown fluency and speed fluency as measures because these two groups of measures are interrelated by equations (Yamaoka, 2013) and function well as a set of indices.

Accuracy is defined in relation to specific target grammar items in this study because our interest is the effects of chunks and chunk combinations on individual grammar items rather than on general accuracy. The target grammar items are articles, verbs, and prepositions that are commonly the sites of grammatical errors by beginner-level learners (Bitchener, Young, & Cameron, 2005; Cai, Hu, Miao, & Song, 2009; Ellis, 2005; Izumi, Uchimoto, & Isahara, 2005; Murata & Isahara, 2004; Tetreault & Chodorow, 2008).

Chunking is closely related to meaning (Ellis, 2001, 2003). Therefore, the author decided that chunks in this study are units related to meaning that can be treated as one whole, including combinations of article and noun, structures with a slot, and phrases. To show some examples, *a dog* is a chunk as a combination of article and noun, *[ ] is running* is a chunk with a slot, and the phrase *on the bench* is also a chunk. As a natural consequence of this definition, chunks in this study can contain a single word as long as that word represents an object. For example, *dogs* is treated as a chunk because it represents objects in the same way as *a dog* does. On the other hand, *dog* is not a chunk because it does not represent any objects; the object it represents is not specified for number and therefore is not a complete representation.

### 3. Measures

Table 1 shows the fluency and accuracy measures used in this study. *Speech time* was used to obtain *speech rate*, *articulation rate*, *phonation time ratio*, and *pause frequency*. Speech time is the total time used for an utterance, from the starting of the first meaningful word to the end of the last meaningful word.

**Table 1. Measures of fluency and accuracy**

Measure [unit]	Definition
<b>Fluency</b>	
Speech Rate [syllables/s]	Total number of syllables divided by speech time, where speech time is total time used for the speech
Articulation Rate [syllables/s]	Total number of syllables divided by total time of phonation, excluding pause
Phonation Time Ratio (PTR)	Total time of phonation divided by speech time
Mean Length of Run (MLR) [syllables]	Total number of syllables divided by number of runs
Pause Frequency [/s]	Total number of pause occurrences divided by speech time
Mean Pause Time [s]	Total pause time divided by total number of pause occurrences
<b>Accuracy</b>	
Article	The number of correct uses (or non-uses in case of plural forms) of articles divided by total number of article uses (or non-uses)
Verb	The number of correct uses of verbs divided by total number of verb uses
Preposition	The number of correct uses of prepositions divided by total number of preposition uses

Two kinds of pause were measured to obtain pause frequency and mean pause time; that is, *silent pause* and *filled pause*. A silent pause was defined as a silent period of longer than 0.25 seconds. Shorter threshold time may allow more pauses to be detected. However, the threshold of 0.25 second already began to cut a few words into smaller pieces in this study; for example, *cat* into *ca +t*. Therefore, this is approximately the shortest borderline threshold that does not break words into sound fragments. Furthermore, this threshold is roughly consistent with those in several previous studies; for example, 0.28 second (Towell, Hawkins, & Bazergui, 1996), 0.25 second (De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2012), and 0.2 second (Kormos & Denis, 2004; Lennon, 1990). A filled pause was defined as a period of non-words, for example, *eh*, *um*, and Japanese *eh-to*. A sequence of silent and filled pause was counted only a single *pause occurrence* because counting successive silent and filled pauses as reflecting the intention to pause two separate times seemed unreasonable. Pause frequency was calculated based on the number of pause

occurrences. Mean pause time was obtained by using *total pause time*, which included both silent and filled pauses. Accuracy measures were obtained by dividing the number of correct uses of the target grammar item by the total number of uses of that grammar item. The correct use of articles included non-use of the indefinite articles *a/an* with plural-form nouns, because it is grammatically incorrect to use the indefinite article *a/an* with a plural noun.

#### 4. Participants

The participants were 31 first-year college students (14 male and 17 female) majoring in health care sciences at a Japanese university. Their English proficiency level corresponded to the Grade 4 of the Society for Testing English Proficiency (STEP) test. Of these 31 participants, 15 students (5 male and 10 female) successfully completed the tasks. Of the others, some were stuck and kept silent for the rest of the allotted time, others did not finish the tasks on time, one participant failed to record the voice, and the rest resorted to consulting the English sentence lists at hand instead of doing it on their own. The failing students received no extra support during the tasks because it was impossible to help them out, since 15 or 16 students were engaged in the tasks simultaneously. From educational viewpoint, however, all of the participants were provided with a chance to practice the same speaking task again in pairs with peer feedbacks later in the course.

#### 5. Task

The experiment incorporated picture description tasks. The participants were instructed to look at a picture of a park silently for 10 seconds and then describe the actions of the people and animals in the park, together with their locations, within 1 minute and 30 seconds. The 10-second preparation period was intended to provide the participants with time enough to understand the situation described in the picture but not enough to pre-construct English sentences. Speech time was limited to 1 minute and 30 seconds in order to press the participants and minimize their on-line planning (Yuan & Ellis, 2003). It is expected that a trade-off relationship between fluency and accuracy will tend to manifest itself under such conditions, that is, of no pre-task planning and limited on-line planning (Yuan & Ellis).

#### 6. Materials

Four pictures were prepared for the picture description tasks, inspired by interview cards used in the STEP Grade Pre-2 speaking test (Obunsha, 2013). This type of picture description task was selected because it seemed suitable for eliciting utterance to measure fluency and accuracy. Unlike the STEP Grade Pre-2 speaking test, necessary words were provided to the participants beforehand with the intention of lowering the difficulty of the tasks. On top of that, verbs required for the tasks were limited to basic, common ones such as *walk*, *stand*, *run*, *sit*, and *read*. However, it must be admitted that the tasks were still too difficult for most of the participants, judging from the fact that only about half of them properly completed the tasks. One of the four pictures showed an empty park with no people or no animals, and was labeled with English words *tree*, *bench*, *pond*, *trashcan*,

and *park entrance* in order to furnish the participants with the English words necessary for the picture description tasks. The other three pictures showed people and animals doing different things in the park. Each of the three pictures contained two single agents and two pairs of agents. For example, in one picture, a dog is sitting under the tree, a man is reading a book on the bench, a woman is walking by the trashcan, two girls are standing at the park entrance, and two cats are running to the pond (See Appendix A for an example of one of the pictures). Expected article uses or non-uses were *a man, a woman, a boy, a girl, a dog, a cat, boys, girls, dogs, cats, the bench, the tree, the pond, the trashcan, and the park entrance*. Expected verb uses were *is standing, is walking, is running, is reading, are standing, are walking, and are running*. Expected preposition uses were *at the park entrance, by the trashcan, on the bench, under the tree, and to the pond*.

In order to attempt to confirm the hypothesis of this study, two treatment materials were prepared for the read-aloud activity as an intervention: the “Random List” and the “Patterned List”. The Random List was a list of English sentences presented in random order (See Appendix B for an excerpt of the Random List). Each sentence was segmented into three chunks with slashes (/) by the author, based on the definition of chunks adopted in this study. A Japanese translation was provided for each sentence. This list contained 80 sentences; they were presented in five sections with 24, 16, 16, 16, and 8 sentences respectively. The Patterned List was intended to present more salient chunk combination patterns than those on the Random List. The sentences were the same as those in the Random List, as were the number of sections and the number of sentences in each section. However, the order of the sentences and the way they were presented were different from the Random List (See Appendix C for an excerpt of the Patterned List). That is, English sentences were presented as combinations of three chunks, whose junctions were marked by slashes (/), and Japanese translation was provided for each chunk. For example, in the case of the list in Appendix C, the participants read the sentences aloud in the following order:

A boy / is standing / at the park entrance.  
 A boy / is standing / by the trashcan.  
 A boy / is walking / at the park entrance.  
 A boy / is walking / by the trashcan.  
 A girl / is standing / at the park entrance.  
 A girl / is standing / by the trashcan.  
 A girl / is walking / at the park entrance.  
 A girl / is walking / by the trashcan.

The junctions between the subject and the verb present chunk combination patterns most saliently with a repeated pattern of “(a + noun) + (is + base form of verb + ing)”. Therefore, if any effect of combination patterns exists, it should be observed in the performance at these junctions; as differences of verb accuracy between the two treatments. The articles show simple patterns, with every sentence starting with the indefinite article *a* or a plural noun. The patterns are not related to

combination of chunks and are similar between the two lists. Therefore, effects to be observed in article use are not likely to reflect the effects of chunk combinations. The prepositions are not related to chunk combinations either: The following noun determines the kind of preposition to be used, not the preceding verb. Therefore, use of prepositions is expected to reflect how correctly the participants memorize the chunks initiated by the prepositions. The sentences on the both lists were selected to cover all the necessary sentences to complete the tasks. The materials were compiled in a booklet in the intended order of procedure, to allow the participants to follow the procedure by turning the pages as instructed. Sixteen IC recorders (Sony ICD-BX332) were used to record participants' individual performance while they spoke simultaneously.

## 7. Procedure

The experiment was conducted as speech practice during a regular lesson. The aim of the course was to expand students' medical vocabulary and expressions to improve their English skills, mainly in reading and listening but not excluding writing and speaking. The experiment served to familiarize students with speaking activities, and was followed by role-playing practice later in the course. The participants were divided into two groups because of the limited number of IC recorders available. One group participated in the tasks and treatment, while the other group took a grammar test in the same classroom. Table 2 shows the procedure for the tasks and treatment.

**Table 2. Procedure for the tasks and treatment**

Sessions	Content
Word Check	Look at a labeled picture and read the words aloud.
Task Practice	Describe a picture.
Task 1	Describe a picture.
Treatment	Read aloud the Random List or the Patterned List.
Task 2	Describe a picture.

The first step for the task-and-treatment group was Word Check, intended to provide them with the words necessary to complete the main tasks. In this session, the group was asked to look at a picture of an empty park labeled with the English words *tree*, *bench*, *pond*, *trashcan*, and *park entrance*. The group was then asked to read the words aloud after the instructor.

In the following Task Practice session, the group members looked at a picture for 10 seconds and then described it in English with a time limit of 1 minute and 30 seconds. Speech time was limited in order to minimize on-line planning and suppress rule-based processing to a minimal level (Ahmadian & Tavakoli, 2011). The group was instructed to describe the actions of the people and animals in the picture together with their locations, under a scenario in which they were explaining the situation in the park to a friend over the telephone. In this first Task Practice session, everyone viewed the same picture. The first main task (Task 1) followed the Task Practice session: The group members engaged in Task 1 simultaneously and recorded their performance individually.

In the Treatment session after Task 1, each participant was assigned to either the Random List treatment or the Patterned List treatment. Participants under the Random List treatment were required to read the sentences on the list aloud as fast as possible, with a pause at each slash. In the case of the Patterned List treatment, participants were told that combining the chunks with the slashes between would allow them to construct full sentences. Then they were instructed to read the sentences as fast as possible, with a pause at each chunk junction marked with a slash. The pauses were intended to familiarize the participants with the chunks and impress chunk content and location of the junctures upon them. The time pressure was expected to encourage holistic rather than analytic processing, thus enhancing the participants' implicit knowledge. In both treatments, no explicit grammatical rules were explained, and the participants read the list aloud once, taking about four minutes. Performance in the Treatment session was also recorded in order to ensure that the participants were really engaged in the treatment.

