

Japanese EFL Learners' Development of Grammatical Competence and its Influence on Critical Writing

TSUTADA, Kazumi
Kansai Gaidai College

Abstract

In the age of globalization at home and abroad, it is important and meaningful to reconsider English education in Japanese universities so that students can advance as social citizens in a global society. In considering the local appropriateness of English education for native Japanese speaking students, this study provided university students with grammar learning opportunities for one semester and then investigated the effects of this training on the development of grammatical competence and the influence of their enhanced competence on criticality in writing. The results demonstrated that grammatical competence increased after grammatical training, despite varied results according to the specific measures, and the enhanced grammatical competence had positive effects on criticality in writing. The study thus concluded by arguing for a certain effect of grammar learning and suggested the importance of grammatical competence for practical writing, which has increased in demand due to accelerating electronic written communication in an increasingly globalized society.

Keywords: grammatical competence, grammatical training, critical writing, criticality

Introduction

Among all the goals and objectives of English education at universities in Japan, one of the principal purposes may be to equip students with adequate practical English communication skills, appropriate to their future courses in their respective professional fields. Amidst a steep rise in demand for writing skills in society due to the rapid prevalence of email communication, the need to develop English writing skills has naturally come to the fore. The present study, based on arguments for the importance of grammar for Japanese EFL (English as Foreign Language) English learners, posits the following two research questions:

- 1) Does the grammatical training exert positive effects on the students' development of grammatical competence?
- 2) Does criticality in writing enhance in parallel with the development of grammatical competence?

Previous Studies

Grammatical Competence

The term “competence” was first introduced by Chomsky (1967) to refer to “knowledge of one's language” (p. 4). However, many linguists objected to Chomsky's definition of

competence (Campbell & Wales, 1970; Hymes, 1972), and instead developed the notion of communicative competence, a term first coined by Dell Hymes, and defined not only as grammatical competence but also the ability to use that grammatical competence in real communication (Hymes, 1972). Canale and Swain (1980) and Canale (1983) followed this up by introducing their theoretical framework of communicative competence, which comprises four skills: grammatical competence (mastery of linguistic code), sociolinguistic competence (knowledge of the sociocultural rules of use and rules of discourse), strategic competence (verbal and non-verbal communication strategies), and discourse competence (ability to connect sentences in stretches of discourse). Despite their insufficient articulation of relationship between grammatical competence and other competencies, Canale and Swain (1980) contended that grammatical competence including the ability to use grammatical knowledge is an important concern for any communicative approach. Subsequently, Bachman (1990) and Bachman and Palmer (1996) proposed two general components of communicative language ability: organizational knowledge and pragmatic knowledge. The former is further categorized into grammatical knowledge and textual knowledge, in which grammatical knowledge refers to grammar on the subsentential and sentential levels, while textual knowledge refers to grammar on the discourse level.

Grounded in the fact that grammatical competence is deeply interconnected with other competencies in language use, and that the boundaries among these competencies remain blurred, this study, in accordance with Canale and Swain's (1980) claim, treats grammatical competence as knowledge of linguistic codes independent of their pragmatic aspects, and the ability to use this knowledge as linguistic infrastructure in actual communication. Furthermore, in accordance with the categorization of grammatical knowledge by Bachman (1990) and Bachman and Palmer (1996), this study focused on grammar at the sentential level when referring to grammatical competence.

Grammar and Writing

Nowadays, many researchers and teachers have become increasingly inclined towards the use of computer-based corpora to help learners improve their L2 writing. Such corpora provide them with a collection of authentic instances of language use, including phrases and sentences. Despite this growing tendency, there remain a number of other factors that are fundamentally important for second language writing, among which is certainly their level of linguistic proficiency in the second language. In particular, many scholars claim that understanding "grammatical structures" might help learners enhance the quality and sophistication of their writing (Celce-Murcia, 2002; Hinkel, 2002; Nation, 2009; Weigle, 2002). For instance, Weigle (2002) contends that "because of the constraints of limited second-language knowledge, writing in a second-language may be hampered because of the need to focus on language rather than content" (p. 35). She adds that inappropriate choices, and lack of time to search for the correct syntactic choices may hinder generation of the written product in accordance with the writer's intention. Many scholars have also claimed that writers' "syntactic variety" is an important construct in writing. For instance, Crowhurst (1980) and Gebhard (1978) claim that the use of more varied sentence structures in essays, including complex sentences, could generally contribute to better text quality. Crowhurst