Next, the group engaged in the second main task (Task 2) following exactly the same procedure as Task 1 but with a different picture (with the order randomized for each participant). After Task 2 was completed, the groups switched roles: The group that engaged in the tasks and treatment took the grammar test, and the other group that had taken the grammar test went through the whole procedure explained above. The entire experimental procedure took about 30 minutes for each group.

## 8. Analysis

The recorded speeches were transcribed for counting syllables and accuracy analysis. If the task was found to have been completed improperly, that participant's set was abandoned and excluded from the analyses. Improper datasets included cases of unfinished speech, recording failure, and cases where the sound of turning pages was recorded—an indication that the participant was consulting the treatment sentence lists. *Praat* (Boersma & Weenink, 2009), free software for speech analysis, was used to calculate speech time and detect pauses. Fluency and accuracy measures were calculated using Microsoft Excel 2008 for Mac (version 12.3.6), and further statistical analyses were conducted using SPSS Version 19. Effect sizes were obtained by dividing *z*-value outputs of SPSS by the total sample size (Mizumoto & Takeuchi, 2008).

## IV. RESULTS

In this section, first, the results from Task 1 are presented to demonstrate the trade-off relationship between fluency and accuracy. Next, task performance before and after the treatment is reported, and the effects of the two treatments are compared.

### 1. Trade-off relationship between fluency and accuracy

In all, 15 out of 31 participants completed the tasks properly. Table 3 shows correlation coefficients between fluency and accuracy measures for Task 1: There was a positive correlation ( $r = 0.526$ ) between mean pause time and verb accuracy while the other measures did not reach

significance level. However, PTR and pause frequency each showed a moderate correlation with verb accuracy ( $r = -0.475$  for PTR;  $r = -0.452$  for pause frequency).

**Table 3. Correlations between fluency and accuracy measures in Task 1 (Pearson's  $r$ )**

Fluency measures[unit]	Accuracy		
	Article	Verb	Preposition
Speech Rate [syllables/s]	.299	-.294	.159
Articulation Rate [syllables/s]	.330	.135	.097
PTR	.172	-.475	.210
MLR [syllables]	.185	.229	.087
Pause Frequency [/s]	.020	-.452	.092
Mean Pause Time [s]	-.160	.526*	-.186

Note.  $n = 15$ . \*  $p < .05$ .

## 2. Comparison of task performance before and after the treatments

Nine participants receiving the Random List treatment completed the tasks properly. Table 4 shows their task performance before and after the Random List treatment.

**Table 4. Task performances before and after Random List treatment**

Measure [unit]	Task 1		Task 2		Effect Size
	Mean	SD	Mean	SD	$r$
Speech Rate [syllables/s]	.89	.19	1.29	.24	.89
Articulation Rate [syllables/s]	2.41	.27	2.74	.33	.89
PTR	.36	.05	.47	.07	.84
MLR [syllables]	2.12	.32	2.81	1.17	.89
Pause Frequency [/s]	.39	.06	.46	.10	.53
Mean Pause Time [s]	1.65	.35	1.20	.31	.77
Accuracy (Article)	.51	.23	.89	.15	.89
Accuracy (Verb)	.55	.36	.56	.40	.14
Accuracy (Preposition)	.67	.22	.78	.15	.42

Note.  $n = 9$ . Wilcoxon signed-rank test (matched pairs). Effect size,  $r$ , was obtained by dividing  $z$ -value by the square root of the sample size.

Effect size values around 0.5, 0.3, and 0.1 are considered large, medium, and small respectively (Mizumoto & Takeuchi, 2008). Judging from these criteria, all the fluency measures except for pause frequency indicated improvement of fluency after the Random List treatment, with large effect sizes. In addition, pause frequency, a dysfluency measure, increased after the Random List treatment, with a large effect size ( $r = .53$ ). The reason for this will be discussed later, in the discussion section. Accuracy of articles improved, with a large effect size ( $r = .89$ ); accuracy of

verbs improved slightly, that is, with a small effect size ( $r = .14$ ); and accuracy of prepositions improved with a relatively large effect size ( $r = .42$ ).

Six participants receiving the Patterned List treatment completed the tasks properly. Table 5 shows their task performance before and after the Patterned List treatment. The Patterned List treatment showed similar tendencies to the Random List treatment with regard to the fluency measures, all of which except for pause frequency, improved after the Patterned List treatment with large effect size. Pause frequency also increased after the treatment, with a large effect size ( $r = .50$ ). Accuracy of articles improved with a large effect size ( $r = .90$ ), as with the Random List treatment; however, the other two accuracy measures showed different degrees of changes compared to the case of the Random List treatment: Accuracy of verbs improved with a medium effect size ( $r = .28$ ), and accuracy of prepositions improved with a small effect size ( $r = .15$ ).

**Table 5. Task performance before and after the Patterned List treatment**

Measures [unit]	Task 1		Task 2		Effect Size
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>r</i>
Speech Rate [syllables/s]	.72	.14	1.23	.18	.90
Articulation Rate [syllables/s]	2.37	.34	2.70	.41	.90
PTR	.31	.06	.47	.09	.90
MLR [syllables]	2.52	.61	3.53	.89	.81
Pause Frequency [/s]	.28	.06	.34	.07	.50
Mean Pause Time [s]	2.63	.65	1.68	.60	.73
Accuracy (Article)	.55	.15	.92	.08	.90
Accuracy (Verb)	.74	.20	.83	.14	.28
Accuracy (Preposition)	.71	.37	.75	.24	.15

*Note.*  $n = 6$ . Wilcoxon signed-rank test (matched pairs). Effect size,  $r$ , was obtained by dividing  $z$ -value by the square root of the sample size.

### 3. Comparison of treatments

Table 6 compares the outcomes (that is, the changes from Task 1 to Task 2) after the Random List and Patterned List treatments. Compared to the Random List treatment, the Patterned List treatment produced fluency measures that indicated greater improvement of fluency except for articulation rate and pause frequency — speech rate and PTR with relatively large effect sizes, MLR and mean pause time with medium effect sizes. The treatments exhibited no differences in articulation rate or in pause frequency ( $r = 0$  for both). Noticeably, the Random List treatment resulted in more improvement in accuracy of verbs with a medium effect size ( $r = .31$ ) than the Random List treatment; while the two treatments showed little difference in accuracy of articles ( $r = .02$ ), accuracy of prepositions was slightly lower in the Patterned List treatment with a small effect size ( $r = .18$ ).

**Table 6. Comparison of outcomes after the treatments**

Measures [unit]	Random List		Patterned List		Effect Size
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>r</i>
Speech Rate [syllables/s]	.41	.17	.51	.20	.37
Articulation Rate [syllables/s]	.33	.21	.33	.18	.00
PTR	.11	.05	.16	.08	.43
MLR [syllables]	.68	.95	1.01	.81	.24
Pause Frequency [/s]	.07	.12	.06	.11	.00
Mean Pause Time [s]	-.46	.44	-.95	1.04	.24
Accuracy (Article)	.38	.23	.36	.20	.02
Accuracy (Verb)	.01	.26	.09	.21	.31
Accuracy (Preposition)	.11	.23	.04	.19	.18

*Note.*  $n = 9$  for Random List.  $n = 6$  for Patterned List. Mann–Whitney U test (two groups). Effect size,  $r$ , was obtained by dividing z-value by the square root of the total sample size.

## V. DISCUSSION

### 1. Trade-off relationship between fluency and accuracy

The trade-off relationship between fluency and accuracy manifests itself in the form of negative correlations between fluency measures (speech rate, articulation rate, PTR, and MLR) and accuracy measures as well as positive correlations between dysfluency measures (pause frequency and mean pause time) and accuracy measures. Therefore, the positive correlation between mean pause time and accuracy of verbs in Table 3 is clear evidence of a trade-off relationship between fluency and accuracy as is the negative correlation between PTR and accuracy of verbs. The negative correlation between pause frequency and accuracy of verbs seemingly contradicts the trade-off relationship. However, this contradiction is quite possibly due to the characteristics of these measures. That is, pause frequency is affected by mean pause time: Shorter mean pause time leads to shorter intervals between pauses, thus resulting in denser pause occurrences (higher pause frequency). In general, mean pause time and pause frequency can represent different factors. In this case, however, mean pause time showed a very strong negative correlation ( $r = -0.950$ ) with pause frequency, suggesting strong domination of mean pause time over pause frequency—to the extent that other factors can hardly be detected by pause frequency. On top of that, pause frequency approximately equals inversion of the sum of mean pause time and *mean duration between pauses*. Stepwise multiple regression analysis by SPSS discarded mean duration between pauses from the variables, and mean pause time remained as the only independent variable to explain pause frequency in this study. To examine the effects of threshold for pauses, correlation coefficients between mean pause time and verb accuracy were calculated for different thresholds: 0.5s, 0.75s, and 1s. The coefficients obtained were 0.471, 0.503, and 0.423 respectively. As can be seen, the

coefficient starts to deteriorate slightly at the threshold of 1s. This might be interpreted as an indication that shorter pauses, below 1s, do not contribute to the trade-off relationship. However, this remains speculative and requires further study.

Thus a trade-off relationship was confirmed only between fluency and accuracy of verbs. This is understandable, considering that verb use is expected to require more cognitive resources for analytic processing than use of articles and prepositions. The forms of a *be* verb must be modified according to the number of the subject and *-ing* must be added to the base form of the verb. As for the articles, only the number of the following noun is concerned. The preposition can be chosen directly based on the following noun.

## 2. Comparison of task performance before and after the treatments

Speech rate, articulation rate, PTR, MLR, mean pause time, and accuracy of articles improved using both the Random List (Table 4) and the Patterned List (Table 5). These results can be explained as follows. Speech rate is a product of articulation rate and PTR. According to the results, PTR underwent larger proportional change than articulation rate; in the Random List treatment, articulation rate increased by 14 percent and PTR by 31 percent, and in the Patterned List treatment, articulation rate increased by 14 percent and PTR by 52 percent. This means that PTR contributed to the increase of speech rate more than articulation rate did. Furthermore, PTR is affected by both mean pause time and pause frequency. Thus, pause is the underlying main factor that caused the changes in speech rate and PTR. MLR is also affected by mean pause time, pause frequency, and articulation rate, and its change can thus be explained by the same logic.

The fact that accuracy of articles improved after both treatments (Tables 4 and 5) can be explained by the saliency of article use in the treatments. The sentences in the two treatments started with either the indefinite article *a* or a plural noun. The other nouns were accompanied with the definite article *the* except in one case, *a book*. This is a very simple and salient pattern. Therefore, accuracy of articles improved in the treatments. In addition, whether the randomized or the patterned list was used, the pattern of article use was the same. The improvement of accuracy of articles can be thus understood as the effect of this salient patterns, not of chunk combination patterns.

It should be noted here that the effects of the treatments in this study are inseparable from those of task repetition. Task repetition has been shown to enhance fluency, although the results are mixed for accuracy (Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Gashan & Almohaisen, 2014; Lynch & Maclean, 2000, 2001). However, the improvement of accuracy of articles in this study can still be assumed to be largely due to the treatments, because it is unlikely that the participants who had been starting sentences with *the* would suddenly begin to start sentences with *a* just because of task repetition.

To determine the effects of chunk combinations in this study, the results of the two treatments must be compared. If there are differences between the effects of the two treatments, these differences should be attributed to the differences between the treatments.

### 3. Comparison of outcomes after treatments

In order to determine the differences between the effects of the two treatments, outcomes after treatment were compared (Table 6). Among fluency measures, speech rate, PTR, MLR, and mean pause time indicated higher fluency outcomes after the Patterned List treatment than after the Random List treatment, with medium or relatively large effect sizes. Considering that all these measures are notably affected by pause, as already mentioned in the comparison of task performance before and after treatments, these results indicate that the two treatments had different effects on pause: It is possible that the Patterned List made chunks more salient than the Random List, thus facilitating processing of implicit knowledge and reducing pause time needed.

As for accuracy, while the outcome for accuracy of articles was almost exactly the same for the two treatments ( $M = 0.38$  for the Random List;  $M = 0.36$  for the Patterned List;  $r = .02$ ), there was some indication of differences in the outcome for accuracy of verbs ( $M = 0.01$  for the Random List;  $M = 0.09$  for the Patterned List;  $r = .31$ ). Articles were not related to chunk combinations; in both treatments, the sentences began with either the article *a* or a plural form noun, for example, *A dog* or *Dogs*. This is a very simple pattern that does not require combining of chunks. For the verbs, on the other hand, the Patterned List was designed to make the combination patterns between subject and verb more salient than those on the Random List. Thus, if there is some difference in the accuracy of verbs and not in the accuracy of articles, it should likely be attributed to the patterned presentation of combinations.

Further, the higher outcome for accuracy of verbs after the Patterned List treatment was accomplished without harming fluency—in fact, with an even with higher fluency outcome than for the Random List treatment, as already mentioned. This can be interpreted as an effect of chunks to improve both fluency and accuracy. That is, in explicit, rule-based processing, verbs are expected to require more cognitive resources to process than articles do. For example, the correct form of a *be* verb (such as *is* or *are*) is chosen depending on the number of the subject and requires the base form of the verb to be modified into the *-ing* form. On the other hand, combining chunks by means of implicit knowledge will require few cognitive resources. Therefore, the difference in accuracy of verbs between the treatments may be explained as follows: The chunks were combined using implicit knowledge after the Patterned List treatment, with fewer cognitive resources required, resulting in higher outcome for accuracy of verbs after the Patterned List treatment than after the Random List treatment, without harming fluency.

The outcome for accuracy of prepositions was also higher after the Random List treatment than after the Patterned List treatment ( $r = .18$ ). One possible explanation for this is the relative saliency of chunks across the treatments. It is possible that chunks with prepositions were less salient in the Patterned List relative to the high saliency of the chunk combination patterns with verbs. Therefore, it is likely that prepositions in the Patterned List did not appeal to the participants' memory so much as those in the Random List did.

Finally, it should be noted that the possibility of processing by using explicit instead of implicit knowledge cannot be totally excluded. In the future study, a delayed posttest is expected to

help determine the kind of knowledge the participants used, namely explicit knowledge or implicit knowledge, because effects of implicit knowledge last longer than those of explicit knowledge (Shin & Christianson, 2011). The future study also needs to investigate the possibility that chunk combination treatments affect each grammatical item differently as was suggested by the results of this study.

## VI. CONCLUSION

This study investigated the existence and nature of an effect of chunk combination practice on fluency and accuracy in second language English learners. A data set of picture description task results for 15 false-beginner university students was analyzed. The results before the treatments revealed a clear trade-off relationship between fluency (as represented by mean pause time) and accuracy (of verbs). Next, two reading-aloud treatments were conducted. One used the Random List, where sentences were listed randomly, and the other used the Patterned List, where sentences were listed according to patterns of chunk combinations. Fluency measures other than pause frequency indicated improved fluency after both treatments; this may be attributed to the effects of the treatments on pause. Interestingly, the Patterned List led to the greater fluency outcome, suggesting the possibility that this list promoted more utilization of implicit knowledge than the Random List. Improvement of accuracy of articles was at the same level in both treatments, probably because of the similar saliency of the patterns of article use in the two treatments. However, comparison of the two treatments showed that the Patterned List treatment resulted in greater improvement in the accuracy of verbs. Combined with the better fluency outcome after the Patterned List, this can be interpreted as showing an effect of chunk use to improve fluency and accuracy. As for preposition accuracy, it showed more improvement after the Random List treatment, possibly because prepositions were less salient relative to the verb patterns in the Patterned List and thus were not memorized very well.

Despite these suggestive results, the possibility of processing by using explicit knowledge instead of enhancing implicit knowledge cannot be totally excluded. Future studies need to examine the kind of knowledge the participants used, for example, by conducting a delayed posttest and adopting task-only group as a comparison group. This should be done before automatization of explicit knowledge, if any, takes place. In addition, there is an indication that chunk combination treatments affect each grammatical item differently. Furthermore, experiments using random and patterned oral repetitions are also required to examine whether the effects are due to reading aloud activity or to the visual pattern of the list.

## ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Osamu Takeuchi for his support and guidance. I would like to extend my appreciation to Ms. Ryoko Yanase, who kindly drew the four pictures used in the experiment. I would also like to thank the anonymous reviewers for their insightful comments.

## REFERENCES

- Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, 15, 35–59. doi: 10.1177/1362168810383329
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14, 191–205. doi: 10.1016/j.jslw.2005.08.001
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10, 245–261. doi: 10.1191/1362168806lr195oa
- Boersma, P., & Weenink D. (2009). Praat: doing phonetics by computer (Version 5.1.07) [Computer software]. Retrieved from <http://www.praat.org/>
- Bygate, M. (1996). Effect of task repetition: Appraising the development of second language learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–46). Oxford: Heinemann.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp. 23–48). Harlow, UK: Longman.
- Cai, D., Hu, Y., Miao, X., & Song, Y. (2009). Dependency grammar based English subject-verb agreement evaluation. In O. Y. Kwong, (Ed.), *Proceedings of the 23rd Pacific Asia conference on language, information and computation*. (pp. 63–71). Hong Kong: City University of Hong Kong Press. Retrieved from <http://www.aclweb.org/anthology-new/Y/Y09/Y09-1008.pdf>
- De Jong, N.H., Steinel, M.P., Florijn, A., Schoonen, R., & Hulstijn, J.H. (2012). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 24, 1–24. doi:10.1017/S0142716412000069
- Ellis, N. C. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33–68). Cambridge: Cambridge University Press.
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63–104). Oxford: Blackwell Publishing.
- Ellis, R. (2004). Definition and measurement of L2 explicit knowledge. *Language Learning*, 54, 227–275. doi:10.1111/j.1467-9922.2004.00255.x
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172. doi: 10.1017/S0272263105050096
- Gashan, A., & Almohaisen, F. (2014). The effect of task repetition on fluency and accuracy of EFL Saudi female learners' oral performance. *Advances in Language and Literary Studies*, 5, 36–41. doi: 10.7575/aiac.alls.v.5n.3p.36

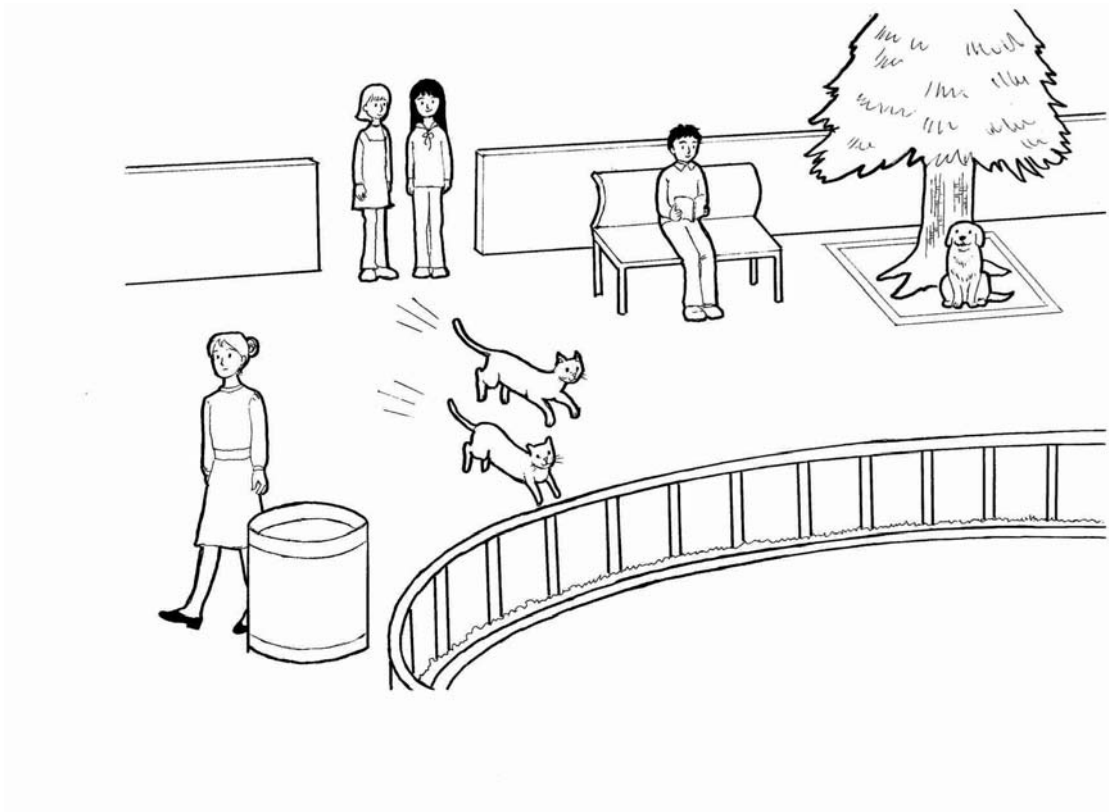
- Izumi, E., Uchimoto, K., & Isahara, H. (2005). *Nihonjin 1200nin no eigo supikingu kopasu* [English speaking corpus of 1200 Japanese]. Tokyo, Japan: ALC.
- Kormos, J., & Denis, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164. doi:10.1016/j.system.2004.01.001
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387–417. doi: 10.1111/j.1467–1770.1990.tb00669.x
- Lewis, M. (1993). The lexical approach: *The state of ELT and a way forward*. Hampshire: UK, Heinle Cengage.
- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221–50.
- Lynch, T., & Maclean, J. (2001). ‘A case of exercising’: Effects of immediate task repetition on learners’ performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks* (pp.141–162). Harlow, UK: Longman.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information, *The Psychological Review*, 63, 81–97. Retrieved from <http://www.musanim.com/miller1956/>
- Mizumoto, A., & Takeuchi, O. (2008). Kenkyuronbun ni okeru kokaryo no hokoku no tameni: Kisoteki gainen to chui ten [Basics and considerations for reporting effect sizes in research papers]. *Studies in English Language Teaching*, 31, 57–66.
- Murata, M., & Isahara, H. (2004). Three English assistance systems using automatic paraphrasing techniques. In *Proceedings of the 18th Pacific Asia conference on language, information and computation* (pp. 299–306). Tokyo, Japan: Logico-Linguistic Society of Japan. Retrieved from <http://www.aclweb.org/anthology/Y04-1032>
- Obunsha (2013). *Eiken jun 2-kyu kako 6-kai zen monndai shuu CD* [Exercise CD for Grade Pre-2 STEP with all the tests in the past six years]. Tokyo, Japan: Author.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York, NY: Routledge.
- Shin J. A., & Christianson, K. (2011). Structural priming and second language learning. *Language Learning*, 62, 931–964. doi: 10.1111/j.1467-9922.2011.00657.x
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Amsterdam: John Benjamins.
- Tetreault, J. R., & Chodorow, M. (2008). Native judgments of non-native usage: Experiments in preposition error detection. In *Proceedings of the workshop on human judgments in computational linguistics*. (pp. 24–32). Manchester: Association for Computational Linguistics. Retrieved from <http://www.aclweb.org/anthology-new/W/W08/W08-12.pdf>
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119. doi: 10.1093/applin/17.1.84
- Vallette, R. M. (1993). The challenge of the future: Teaching students to speak fluently and

accurately. *The Canadian Modern Language Review*, 50, 173–178.