(1980) also found a positive relationship between syntactic complexity and text quality. From the same perspective, Beers and Nagy (2009) claim that “good writing requires, among many other things, the effective use of syntactic structure.” More recently, Ortega (2015) contended that “syntactic complexity ... is the sophistication, variety, diversity, or elaboratedness of grammatical resources exhibited in language production” (p. 86). Furthermore, syntactic variety has been recognized as a valuable indicator of writers’ developmental paths (Ortega, 2003; Vyatkina, 2012).

The place of grammar is of particular relevance for writing in academic contexts (Celce-Murcia, 1991; Hinkel, 2011, 2013; Weigle, 2002), where writing is frequently seen as a key to entering the social discourse community. Celce-Murcia (1991) emphasizes that grammar is a crucial factor for learners to attain their educational and professional goals in writing. Along the same line, Hinkel (2013) presents a list of grammar constructions to be learned for academic writing, topped by “sentence construction.” Emphasizing the importance of appropriate use of syntactic knowledge in academic writing, Hinkel claims that “English has a rigid word order...The simplest approach to teaching the basic sentence structure can take advantage of the relative rigidity in English sentence structure” (p. 7). Meanwhile, in the business context, Beason’s survey (2001) reported that business people feel that careless writers will generally be careless in conducting business. Beason claims that form should not be considered an added value to writing, but is, rather, an integral part of communication.

In addition, Weigle (2002) argues for the need to take into account the linguistic difference between the native language and the target language. Based on Rutherford’s categories, Elder and Davies (1998) underlined the syntactical differences between English and Japanese in three categories: subject/verb/object, subject-prominence, and grammatical word-order. They further presented the linguistic distance scale for Japanese as rank 5, which connotes the furthest linguistic distance from English. Given these results of typological classification, it may well be considered that one of the most crucial components in attaining an adequate level of English proficiency, particularly for Japanese EFL learners, is a focus on syntactic grammar; in other words, the established syntactic rules of English language. It may seem that having systematic grammatical knowledge can facilitate students’ writing of clear and appropriate sentences, while saving them a great deal of time in trying to discover the rules by themselves.

Heeding the findings of Japanese researchers, Muranoi (2006) proposed the cognition process from input to output, that is, through grammatical noticing, comprehension, intake, and integration. This proposal contains inductive grammar teaching and explicit grammatical explanation as a foundation for output activities. Chujo, Nishigaki, Nishibori, & Oghigian et al. (2008) contend that such teaching strategies work effectively in an EFL environment where the exposure to the foreign language is extremely limited, and strongly support the transfer of grammatical knowledge into output activities. Considering the great language distance between English and Japanese, their assertion is fairly applicable to add to significance of grammar teaching for English writing by Japanese students.

In line with all of the above assertions, the current study attaches a pedagogical value to laying a foundation of grammatical competence, with a focus on syntactical competence, for writing. It also explores the effect of grammatical competence on criticality in writing.

Critical Writing

In the conventional literature, “critical writing” has not been generally acknowledged as accepted pedagogical terminology. However, this study uses the term “critical writing” to denote writing endorsed by critical thinking, which is an important educational goal in higher education for fostering social citizens who can act independently with individual thought and actions (Barnett, 1997). It is also a component that is greatly required to fulfill the needs of an increasingly globalized world, and is “a fundamental aspect of ab initio language teaching” (Houghton & Yamada, 2012, pp. 3 - 4).