- Yamaoka, K. (2013). Exploring the trade-off relationship between fluency and accuracy: Identifying the key factor that determines this relationship. *LET Kansai Chapter Collected Papers*, 14, 41–61.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1–27. doi: 10.1093/applin/24.1.1

Appendix A

An example of the task pictures



Appendix B  
An excerpt from the Random List

/ でポーズを入れながら、できるだけ速く音読してください。

セクション 1

A boy / is standing / at the park entrance.	1 人の少年が / 公園の入り口で / 立っている。
A girl / is walking / at the park entrance.	1 人の少女が / 公園の入り口で / 歩いている。
A cat / is sitting / under the tree.	1 匹の猫が / 木の下で / 座っている。
A cat / is running / to the pond.	1 匹の猫が / 池に向かって / 走っている。
A dog / is running / at the park entrance.	1 匹の犬が / 公園の入り口で / 走っている。
A man / is standing / at the park entrance.	1 人の男性が / 公園の入り口で / 立っている。
Cats / are running / to the pond.	猫たちが / 池に向かって / 走っている。
A girl / is standing / by the trash can.	1 人の少女が / ゴミかごの側で / 立っている。
A man / is standing / by the trash can.	1 人の男性が / ゴミかごの側で / 立っている。
A boy / is running / at the park entrance.	1 人の少年が / 公園の入り口で / 走っている。
Dogs / are running / by the trash can.	犬たちが / ゴミかごの側で / 走っている。
A woman / is walking / at the park entrance.	1 人の女性が / 公園の入り口で / 歩いている。
A boy / is walking / by the trash can.	1 人の少年が / ゴミかごの側で / 歩いている。
Girls / are running / at the park entrance.	少女たちが / 公園の入り口で / 走っている。
Boys / are walking / at the park entrance.	少年たちが / 公園の入り口で / 歩いている。
Girls / are standing / by the trash can.	少女たちが / ゴミかごの側で / 立っている。
Girls / are standing / at the park entrance.	少女たちが / 公園の入り口で / 立っている。
Boys / are running / to the pond.	少年たちが / 池に向かって / 走っている。
A woman / is standing / at the park entrance.	1 人の女性が / 公園の入り口で / 立っている。
Cats / are running / at the park entrance.	猫たちが / 公園の入り口で / 走っている。
A cat / is walking / to the pond.	1 匹の猫が / 池に向かって / 歩いている。
A girl / is walking / by the trash can.	1 人の少女が / ゴミかごの側で / 歩いている。
A dog / is running / to the pond.	1 匹の犬が / 池に向かって / 走っている。
Dogs / are walking / to the pond.	犬たちが / 池に向かって / 歩いている。

## Appendix C

### An excerpt from the Patterned List

次の3種類のパーツをそれぞれ組み合わせると、英文が出来上がります。  
/ でポーズを入れながら、全ての組み合わせをできるだけ速く音読してください。

### セクション 1

A boy	/	is standing	/	at the park entrance.
A girl	/	is walking	/	by the trash can.

1 人の少年 / ( ) が立っている / 公園の入り口で

1 人の少女 / ( ) が歩いている / ゴミかごの側で

A boy / is standing / at the park entrance.  
/ by the trash can.

A boy / is walking / at the park entrance.  
/ by the trash can.

A girl / is standing / at the park entrance.  
/ by the trash can.

A girl / is walking / at the park entrance.  
/ by the trash can.

A man	/	is standing	/	at the park entrance.
A woman	/	is walking	/	by the trash can.

1 人の男性 / ( ) が立っている / 公園の入り口で

1 人の女性 / ( ) が歩いている / ゴミかごの側で

A man / is standing / at the park entrance.  
by the trash can.

A man / is walking / at the park entrance.  
/ by the trash can.

A woman / is standing / at the park entrance.  
/ by the trash can.

A woman / is walking / at the park entrance.  
/ by the trash can.

次のページに続く

## **An Analysis of English Argumentative Essays Written by Japanese University Students Based on Four Analytical Frameworks to Identify Coherence Breaks<sup>1</sup>**

Miho Yamashita

Graduate School, Kansai University

### **ABSTRACT**

Numerous studies have investigated organizational patterns of English texts written by Japanese university students and their L1 interference in second language writing. There have been few studies, however, to clarify the nature of rhetorical anomalies, if not errors, observed in such texts and the reasons behind them. In order to identify these anomalies, this study investigated English argumentative essays written by two undergraduate students at different English proficiency levels. The analytical frameworks used for this study include 1) Structural-Functional Analysis, 2) Keywords-Chain Analysis, 3) Metadiscourse Markers (MDM) Mapping, and 4) Rhetorical Structure Theory (RST). The purpose of this study is two-fold: to test the viability of the above frameworks in the analysis of non-native English texts, and to identify rhetorical anomalies among students' English essays. Results demonstrated that the proposed analytical frameworks were useful although they need further refinements, and that there were three types of coherent breaks; namely, 1) irrelevant ideas, 2) insufficient information, and 3) incomprehensible content due to inappropriate word choice. By comparing the English essays with their Japanese counterparts, it was hypothesized that the differences in the basic logical frames between English and Japanese were partly responsible for the coherence breaks. Pedagogical implications of "writing bilingually" in second language writing will also be discussed.

**Keywords:** coherence breaks, keywords-chain, metadiscourse markers, RST

## **I. INTRODUCTION**

### **1. Previous Research**

There have been a number of studies to identify the features of English texts written by Japanese students in terms of lexis, syntax and rhetorical organization. As for organization, contrastive rhetoric study, which seeks to identify organizational characteristics of either or both L1 and L2 texts, has been widespread since Kaplan (1966). As for Japanese texts, past studies support the idea that they are organized inductively, unlike English texts, which are organized deductively. Hinds (1983) points out the *ki-sho-ten-ketsu* pattern that is unique to Japanese compositions as well as such characteristics as *quasi-inductive* or delayed introduction of purpose and *reader responsibility*. Based on the assumption that each language and culture has unique rhetorical conventions and that they negatively affect L2 writing (Grabe & Kaplan, 1989; Kaplan, 1966, cited in Kubota, 1998), there have also been studies to probe into L1 (Japanese) transfer to L2 (English)

in organization patterns (Hirose, 2005; Kubota, 1998). Since the characters of groups analyzed or the genres of composition among the target groups differ in each study, no conclusive results have been obtained in this regard; however, they have confirmed that organizational patterns are not always negatively transferred from L1 to L2. Kubota (1998) revealed that half of the students wrote L1 and L2 essays with different organizational patterns, and even among their L2 essays, those written with an inductive pattern were found to be evaluated positively.

These studies have sought to analyze the organizational patterns of L1 (Japanese) and L2 (English) characteristics of Japanese writers; however, studies which have analyzed rhetorical features or logical anomalies, if not errors, specific to English texts written by Japanese, have been scarce. Oi (2005) compared argumentative essays written by Japanese students and American students based on the Toulmin Model, and revealed some features of argumentative styles among the Japanese students. They include indecisive, inconsistent, and undeveloped arguments, and she argues that these present problems in writing a strong and effective argumentative essay in academic English. Hinkel (1997) compared essays written by Asian-background ESL writers and those written by native speakers of English, based on corpus analysis, in order to examine the use of indirect strategies with metadiscourse markers. The results showed that Asian writers, including Japanese, use indirect strategies such as vagueness, repetition, hedges, etc., more frequently than native speakers. The author (Hinkel, 1997) argues that these features of indirectness violate the expectations that English-speaking academia has about writing. Thus, ESL writers should be taught to write explicitly, with an appropriate style or tone of writing that meets such expectations.

As for the studies on investigating rhetorical problems, English texts written by non-native speakers of English other than Japanese have been conducted. Wikborg (1990) identified coherence breaks from an analysis of 114 essays and papers written by Swedish students. "Coherence breaks" are the disrupted sequences of ideas within paragraphs, and those found by Wikborg (1990) were of two types: topic-structuring problems and cohesion problems. Maghfiroh (2013) found that disunity of thought and disorganization of ideas were the most frequent causes of coherence breaks among English texts written by Indonesian ESL students. Skoufaki (2009) detected coherence errors in 45 paragraphs written by Chinese ESL students, using Rhetorical Structure Theory (Mann & Thompson, 1988).

Although much research has been conducted in the past to identify rhetorical features of Japanese L2 essays, there has been little research to pinpoint what kinds of rhetorical anomalies or coherence breaks occur in English texts written by Japanese, and the reasons behind them. It can be hypothesized that the transfer of organization or any type of coherence breaks detected in English texts may be accounted for by the rhetorical convention characteristics of Japanese writing; therefore, a comparison of their L1 and L2 texts is necessary.

The purpose of the current study is two-fold: to test the applicability of the four analytical frameworks which are briefly described in Section II and, using those frameworks, to identify logical anomalies or coherence breaks in English essays written by Japanese university students. This is a pilot study in which essays written by only two students are examined. However, the

author believes that the usefulness of the proposed analytical models to investigate rhetorical features of non-native English texts can still be demonstrated with a limited number of students, and this is what she hopes to prove in the current study.

In the next section, a brief outline of the writing class from which the data for this study were collected will be presented. This is followed by the description of the four analytical models in Section II. Then, the research findings and discussion will be presented in Sections III and IV respectively. The paper will conclude with the author's remarks on some of the limitations of the current study as well as its pedagogical implications.

## **2. Writing Classes <sup>2</sup>**

The data for this study were collected from English writing classes at Kansai University in Osaka. These classes were developed for the Kansai University Bilingual Essay Corpus (KUBEC) Project in 2012 and were scheduled to continue until 2014. The purpose of the project is to compile a large scale corpus of essays written by undergraduate students in both English and Japanese. The participants consist of two groups: students from the Faculty of Foreign Language Studies (G-class hereafter) and students from the Faculty of Law (L-class hereafter). G-class students take part in a "Study Abroad (SA)" program when they are sophomore students. They study at one of the designated universities in English-speaking countries for about nine months where they receive extensive training in English. The L-class, on the other hand, has no prior experience of academic English writing, nor of studying abroad. G-class students are juniors, and L-class students are sophomores. The English proficiency of the G-class students is considered at intermediate level, as their TOEIC scores range from 660 to 840. The L-class, on the other hand, belongs to the beginner level with scores ranging from 460 to 540. A unique feature of these classes is that students are required to write essays both in English and in Japanese on the same 13 topics.<sup>3</sup> In each class, they first write an English essay in 60 minutes and a Japanese counterpart in 30 minutes. The students whose data were collected for this study are not able to use any reference materials, including dictionaries, while writing in class.

## **II. THE STUDY**

This study looks into rhetorical features of English essays written by Japanese students, and identifies coherence breaks, if any, in their English texts with a partial attention to their Japanese counterparts. Before the analytical frameworks are discussed, the definitions of coherence and cohesion, which are the key notions in the current analysis, will first be described in the next section.

### **1. Coherence and cohesion of a text**

"A discourse is not an arbitrary collection of unrelated units; rather, it is an accumulation of various parts that work together conceptually in a particular rhetorical context" (Silva, 2002, cited in Maghfiroh, 2013). In other words, units of discourse are put together under a certain principle of connectedness, and under that principle, there are two important elements: coherence and cohesion

(Maghfiroh, 2013). Cohesion is defined as the linguistic feature which helps to make a sequence of sentences a text (Halliday & Hasan, 1976). According to Halliday and Hasan (1976), there are five cohesive relations identified in a text: reference, substitution, ellipsis, conjunction, and lexical cohesion. Coherence, on the other hand, refers to the overall discourse level property of unity, or how well a text holds together (Hasan, 1984; van Dijk, 1980). Kintsch and Van Dijk (1978) argues that the semantic structure of a text is characterized on two levels: namely, at the levels of microstructure and macrostructure. The microstructure refers to the local level of the discourse, or the structure of the individual propositions and their relations. The macrostructure refers to the global level of connectivity. A discourse is coherent only if its respective sentences and propositions are properly connected, and if these propositions are organized globally at the macrostructure level (Kintsch & van Dijk, 1978). Having confirmed the definitions of coherence and cohesion used in this study, the proposed four analytical frameworks will be described in the next section.

## **2. The four analytical frameworks**

### **2.1. Structural-Functional Analysis**

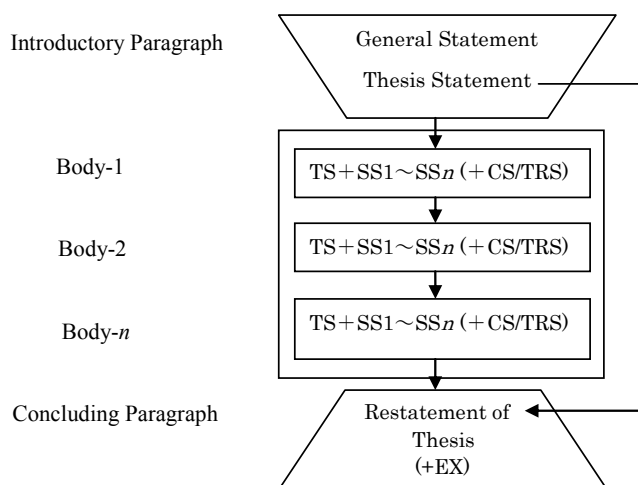
The basic structural organization of an essay is first looked into. Figure 1 is a diagrammatic representation of a typical structure of English essays (Oshima et al., 2006; Someya, 1994). As this diagram indicates, an essay consists of three structural parts: an introductory paragraph, a body, and a concluding paragraph. The body can include as many numbers of paragraphs as required, but three is the most typical and recommended number of body paragraphs (Someya, 1994). Each paragraph also has three major functional parts: a topic sentence (TS), supporting sentences (SSs), and either a concluding sentence or a transitional sentence (CS/TRS). Each SS can be supplemented by extensions (or EXs) which elaborate on the SS. The TS and SS are obligatory, but the CS, TRS and EX are optional. The introductory paragraph generally starts with a general statement (or, a hook) followed by a thesis statement which is an obligatory component. The thesis statement is usually restated in the first part of the concluding paragraph in the form of a summary. The EX at the end of the concluding paragraph is optional, but students are recommended to conclude an essay by offering a suggestion, giving an opinion, or making a prediction relevant to the thesis statement.

With this model, the author first conducts what she calls “Structural-Functional Analysis” to examine whether essays have these structural and functional components. Her assumption is that this simple analysis will provide a good indicator as to the successful organization of an essay.

### **2.2. Keywords-chain Analysis**

The second analysis focuses on the content, in particular the topic of propositions, or the topical/thematic coherence (Hymes, 1974). This is because the topic is a crucial aspect of context (Hymes, 1974) and governs how language is used to help the reader understand the context smoothly (Cazden, 1970 cited in Watson, 1998). The identification of topics relies on the identification of predominant keywords (Scott, 1997). Key lexical items recur with a frequency indicative of topic prominence, and Watson (1998) approached topic identification by drawing up a

semantic network of lexical items. He made up line diagrams with key lexical items categorized by Halliday and Hasan (1976) in order to highlight the semantic relations between the topics.



**Figure 1: Typical structure of an essay**

In this study, two types of keywords are investigated: theme-setting keywords and argument-setting keywords. The former is related to the given theme of an essay or the one chosen by the writer, and the latter is related to the arguments developed under the theme. They are the words which appear recurrently, often in their inflected forms, synonyms or near-synonyms, or in different words that nevertheless belong to the same semantic category. As for the words in the same semantic category, this study also tries to investigate the items related to content schemata (Hudson, 1982, Watson, 1998). For example, *studies*, *grades*, *class*, and *course* in the second paragraph of Student A's essay (Appendix A) are all considered to belong to the "school schema;" thus, they are in the same semantic category. These keywords are connected with lines, which result in a "keywords-chain" – a visual representation of the connectivity or cohesion of the target text (Halliday & Hasan, 1976). The hypothesis here is that a well-connected network chain of keywords is an indicator of the essay being well developed.

### 2.3. Metadiscourse Markers (MDM) Mapping

Adopting the idea presented by Hyland (2005), which is "writing or speaking is viewed as a social and communicative process between writers or speakers and readers or listeners" (p. 3), this study investigates the writer's linguistic strategy with readers in mind. These strategies are usually manifested in a text by the use and distribution of so-called "metadiscourse markers" or MDMs (Hyland, 2005). Among them, frame markers including sequencing (*firstly*, *secondly*, etc.) and transition markers (*therefore*, *however*, *moreover*, *in addition*, etc.) are mainly focused on in this current investigation since students have been taught in class that an appropriate usage of transition signals or conjunctions plays an important role in creating cohesion in paragraphs and throughout a

whole essay. Other MDMs including those categorized as interpersonal MDMs are not included in the current pilot study.

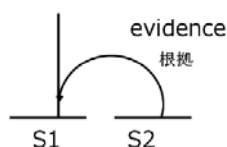
## 2.4. Rhetorical Structure Theory

The last framework used here is Rhetorical Structure Theory (RST hereafter). This is a theory of text organization originally proposed by Mann and Thompson (1988). The basic idea of RST is that a text can be divided into parts, and the theory describes how they can be arranged and connected to form a cohesive text. In other words, this theory basically looks into the microstructure of sentence-level organization. However, it also relates to the macrostructure, or the coherence of a whole text.

In RST, relationships between parts of a text, termed *text spans*, are crucial in making a single unit of a text, and they can be clauses, sentences, or units larger than sentences. A text is generally broken into two spans: namely, *nucleus* and *satellite*. The former presents the more essential information in the text, while the other (subordinate span) presents supporting or background information (Mann & Thompson, 1988). The original list of rhetorical relations of the spans introduced 25 relations; however, this was subsequently expanded to 75 relations (see Carlson & Marcu, 2001). The relations between the following sentences, for example, can be described as in Figure 2 by using an RST diagram. These two sentences are connected by the relation called *Evidence*; S1 is the main statement, or nucleus, and S2, the supporting evidence to S1, or satellite.

S1: Jogging is not as easy as it appears.

S2: Ninety-seven percent of people cannot jog three miles without stopping.



**Figure 2. The RST Diagram showing the logical relationship between S1 and S2**

The RST framework was originally developed to analyze coherent texts written by native speakers of English, or otherwise “well-constructed” pieces of texts written by non-native writers for cross-linguistic comparison. The use of RST was considered unsuitable for analyzing L2 learners’ texts, which can be full of logical anomalies; however, Skoufaki (2009) attempted to use RST for coherence error or anomaly detection in 45 paragraphs written by Chinese ESL students. Following his research, this study seeks to examine the applicability of the theory to Japanese students’ English texts in order to identify coherence breaks. In this study, a sentence is regarded as a basic unit of span, and the list of RST used here is the original version proposed by Mann and Thompson (1988). The author’s assumption is that there may be some sentence relations that cannot be neatly labeled by the original 25 relation labels.

### 3. Research Questions

The present author formulated three RQs for this study.

1. Are the four analytical frameworks effective in analyzing structural and rhetorical features of English essays written by Japanese students?
2. Is RST useful to identify “coherence breaks” unique to Japanese students’ argumentative essays?
3. Are the “coherence breaks” identified in students’ English essays specific to their English essays, or are they also observed in their Japanese essays?

### 4. Subjects and Procedure of analysis

The bilingual essays written by two students in the classes of 2013 were chosen for this pilot research. Student A is a female student from G-class, and Student B is a male student from L-class. They were selected because they represented typical students in their respective classes. The data used for the analysis were from argumentative essays written about Topic 12: “Part-time Job.” Students were required to take a position over the proposition “Should college students have a part-time job or not?” and they had to defend the position they took. This particular topic was chosen because the author believes that argumentative essays are more suitable for the proposed analysis than descriptive essays.

Three researchers conducted analyses according to each of the four frameworks. As for the first three frameworks, only the English essays were analyzed. As for RST analysis, both the English and Japanese essays were examined to see the relationships between L1 and L2. The first researcher, a Japanese college teacher, mainly led the analysis of both the English and Japanese essays, and the second researcher, an American writer, examined the English essays. The third researcher, who is Japanese and teaches Japanese to foreign students, analyzed the Japanese essays. When disagreement occurred, they conferred until they reached an agreement. All the surface errors in the texts remained unchanged.

## III. FINDINGS

### 1. Structural-Functional Analysis

First, the basic structure of the essays was analyzed. In class, students were encouraged to organize their ideas along with this typical essay framework. Thus, they understood the importance of the *introduction-body-conclusion* structure to help their arguments flow logically.

Student A’s essay consists of 5 paragraphs while Student B’s essay comprises 4 paragraphs. Both of the English essays are appropriately structured from introduction to conclusion. Student A’s essay (see Appendix A) begins with both the proponents’ view, which is her position, and the opponents’ view in the introduction paragraph, followed by evidence to support these two positions in the 1st and 2nd body paragraphs. After inserting a counterargument to provide opponents with a solution in the 3rd body paragraph, she restates her position in the concluding paragraph. Student B (see Appendix B) also takes the proponents’ position, and compares both of the positions in the 2nd and 3rd paragraphs. Similar to Student A, he briefly offers suggestions to those who oppose the

proposition at the end of the body paragraph, then he concludes his essay with a restatement of his position. As for the functions of the sentences in each paragraph, however, a big difference can be found between the two. Table 1 shows the number of TS, SS and CS, which constitute the English essays written by Students A and B. Neither of them included the TRS or EX.

**Table 1. Number of sentences per paragraph according to their functions**

Paragraph	A					B				
	TS	SS	CS	Non-SS	Total	TS	SS	CS	Non-SS	Total
1	1	3	0	0	4	1	0	0	4	5
2	1	3	1	0	5	1	4	0	1	6
3	1	5	1	0	7	1	4	0	3	8
4	1	2	1	0	4	1	0	1	5	7
5	1	2	0	0	3	-	-	-	-	-
Total	5	15	3	0	23	4	8	1	13	26

Note: TS= topic sentence; SS= supporting sentence; CS= concluding sentence

Non-SS= sentences which do not function properly as a supporting sentence

Both of the essays have one TS in each paragraph. In Student A's essay, each body paragraph has a CS. She starts with her opinion in the TS and either restates or summarizes it in the CS. Student B's essay, on the contrary, has only one CS. The big difference between the two here is that while the rest of the 15 sentences are all functioning as supporting sentences in Student A's essay, there are only 8 SSs in Student B's essay, and 13 sentences are considered non-SSs. In other words, these 13 sentences are not fully explaining, describing, or giving reasons, facts or examples to support the topic sentence in each paragraph. In particular, in both the introductory and concluding paragraphs, there are 4 and 5 sentences respectively that are regarded as non-SSs, making no meaningful contribution to the entire essay. In order to identify this particular problem in more detail, the semantic roles of each sentence will be examined in Section 4.