To understand critical writing, Canagarajah (2002) claims that “the label ‘critical’...develops an attitude and a perspective that enables us to see some of the hidden components of text construction and the subtle ramifications of writing” (p. 1), and also explains the term “critical” as an antonym of “detached,” “disinterested,” or “abstract.” Meanwhile, Wallace and Wray (2011) argue that “self-critical writing” is underpinned by writers’ own arguments strongly and clearly so as to effectively communicate with and convince their target readers.

According to these definitions and specification of its features, the present article uses the term “critical writing” to encompass the importance of being critical, or criticality in writing to mean writing one’s own ideas and claims subjectively and critically on given information or circumstances. This seems meaningful as criticalness has not been among the common scales for writing assessment, even at universities in Japan.

Assessment Measures

The three measurement dimensions that have generally been identified, i.e., complexity, accuracy, and fluency (CAF), have long been controversial in relation to their respective definition as constructs. They have been evaluated by various measures including holistic and subjective measures, and objective quantitative measures (Wolf-Quintero, Inagaki & Kim, 1998). However, because this current research aimed to evaluate grammatical competence on specific syntactic grammatical items, it was considered most appropriate to employ specific quantitative measures targeting the relevant grammatical items for the assessment of complexity and accuracy.

Meanwhile, fluency concerns automatic language use with limited conscious attention, which is generally regarded as one of the principal ultimate goals of language learning, possibly making it another effective measure. However, its usability as a writing assessment indicator for this study is questionable for the following reasons. First, fluency is generally seen as “mainly a phonological phenomenon” (Housen, Kuiken, & Vedder, 2012, p. 5). Second, in many studies, writing fluency has been measured based on product-based indicators, which are often problematic when word redundancy, level of background knowledge, and emotional factors are taken into account (Bonzo, 2008). Due to these ambiguities and concerns, the current research determined not to include a fluency measure as a writing assessment indicator.

When it comes to a measure for criticality, nearly no such terminology has been commonly introduced as a scale for communicative skills including writing. Although a similar term “argumentativeness” has often been employed to measure how one is argumentative in

communication, its principal concept is to measure a personality trait and mostly in verbal communication rather than written products (Infante, Rancer & Wigley, 2011). Accordingly, this study uses “criticality” as a term for and an important construct of the writing assessment measure to evaluate how much a writer demonstrates his/her critical views on the prompts provided.

Method

Research Design

This study implemented a 10-week grammatical training, using news English, and examined the longitudinal development of grammatical competence, and the effects of grammatical competence on criticality in writing. (See Figure 1.)

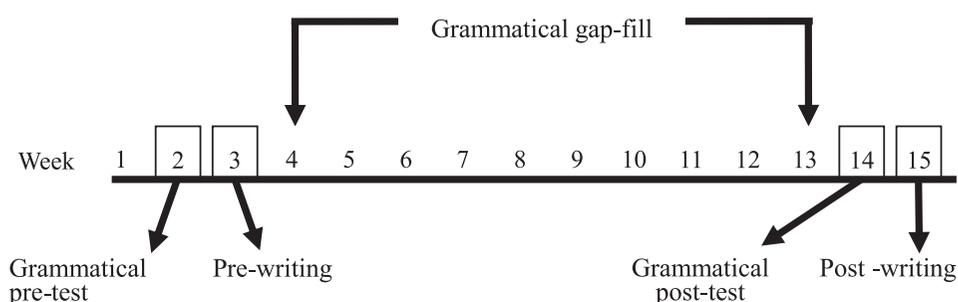


Figure 1. Design of the schedule.