## 2. Keywords-Chain Analysis

The English essays written by both A and B have theme- and argument-setting keywords so that the readers can follow what they are discussing to support their argument. Closely examined, however, there is a difference between A and B. A's essay has more keywords than Student B's, creating a more complicated network chain of keywords (see Appendices A and B). Table 2 shows both theme- and argument-setting keywords used in Student A's essay and Student B's. The numbers in brackets indicate frequency. The words and/or phrases in the same meaning groups, such as *college student(s)* and *student(s)* or *study* and *studies*, are lumped together into the same word type.

Both of the theme-setting keywords in A and B are *college student(s)* and *part-time job(s)*. *Part-time job(s)* and *work (working, work for part-time)* are intertwined as both theme-setting keywords and argument-setting keywords in both of the essays as indicated by line types 1 and 2 (see notes to Appendix B). These words are contrasted with *study (studying, academics)* in line type

3, indicating that their essays are debating between the positive influence of work (part-time job) and the negative influence on study (at school).

**Table 2. Keywords in Students A and B**

	Student A	Total	Student B	Total
Theme-setting keywords	<i>college student(s)</i> [14] <i>part-time job(s)</i> [16]	30	<i>college student(s)</i> [8] <i>part-time job(s)</i> [17]	25
Argument-setting keywords	<i>Work</i> [16] <i>study</i> [10], <i>experience</i> [9] <i>time</i> [5]	40	<i>Work</i> [17] <i>study (academics)</i> [7] <i>useful</i> [2] <i>knowledge</i> [4]	30

What is noticeable with Student A is that all of the keywords which appear in the 1st paragraph continuously appear in other paragraphs with a variety of vocabulary or in other forms, and the lines are linked from beginning to end. It indicates that the keywords are effectively used and distributed properly in order to construct cohesiveness in her essay. For example, one of her argument-setting keywords, *experience*, appears throughout her essay, because she emphasizes the importance of experience gained from a part-time job as a supporting reason for her opinion in the 1st paragraph. She discusses it in the 3rd paragraph, and restates it in the last paragraph. In addition, the words belonging to the “school schema,” such as *studies*, *class*, *grades* and *course*, are effectively used to deliver the opposing view in the 2nd paragraph.

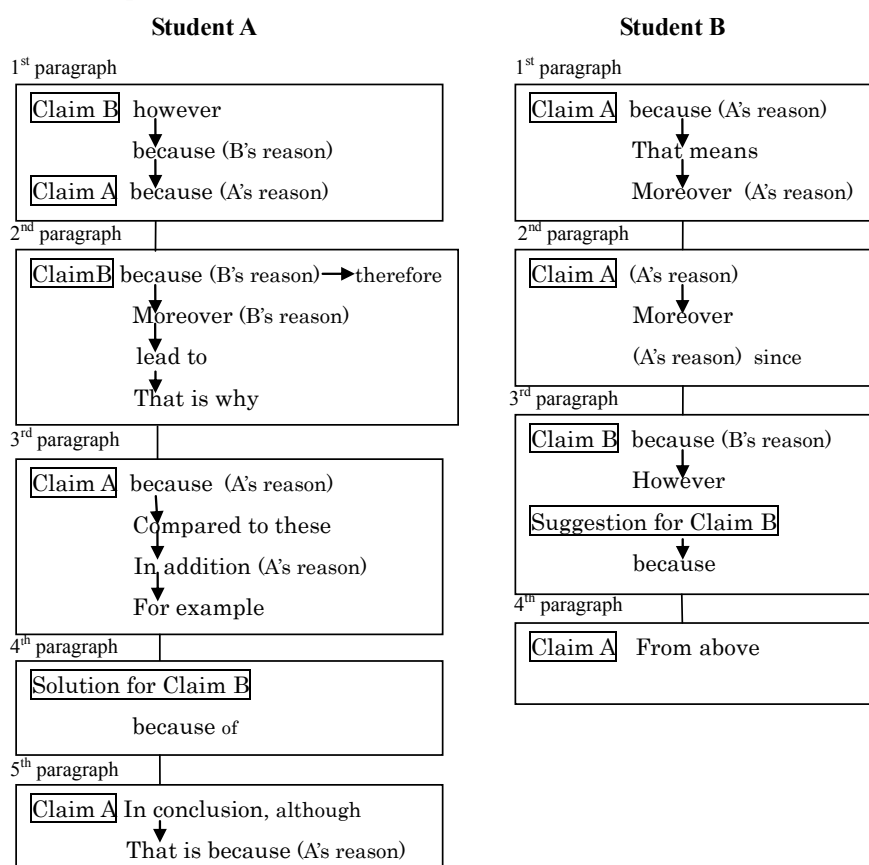
On the other hand, in the essay written by Student B, although *work* and *study* appear in almost equal number with Student A as argument-setting keywords, the frequencies of *knowledge* (see line type 4) and *useful* (see line type 5) are low, and they are repeated only four times and twice respectively. *Useful*, which appears in his thesis statement, and reappears in the 2nd paragraph, could be an important argument-setting keyword; however, it is not used in other paragraphs. This indicates that he failed to focus on the usefulness of part-time work but instead jumped off in a different direction, which resulted in a major coherence break in his argument.

### 3. Metadiscourse Markers (MDM) Mapping

As for the use of metadiscourse markers, there is a stark contrast between Students A and B. The numbers of MDMs are 15 and 9 respectively, and they are all frame markers which are mainly transition markers. Figure 3 shows MDM mapping of Student A’s essay and that of Student B’s essay. At a glance, Student B has a simpler distribution of MDMs than Student A, with the use of *because*, *moreover*, and *however* most evident. The writer tries to present reasons for both the proponents’ view (claim A) and opponents’ view (claim B) using *because* and adds more evidence with *moreover*. Although the usages of these transition markers are basically appropriate, the logical structure of his essay appears to be rather simple and the variation of MDMs is quite limited. A simple examination of the use of MDMs shows that the argument is not fully developed in his essay.

The MDM mapping of Student A, on the other hand, indicates that her essay has a more

layered structure of argument or in-depth discussion. The frequent use of *because* indicates both positions are supported by sufficient reasons or evidence. In addition, she uses a variety of transition markers appropriately and effectively to construct a logical flow for her argument. For example, in the 2nd paragraph, she starts with the opponents' view and gives two reasons using *because* and *moreover*. The use of *lead to* implies a significant consequence which results from having a part-time job, and concludes the paragraph with *that is why* in order to strengthen the opposing view. The 3rd paragraph is discussed with a sequence of transition markers *because*→*compared to*→*in addition*→*for example* in this order. The order of these transition markers also indicates that the proponents' view is fully discussed, providing enough evidence to support her position. This effective usage of transition markers implies that her essay is more or less convincing to the readers, compared to that of Student B.

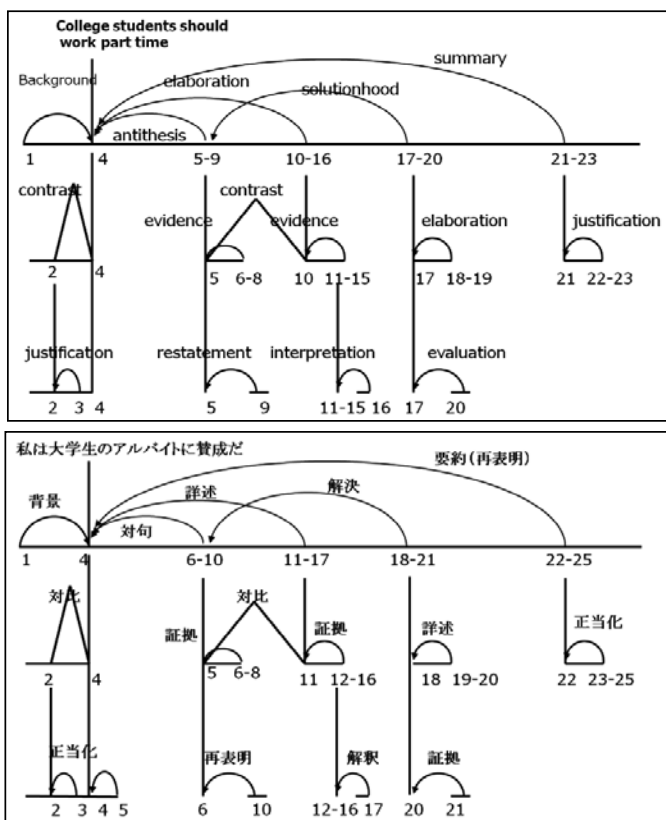


**Figure 3: MDM mappings of Student A's essay and Student B's essay**

#### 4. RST analysis

The rhetorical relations between the sentences in both the English and Japanese essays were analyzed using RST. Figure 4 shows that rhetorical structures of both of the essays written by Student A are completely matched. All the sentences are logically linked, by means of comparing contrasting claims (*Contrast*), providing evidential support (*Evidence*), and cementing her claim

with a *Restatement* of her opinion. She also provides *Interpretation* and *Elaboration* of what she wants to say in order to persuade the readers to fully understand her argument.



**Figure 4: The RST diagrams of Student A's essays**

*Note:* The RST diagram of the English essay is above while the Japanese essay is below.

Student B's essay, on the other hand, exhibits five major "anomalies" in his RST diagram as shown in Figure 5. These five parts are not connected logically to any of the other parts of his English essay. These are typical "coherence breaks" which Skoufaki (2009) calls "dangling units of analysis." Skoufaki (2009) identified coherence breaks indicated by RST diagram abnormalities among paragraphs written by Taiwanese lower-intermediate learners of English, and found that these happen when the paragraph contains "irrelevant content," "incomprehensible content," and suffers from what he calls "self-sufficiency." With five dangling parts closely examined, three types of coherence breaks can be detected. In the 1st and the 5th dangling parts, which are marked with circles in Figure 5, "irrelevant ideas" are inserted. In order to see exactly what has happened, Student B's essay needs to be examined in more detail.

Table 3 below shows English sentences and their corresponding Japanese sentences in the 1st paragraph of Student B's essays. As this indicates, Student B started his English essay with the thesis statement in E-S1 and the reader expects the writer to explain how useful a part-time job is

from E-S2 onward. However, Student B writes about a situation where part-time workers are treated

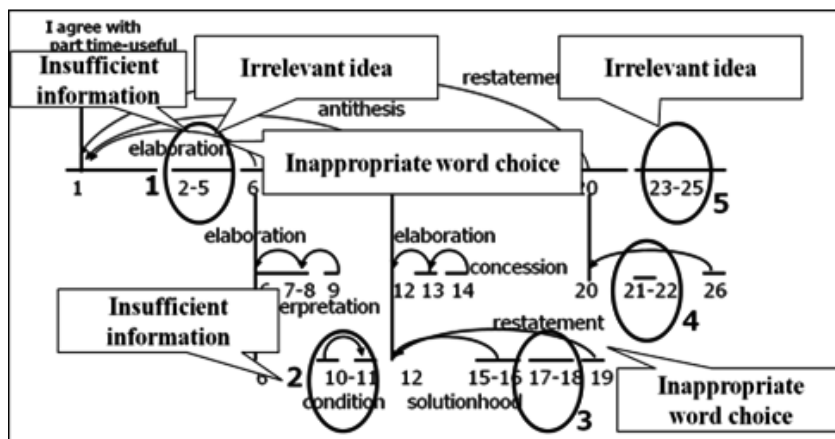


Figure 5: The RST diagram of Student B's English essay

in Japan as if they were regular employees. The 4th sentence “I define part-time job as a study.” further confuses readers about what the writer wants to say. What he intends to state here is that a part-time job provides an opportunity where students can learn something about society, or *shakaibenkyo* in Japanese. These sentences which include the idea of a part-time job being also *shakaibenkyo*, are not, however, effective enough to support the thesis statement. This idea can therefore be considered an irrelevant idea or a new topic abruptly inserted. This confusion is further aggravated by inappropriate word choice or missing definitions of the word chosen. Since he repeatedly uses the word ‘study,’ he should have defined it in the first paragraph; otherwise, readers might be confused as to whether ‘study’ means *shakaibenkyo* or schoolwork. A similar phenomenon can be detected in the 5th dangling part in Figure 5.

Table 3. Sentence Alignment between English and Japanese Versions (Student B)

E-S1	I agree with part-time job that student have because part-time job is useful for students after graduation.	J-S1	学生がアルバイトをすることに賛成である
		J-S2	というのも、学生時代のアルバイトは卒業して社会に出たときに役に立つであろうと考える
E-S2	In Japan, those who work part-time are often treated as company members.	J-S3	日本ではアルバイトの人も社員のごとく扱うことがしばしば見られる。
E-S3	That means part-time job is a part of real work.	J-S4	つまり、アルバイトとはいえど仕事であるのだ。
		J-S5	学生の仕事とは学ぶことである。
E-S4	I define part-time job as study and this meaning will be explained in last paragraph.	J-S6	アルバイトを勉強と位置付けようではないか、この意味は最終段落で述べることにする。
E-S5	Moreover, task of students is to study and to get a lot of knowledge.		

The second problem is “insufficient information” which is most evident in the 2nd dangling part in Figure 5. The paragraph should have provided evidence to support the opinion or claim

given in E-S9; however, the writer failed to do so in the subsequent sentences as shown in Table 4.

**Table 4. Sentence Alignment between English and Japanese Versions (Student B)**

E-S9	This knowledge will be useful after graduating school.	J-S10	ここで得た知識は卒業してから就職したときに必ず役に立つであろう。
E-S10	If you have no part-time job till graduating school, you may be panic when you face difficulties.	J-S11	もし、在学中アルバイトを一つもせずに卒業し就職してから困難に直面した時が大変であろう。
E-S11	Part-time job enable them to solve difficulty easier than those who have no part-time job since those who have part-time job know how to solve these problems.	J-S12	しかし、アルバイトをしていれば問題の解決もできるであろう。
		J-S13	アルバイトで得た知識が問題の解決方法方法である。

As one of the advantages of having a part-time job, the writer introduces his opinion about how the knowledge acquired through a part-time job is useful to overcome possible difficulties in the workplace. These sentences, however, are short of persuasive evidence because of a lack of an explanation about the *difficulties* referred to in E-S10. In addition, the writer needs to mention the logical linkage between the *knowledge* acquired through a part-time job and overcoming difficulties, or how useful this knowledge is in overcoming these difficulties.

The third problem is “incomprehensible contents due to wrong word choice.” As previously mentioned, some words such as *study* or *knowledge* are used without definition. As in the 3rd and the 4th dangling parts in Figure 5, three words, *plan*, *merit*, and *demerit* appear suddenly, which confounds readers as to what these words are referring to. This ambiguity in meaning makes content incomprehensible, and also breaks the lexical cohesion.

When looking into the Japanese essay written by Student B, on the other hand, interesting phenomena can be observed. There does not seem to be a significant breakdown of coherence in the RST diagram in Figure 6, and all the parts appear to be connected. In Japanese, readers can infer what the writer intends to say even without the intention being specifically explained. As for the first coherence break in his English essay, the corresponding Japanese sentences from J-S4 to J-S6 in Table 3 can be literally translated to “A part-time job can be considered as work in the society. A students’ job is to study; thus, a part-time job can be defined as study.” Although these sentences in Japanese contain some ambiguity, readers can infer that the writer wants to say “a part-time job helps students learn something about society; thus, a part-time job can be regarded as one type of *shakaibenkyo*. (Therefore, a part-time job is useful.) The word *manabu* (to study or to learn) has an ambiguity in its meaning here, but readers can infer *manabu* to mean *shakaibenkyo* from the context in Japanese. As in the case of the 2nd dangling part, without an explanation of *difficulty* or *knowledge* gained from a part-time job in detail, Japanese readers can try to understand what *knowledge* is expected to be gained, or what *difficulty* students expect to face in the workplace from common experience or from the context. With regard to inappropriate word choice, the word *study*, which is repeatedly used in the essays, confuses readers in English but not in Japanese because the word *manabu* contains two connotations of either schoolwork or *shakaibenkyo* and the reader can

infer which meaning is being used from the context.

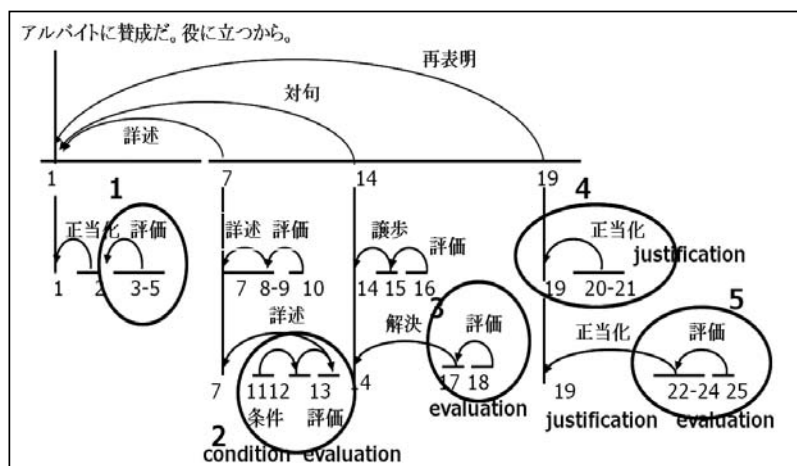


Figure 6: The RST diagram of Student B's Japanese essay

#### IV. DISCUSSION

The investigation into L1 and L2 texts written by two students with the four frameworks has revealed some contrasting results. The possible reasons are described below.

The English essay written by Student A is structurally and logically organized with appropriate choice of both theme-setting and argument-setting keywords, which are linked across the paragraphs, demonstrating that her argument is well developed. Her RST diagrams do not show any particular coherence breaks. Previous research on L2 writing (Kraples, 1990; Hirose & Sasaki, 1994) has shown that the variables influencing the quality of L2 writing include L2 proficiency, L1 ability, composition competence, meta-linguistic ability (awareness of the system of the language), meta-knowledge (strategy of how to write), etc., and complicated interactions among many variables determine the quality of L2 texts. Taking these variables into consideration, Student A is considered to have a high composition competence in both L1 and L2. Also, with the order of compositional assignment from English to Japanese in mind, she might have a high meta-linguistic ability to appropriately express herself in both languages. In addition to her relatively high L2 proficiency, she learned how to write English essays in her SA program the year before as well as on the current course in which she was enrolled. Accordingly, she seems to have already acquired the basic prerequisites to be a good writer and was able to divert her attention to both micro (vocabulary or grammar) and macro (rhetorical structure) levels of writing. The English essay written by Student B has more or less the proper organizational structure of an argumentative essay. However, his essay has weak lexical linkage, especially in terms of the distribution and connectivity of argument-setting keywords. In addition, three types of coherence breaks have been detected in the RST analysis of his English essay as discussed in the previous section. These coherence breaks were also identified among the English essays written by other ESL students of different nationalities such as Swedish (Wikborg, 1990), Chinese (Skoufaki, 2009), and Indonesian

(Maghfiroh, 2013). The last of these three studies revealed “irrelevant ideas” or “disunity of thought” as a major cause of problems with writing in general and with coherence and cohesion in particular, and the author suggests transfer from L1 to L2 as accountable. In the case of Student B, both his L2 proficiency and meta-linguistic ability need to be improved. In writing in English, he needs to streamline the information and focus on what he wants to write the most. For instance, the word *useful* could be an argument-setting keyword in the thesis statement; thus, his essay could have focused on how useful a part-time experience was. Then, he could have described concretely about the knowledge and skills gained through a part-time job in separate body paragraphs. It is also necessary for him to use the words whose meaning he knows well and use them in a proper context.

The difference between L1 and L2 in the RST diagram of Student B is attributable to the differences in the basic logical frame between English and Japanese. According to Hall (1976), English has a high-context culture where messages need to be conveyed explicitly, while Japanese is a low-context culture where messages are often expressed implicitly. Such factors as topic shift, vagueness, organization of ideas without a clear linkage, and a shortage of explanation found in Japanese written texts are not necessarily regarded as fatal errors, and the messages can be understood from inference in the context (Tomioka, 2003). These factors, on the contrary, are regarded as being inappropriate or illogical in English texts. Student B’s English essay exhibits this particular problem. He needs to learn how to express himself explicitly in English with his target readers in mind. Although students write essays from English to Japanese in this current writing class, they probably construct their ideas in Japanese. Under this assumption, his Japanese essay may have some ambiguity in the content, but this needs to be re-interpreted so as to convey exactly what the writer intends to say in his English essay. In other words, when learning L2, learners need to be aware of the difference between their L1 and L2 in terms of logical frame, and need to acquire mediation-ability<sup>4</sup> or an ability to negotiate “meaning” between two languages. In this sense, writing in both L1 and L2 is meaningful in nurturing the mediation ability of the students. By writing bilingually in this class, there have been some students who say L1 writing is useful to review their English essay or vice versa. This is an interesting phenomenon; however, the positive effects of writing bilingually need to be further investigated.

Lastly, the answers to the research questions are described.

RQ1: Are the four analytical frameworks effective in analyzing structural and rhetorical features of English essays written by Japanese students? -- The current study proposed four frameworks for the analysis of English argumentative essays: 1) Structural-Functional Analysis, 2) Keywords-Chain Analysis, 3) Metadiscourse Markers (MDM) Mapping, and 4) Rhetorical Structure Theory (RST). The basic structural/functional properties of the sentences in each paragraph can be explored by the Structural-Functional Analysis. Cohesion of a text and the logical development of the argument can be observed by the Keywords-Chain Analysis and MDM Mapping. These investigations are important to determine if the essay in question is properly constructed or not. However, without a semantic analysis of sentences with RST, the logical problems cannot be thoroughly recognized. In this study, the non-SSs detected from the first framework were found to

coincide with the coherence breaks identified in the RST analysis. The author concludes that the four frameworks altogether are very useful in analyzing a text for both structural and logical properties.

RQ2: Is RST useful in identifying “coherence breaks” unique to Japanese students’ argumentative essays? -- The results show that RST can be applied to the analysis of texts written by Japanese students, and can be used to identify coherence breaks. Due to its prescriptive nature, RST is best suited to the analysis of well-constructed texts including those written by native speakers of English. The question was whether it can be used for the analysis of L2 learners’ texts (especially those by Japanese college students) whose logical properties are less than satisfactory and often deviate largely from western standards. This pilot study has indicated that the RST framework is still applicable in detecting non-standard rhetorical relations. The author believes that the usefulness of RST can be expanded if such non-standard rhetorical relations are clearly identified and corresponding “relation labels” are added to the existing relations set. This needs extensive study with larger sample sets.