Participants

This research was originally attempted with 188 students from two universities. Of these students, one multivariate outlier who was distinctly different in all of the observed variables from the rest of the population was removed from the sample, rather than transforming the relevant data. This was because, as Kline (2011) claims, it is difficult to obtain such data that works with the rest of the scores by converting it based on mathematical operation; moreover, if the outlier’s scores are severely non-normal, transformations basically do not work. In addition, with regard to missing data, listwise rather than deletion was adopted; thus, cases with missing scores on any variable were excluded from all analyses. This was because listwise deletion is generally recommended unless the number of missing data is small. Thus, in this study, the data of a further 27 students were excluded to give a final number of 160 participants including 153 first- and second-year students, and 7 students in their third and fourth years. Participants’ majors varied widely across both humanities and science departments, as did their level of English proficiency with TOEIC® scores ranging from 410 to 840. All of the participating students granted permission to use the data collected for the research.

Grammatical Training

The grammatical training was executed using gap-fill exercises, which were executed in one class per week for a period of 10 weeks along with other class activities including

vocabulary learning and reading. The most recent online news articles from NHK World's website were used as grammar training material due to their various distinct features, including their authenticity and appropriate levels of readability for the participants. This enabled students to be exposed to real, genuine English, a degree of up-to-datedness that is difficult for textbooks to replicate, and also facilitated the possible development of critical viewpoints on various global issues. The current author made weekly gap-fill exercises (each of which consisted of ten questions) from the news articles, with a focus on the specific syntactical structure of a sentence: the subject-verb (SV) construction, dependent clauses, and post-modifiers (participles, relative clauses, and infinitives).

First, with regard to SV construction, the English language distinctively differs from Japanese in subject/verb/object, subject-prominence, and grammatical word-order (Elder and Davies, 1998), requiring careful attention toward the use of proper verb forms, word order, etc. However, it is fairly common that writers become lost in forming correct SV construction, especially when a subject and verb are far apart in a relatively long sentence. When it comes to the use of dependent clauses, what has been evident in their writing is that many are unable to properly form dependent clauses due to a lack of syntactical knowledge regarding the use of a subject and verb after a dependent conjunction. In addition, frequent mistakes were seen in the past where writers placed a dependent conjunction at the beginning of a main clause, making it difficult for a reader to understand the idea or opinion the writer intends to communicate. The third item, post-modification, is one of the most conspicuous differences from Japanese grammar, making it a crucial factor for Japanese students to learn. For those reasons, conscious attention towards the correct use of these three syntactical elements seems to be of vital importance.

Due to the short duration of this study, the need to acquire relevant knowledge and transfer it into productive use, and the general proficiency level of the participants, it was considered more effective to restrict the number of items on which to concentrate. Thus, the grammatical items selected for testing accorded with those to be linguistically assessed in writing. The entire training was conducted with the intention of encouraging explicit understanding of each grammatical item through intra-group and inter-group discussion, and interaction between students and the teacher about the answers and the reasons for them. The following are two examples of grammatical training materials, edited by the current author. (Only those sentences edited as questions are excerpted from each article, with answers marked by circles.)

Example 1: US Congressman urges Obama to visit Hiroshima

A US Congressman has urged President Barack Obama to visit the atomic-bombed city of Hiroshima when he travels to Japan for the Group of Seven summit next month.

...

Takano encouraged Obama to become the first US president to visit Hiroshima. He said that a visit to Hiroshima is not an apology, but a signal that the leader of the world's largest arsenal of nuclear weapons ___ their power.

(A) appreciation (B) appreciating (C) appreciate (D) appreciates

...

The view that the bombing was necessary to bring the war to a swift end ___ deeply rooted among many Americans.

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

Critics, including veterans, say a presidential visit would represent an apology by the United States for dropping the bomb. Major US newspapers such as the New York Times and the Washington Post have printed editorials ___ the idea of a visit by the US leader.

(A) support (B) supporting (C) supports (D) supported

...

(“US Congressman urges,” 2016)

Example 2: California’s last nuclear power plant to shut down

The operator of California’s last active nuclear power plant ___ it down by 2025.

(A) shut it down (B) to shut it down (C) will shut it down (D) shutting it down

...

PG&E says it decided to decommission the reactors ___ requirements in a joint proposal with labor and environmental groups.

(A) to meet (B) met (C) meet (D) meets

...

Local residents had been demanding their shutdown ___ the accident at Japan’s Fukushima Daiichi nuclear plant in March 2011.

(A) as (B) despite (C) since (D) because

...

In recent years an increasing number of utilities in the US have shut down their nuclear plants, or have decided to, ___ economic and other reasons. The US has more commercial reactors than any other country.

(A) cite (B) cites (C) which site (D) citing

...

(“California’s last nuclear,” 2016)

Grammatical Tests

The grammatical pre- and post-tests aimed to measure the level of grammatical knowledge before and after the grammatical gap-fill training. Those items tested here were the same as those learned through grammatical gap-fill exercises (SV structure, dependent clauses, and post-modifiers). The same tests were used on both occasions, as it seemed impossible from a practical perspective to make two different tests that would be completely identical in terms of difficulty level. In addition, the time between the tests (12 weeks) seemed sufficient to counter any memory effect of the pre-test, although it is practically impossible to account for this. Three different tests were formulated (with reference to *English Grammar in Use* and *Advanced Grammar in Use*): Test 1. Multiple choice to choose one correct word or phrase (30 questions); Test 2. Multiple choice to select a correct sentence (20 questions); and Test 3. Error correction to indicate incorrect usages and correct them (30 questions). The tests were examined through preliminary pilot tests, and their results were analyzed using item analysis (Classical Test Theory). Necessary modifications were then made to improve the level of test

usefulness. These tests were administered in accordance with the schedule (Figure 1) under timed conditions (15 minutes for Test 1, 10 minutes for Test 2, and 15 minutes for Test 3). Example questions of Test 1, Test 2, and Test 3 are shown below with answers.

To examine the extent of development of grammatical knowledge after the grammatical training, the results of these three sets of pre- and post-tests were analyzed by multiple comparisons.

Excerpts from Grammatical Test 1 (Choose the correct word)

1. The product was not a commercial success ___ a lot of money was spent on advertising.
(A) despite (B) in spite of (C) even though (D) because of
2. Due to his father's bankruptcy, the funds ___ for entry into the expensive university were unavailable.
(A) requirement (B) require (C) requiring (D) required
3. Ms. Parker was very impressed with the answers the job applicant ___ during the interview.
(A) giving (B) to give (C) gave (D) give
4. When you buy a house, be sure to use a real estate agent ___ knowledge of the local market is comprehensive.
(A) who (B) whose (C) that (D) which
5. My cousin's amazing advice about selling our stocks ___ us thousands of dollars.
(A) saving (B) saved (C) to save (D) save
6. Consumers are spending less these days ___ reports that the economy is steadily improving.
(A) in spite (B) because (C) although (D) despite
7. All new staff members should become familiar with the standard office procedures ___ in the new employee manual.
(A) write (B) written (C) writing (D) to write
8. In spite of several years of experience in management, Pedro did not get the position ___ he applied.
(A) that (B) which (C) for which (D) whose

Excerpts from Grammatical Test 2 (Choose the correct sentence)

1. (A) Some paintings being stolen from the museum haven't been found yet.
(B) Some paintings stealing from the museum haven't been found yet.
(C) Some paintings stolen from the museum haven't been found yet.
2. (A) The car you rented broke down after a few miles.
(B) The car you rented and broke down after a few miles.
(C) The car rented and broke down after a few miles.
3. (A) During we climbed the hill, we got more and more tired.
(B) As we climbed the hill, we got more and more tired.
(C) Because of we climbed the hill, we got more and more tired.

4. (A) I recently went back to the small town where I grew up.
(B) I recently went back to where I grew up in the small town.
(C) I recently went back to the small town which I grew up.
5. (A) The college offers English intensive course in the evening beginning next Monday.
(B) The college offers English intensive course in the evening begins next Monday.
(C) The college offers English intensive course in the evening will begin next Monday.