RQ3: Are the “coherence breaks” identified in students’ English essays specific to their English essays, or are they also observed in their Japanese essays? -- The coherence breaks found in this study were specific to the English essay and not found in the Japanese essay. This can be ascribed to the differences in the basic logical frame between English and Japanese; however, further investigation into more bilingual texts is needed to give any definite answer to this question.

## **V. LIMITATIONS AND PEDAGOGICAL IMPLICATIONS**

Given that this is meant to be a pilot study, there are nevertheless several limitations to this current study. One limitation was that RST analysis tends to be subjective. Although researchers conferred when there was a disagreement, to fully understand the relations presented by RST is a quite complicated and tedious task. The second limitation is that since the essays written by only two students were analyzed this time, the difference in their English proficiency levels stood out. The author believes that the proficiency level is only one factor influencing the quality of L2 essays, and the result of this study implies that even if the proficiency level improves, the quality of L2 essays will not improve among the students who are trapped in a Japanese logical frame. In fact, in this on-going analysis, it has been found that some of the essays written by other G-class students also have several coherence breaks; on the other hand, even among L-class, there are students whose essays are written without any major logical breaks. The third limitation is about text analysis. Considering the numerous variables involved in order to determine the quality of L2 texts, an evaluation of essays needs to be included, in addition to interviews with writers to know more about their intended meaning in the texts as well as their educational background.

As for the pedagogical implications, the author suggests that the frameworks presented in this study be used in writing classes. Although the basic structure of an essay and the functions in each paragraph are taught to the students, they often fail to apply them to their essays. By having them analyze their own text using these frameworks, students will be able to acquire an ability to analyze

their own texts, as well as a self-editing ability. RST analysis might be difficult to implement, but the first three frameworks can be introduced when their essays are reviewed in class. Another implication for writing classes is the benefits of writing essays bilingually. In the writing class at Kansai University, we believe that students will be able to improve their linguistic competence and metacognitive skills by writing essays bilingually. As previously mentioned, past research suggests that improved writing competence in L1 is transferable to that of L2 and vice versa. The bilingual approach like this may be more effective in improving the quality of essays than simply writing in English alone, although this hypothesis has yet to be proven. In this paper, the English essays written by two students were analyzed based on four analytical frameworks, and their Japanese counterparts were also examined. It has been demonstrated that the four frameworks are useful for the proposed analyses. Based on this result, the author has already started to conduct a large scale investigation in order to identify structural and rhetorical features or problems of English essays written by Japanese college students. She hopes that this study on writing bilingually will make a favorable contribution to second language writing education in Japan.

## ACKNOWLEDGEMENT

I would like to express my gratitude to Professor Yasumasa Someya and the anonymous reviewers.

## NOTES

- <sup>1</sup> Part of the present study was presented at JACET Kansai Chapter 2014 Spring Conference held at Osaka University of Pharmaceutical Sciences in June 14th, 2014.
- <sup>2</sup> The same description for the writing classes and the details of the project are found in Yamashita, M. (2014). The Design, Development and Research Potential of Kansai University Bilingual Essay Corpus. *English Corpus Studies*, 21, 19-35.
- <sup>3</sup> The essay topics are 1) Environmental pollution, 2) Violence on TV, 3) Young people today, 4) Suicide, 5) Sports, 6) School education, 7) Recycling reusable materials, 8) Money, 9) Divorce, 10) Death penalty, 11) Crime, 12) Part-time job, and 13) Smoking.
- <sup>4</sup> The idea of “mediation ability” was first included in CEFR (Common European Framework of Reference for Language) and this study interprets this term as “negotiation of meaning between different languages.”

## REFERENCES

- Carlson, L. & Marcu, D. (2001). *Discourse Tagging Reference Manual*. ISI Technical Report. ISI-TR-545. Retrieved from <http://www.isi.edu/~marcu/discourse/tagging-ref-manual.pdf#search='Discourse+Tagging+Reference+Manual>
- Hall, E. T. (1976). *Beyond Culture*. New York: Anchor Press.
- Halliday, MAK. & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hasan, R. (1984). Coherence and cohesive harmony. In J. Flood (Ed.), *Understanding reading comprehension* (pp.181-219). Newark, DE: International Reading Association.
- Hinds, J. (1983). Contrastive rhetoric: Japanese and English. *Text*, 3(2), 183-195.
- Hinkel, E. (1997). Indirectness in L1 and L2 academic writing. *Journal of Pragmatics*, 27, 361-386.
- Hirose, K. (2005). *Product and Process in the L1 and L2 Writing of Japanese Students of English*

Hiroshima: Kisuisha.

- 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.
- Hudson, T. (1982). The effects of induced schemata on the "short circuit" in L2 reading: non-decoding factors in L2 reading performance. In P. Carrell, J. Devine & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 183-205). Cambridge: CUP.
- Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.
- Hymes, D. (1974). Ways of speaking. In R. Bauman & J. Sherzer (Eds.), *Explorations in the ethnography of speaking* (pp.433-452). Cambridge: CUP.
- Kaplan, R.B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, 16, 1-20.
- Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363-394.
- Kraples, R. (1990). An overview of second language writing process research. In B. Kroll (Ed.), *Second Language Writing* (pp. 37-56). New York: CUP.
- Kubota, R. (1998). An Investigation of L1-L2 transfer in writing among Japanese university students: Implications for contrastive rhetoric. *Journal of Second Language Writing*, 7, 69-100.
- Maghfiroh, S. (2013). *Coherence Breaks in Essay Made by the 6<sup>th</sup> Semester Students of English Department of STAIN Salatiga in the Academic Year of 2012/2013*. Retrieved from <http://eprints.stainsalatiga.ac.id/view/subjects/PBI.html>.
- Mann, W. C. & S.A.Thompson. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *TEXT*, 8, 243-281.
- Oshima, A. & Hogue, A. (2006). *Writing Academic English*. New York: Pearson Education
- Oi, K. (2005). Comparison of Argumentative Styles: Japanese College Students vs. American College Students- An Analysis Using the Toulmin Model. *JACET Bulletin*, 41, 85-100.
- Scott, M. (1997). PC analysis of key words-and key key words. *System*, 25, 233-245.
- Skoufaki, S. (2009). An Exploratory Application of Rhetorical Structure Theory to Detect Coherence Errors in L2 English Writing: Possible Implication for Automated Writing Evaluation Software. *Computational Linguistics and Chinese Language Processing*, 14(2), 181-204.
- Someya, Y. (1994). 「第 1 5 週 : 英文レポートの書き方とパラグラフライティングの技法」『ライティングマラソン』4, 121-176. 東京 : アルク
- Tomioka, T. (2003). 『論理思考を鍛える英文ライティング』 東京 : 研究社
- van Dijk, T. A. (1980). *Macrostructures*. Hillsdale, N.J.: Erlbaum.
- Watson, T. (1998). Topic-based analysis of classroom discourse. *System*, 26(3), 303-318.
- Wikborg, E. (1990). Types of coherence breaks in Swedish student writing: Misleading paragraph division. In U. Connor and A. M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 131-148). Alexandria: TESOL.

## Appendix A: Keywords-Chain Analysis (Student A's Essay)

Title: College Student should Work at a Part-time Job

<1<sup>st</sup> paragraph>

1. In Japan, a lot of **college student** work at part-time jobs.
2. Some people say **college students**, however, should stop working.
3. That is because these people think that **college students** cannot concentrate on their **studies**.
4. In my opinion, **college students** should experience working at part-time jobs because they can gain **experience** that they cannot attain only by **studying** in the university.

<2<sup>nd</sup> paragraph>

5. Some people say that **college students** cannot perform well at **school** because they work at part-time jobs.
6. It is said that **college students** use all of their free **time** for working and therefore they do not have the **time** to review for their academic **studies**.
7. Moreover, some of them prioritize their part-time job that sometimes they do not attend their **class**.
8. These lead them to get poor **grades** at their **studies** and some even fail their **course**.
9. That is why some people believe part-time jobs are not good for **college students** to do.

<3<sup>rd</sup> paragraph>

10. I believe **college students** should work for part-time job because they can **experience** things that they cannot do inside the university.
11. In the university, **students** can **experience** to **study** academic things, join clubs or circles and participate in the university's events.
12. Compared to these, **students** can **experience** the hardships to earn money when they work.
13. In addition, the people they communicate with at part-time jobs are not only teachers and fellow students.
14. The **experiences** they get from part-time jobs can help them improve themselves in different ways.
15. For example, they can learn how to communicate effectively with people who have different range of age.
16. This cannot be **experienced** by only **studying** inside the university.

<4<sup>th</sup> paragraph>

17. The appropriate solution for **college students** who fail at their **studies** because of part-time jobs is to make these **students** learn how to discipline themselves.
18. They should learn how to manage their **time**.
19. It is important for them to know how much **time** they should use for **studying** and for working at part-time jobs.
20. Making them stop working will only result for them to **lose** the chance of **experiencing** many things outside the university.

<5<sup>th</sup> paragraph>

21. In conclusion, although some people say that part-time jobs lower the performance of **college students** at their **studies**, I believe these **students** should still work for part-time job.
22. That is because they can gain a lot of valuable **experiences** in their job.
23. The only thing **students** should be careful when they are working at part-time jobs is to learn how to divide the **time** they use for work and for **study**.

## Appendix B: Keywords-Chain Analysis (Student B's Essay)

Title: Significance of part-time job

<1<sup>st</sup> paragraph>

1. I agree with part-time job that **students** have because part-time job is useful for **students** after graduation.
2. In Japan, those who work part-time are often treated as company members.
3. That means part-time job is a part of real work
4. I define part-time job as study and this meaning will be explained in last paragraph.
5. Moreover, task of **students** is to **study** and to get a lot of **knowledge**.

<2<sup>nd</sup> paragraph>

6. Part-time job is important for **students**.
7. They can learn a lot of things, such as communication skill, patience and importance of money through part-time job.
8. Moreover, it teaches them what is working.
9. This **knowledge** will be useful after graduating school.
10. If you have no part-time job till graduating school, you may be panic when you face difficulties.
11. Part-time job enable them to solve difficulty easier than those who have no part-time job since those who have part-time job know how to solve these problems.

<3<sup>rd</sup> paragraph>

12. Some people say **students** should not work part-time because priorities of them are first **academics** and second others.
13. Part-time job may make **students** negligence of **academics**.
14. Surely, the argument is correct and I grant it.
15. However, this argument is a little irrelevant and part-time job never conflict with **study**.
16. It is difficult to say **students** should not work part-time because they should work part-time and make an effort not to be negligence of **academics**.
17. The plan enables them to protect their priorities.
18. It clearly outweighs the argument.
19. My plan makes an effort not to be negligence of **academics** and can get a lot of **knowledge**.

<4<sup>th</sup> paragraph>

20. From above, you can say **students** should work part-time.
21. I show merit and demerit of part-time job and compare merit with demerit.
22. You can conclude merit outweighs demerit.
23. Part-time job is also study of a kind.
24. Task of students is to **study** and to get a lot of **knowledge**.
25. Part-time job meet criteria **Students** should work part-time so that you spend better school life.

Note: Line Type: Type1 —; Type2 —>; Type3 —·—; Type4 - - - -; Type5 ······

Words in bold type, underlined words, words in a box,       are either theme- or argument-setting keywords.