Excerpts from Grammatical Test 3 (Correct an error)

1. The direct flight from Tokyo to New York has been canceled because (→because of) the local weather forecast that a devastating hurricane is approaching.
2. The attaching (→attached) file contains more information on the new product.
3. Drivers have to make a detour during (→while) the main street is under construction.
4. According to the classified ad in today's newspaper, one of the best apparel companies have (→has) some openings in the sales section now.
5. However (→Although) he has a very important job, he isn't particularly well-paid.

Writing

Writing was assessed from two aspects. Regarding “grammatical” elements, participants’ writing proficiency was evaluated in respect of the same grammatical components as those learned through gap-fill grammatical training. Meanwhile, as a subjective measure, “criticality” was evaluated encompassing cohesion and organization. In fact, although the assessment was also made on comprehensibility along with criticality for another independent analysis, the current approach focuses on criticality in accordance with the objective of the study. Consequently, the construct of the critical writing test was to examine the development in linguistic proficiency in writing with regard to the relevant grammatical features, and criticality.

Writing conditions. The writing test was executed in a 30-minute timed impromptu manner. Despite the difficulty of choosing prompts that would be equally accessible to all participants, “general” rather than “personal” topics were selected to allow the participants to construct and develop their arguments in critical writing. Specifically, the prompts were chosen from the following news stories: “Paternity leave” for the pre-writing task and “Right to vote for the 18-year-old” for the post-writing task. Both topics were socially controversial on the respective occasions.

Regarding dictionary use, to reflect natural, authentic writing environments, participants were allowed free access to dictionaries. Furthermore, as East (2006) and Bruton (2007) argue, allowing dictionary use helps reduce the psychological burden on writers, enhances their motivation, and encourages writer autonomy as they have rich resources at hand. Regarding the writing format, the essays were handwritten on one or both sides of A4 paper.

Assessment measures. The writing tasks were assessed using linguistic measures (complexity and accuracy), and a criticality measure.

The linguistic measures were determined so that the relevant grammatical knowledge

learned through gap-fill exercises could be generally tested as productive knowledge. Complexity was first evaluated using two measures: one to measure the relative ratio of correctly applied post modifiers used in an essay (post-modifying relative clauses, participles, and infinitives) (coded as C1); the other to address the relative ratio of properly used dependent conjunctions in an essay (coded as C2). Accuracy was then evaluated using one measure, to indicate the extent to which the SV construction was correctly formed in each clause, and represented by the relative ratio of correct clauses (on the use of SV structure) (coded as A).

With regard to a criticality measure (coded as CR), as it is a subjective measure, the evaluation was conducted by independent native English speaking raters and based on a rubric originally formulated by the present author with reference to the Educational Testing Service's (ETS) 6-point Test of Written English (TWE®) rating scale, which is the essay component of the TOEFL® and a well-known holistic scoring rubric. (See Appendix for the criticality rubric). In addition, in order to orchestrate necessary factors to make a reasonably tangible and accessible criticality rubric for actual evaluation it was supplemented at the author's discretion, with assessment criteria based on the conclusions. The rating process started with the selection of the anchor essays to be used as general yardsticks in the actual evaluation. These were selected from another writing pool collected from a group of students at another university, who wrote on the same prompt under identical conditions. Once all of the writing samples were typed by the author, the raters were asked to assemble for orientation and rater training, which Weigle (1994) claims enhances the reliability of the raters. One thing to note is that the raters were advised to engage in the assignment independently so as not to be influenced by other raters' scores. What was fortunate was that the second set of essays (post-writing samples) were able to be assessed by the same six raters as the first set (pre-writing samples).

The results of the pre- and post-writing tests were analyzed using multiple comparisons for three measures - complexity 1 (C1), complexity 2 (C2), and accuracy (A), and a paired *t*-test for criticality (CR) - to examine the significant difference between the two tests.

Results and Discussions

Grammatical Competence (G1, G2, G3, C1, C2 and A)

First, the gap-fill scores for weekly grammatical training showed a general increase, with all of them including only a couple of minor drops, indicating that the students' skills at solving the grammatical questions embedded in news articles increased steadily over the 10 weeks. Particularly, after the sixth week, the scores rose rather noticeably. This may be due to their accumulation of grammatical knowledge through grammatical discussions and explicit instruction by the teacher. Mean values of weekly results with standard deviations are shown in Table 1.

Table 1

Weekly Gap-fill Scores (N = 160)

Week	1	2	3	4	5	6	7	8	9	10
<i>Mean</i>	3.02	3.08	3.32	3.76	4.69	4.44	5.44	6.10	7.14	7.19
<i>SD</i>	0.94	0.84	0.98	0.83	0.68	0.75	0.75	0.91	1.04	0.94

Regarding the results of the six indicators of grammatical competence (i.e., the scores of the three grammatical tests, and the three assessment measures for writing), Table 2 shows the descriptive statistics of the pre- and post-results. In the table, skewness and kurtosis, all within ± 2 , indicate the normal distribution of the statistics. The results of multiple comparisons using paired *t*-tests are shown in Table 3.

As shown in Table 3, the results of all three types of grammatical tests (G1, G2, and G3) showed a significant difference, with medium effect sizes, which indicated some improvement in grammatical knowledge in the post-test. The ability to use English (writing complexity and accuracy) revealed a mixed outcome among the three assessment measures. First, the two complexity measures (C1 and C2) showed a significant difference with large effect sizes. While it might not be appropriate to straightforwardly judge that higher complexity corresponds to more advanced writers, based on the claims of previous research regarding the relationship between syntactic variety and text quality (Crowhurst, 1980; Gebhard, 1978; Ortega, 2003, 2015; Vyatkina, 2012), it was considered an important phase of the students' developmental path that they be equipped with some ability to incorporate post-modification such as relative clauses, participles, infinitives, and subordinate conjunctions (all focused on in this study) in their English production. As for the accuracy measure, errors were counted based on the definition of accuracy formulated in this study, i.e., the correct formation of the SV structure in a clause. Table 3 indicates the significant difference between the pre- and post-writing samples with a small effect size. The lower improvement in accuracy than in complexity may be partly explained by the Trade-off Hypothesis. As Larsen-Freeman (2009) claims, "there is a competition for resources, which leads to performance which is either more complex or more accurate, but not both" (p. 583), and it might have been too demanding for students to simultaneously increase their proficiency in all assessment factors over the limited study period.

Table 2

Descriptive Statistics of Pre- and Post-tests by Measure (N =160)

		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>Skewness</i>	<i>Kurtosis</i>
G1	Pre	160	18	4.17	7	27	-0.18	-0.41
	Post	160	20.74	4.11	12	29	-0.17	-0.89
G2	Pre	160	11.63	3.40	3	20	0	-0.27
	Post	160	13.86	2.89	8	20	0	-0.80
G3	Pre	160	12.26	5.87	1	27	0.27	-0.57
	Post	160	15.29	5.20	5	26	0.21	-0.76
C1	Pre	160	1.59	0.73	0	3.92	0.55	0.14
	Post	160	3.75	1.72	0	8.65	0.49	-0.38
C2	Pre	160	1.59	0.73	0	3.91	0.55	0.14
	Post	160	2.22	0.79	0	4.60	0.57	0.35
A	Pre	160	9.45	2.55	4.44	15.54	0.40	-0.54
	Post	160	10.2	2.55	4.50	16.18	0.42	-0.36

Table 3

Results of Multiple Comparisons between Pre- and Post-tests by Measures (N =160)

	<i>Mean difference</i>	<i>t</i>	<i>df</i>	<i>p</i> ^a	<i>d</i> ^b
G1	2.74	22.44	159	< .001	0.66
G2	2.23	19.88	159	< .001	0.67
G3	3.03	14.70	159	< .001	0.53
C1	2.16	22.36	159	< .001	1.13
C2	0.62	8.72	159	< .001	0.82
A	0.74	3.55	159	< .001	0.29

Note. ^a α level was adjusted to .008 by Bonferroni correction. ^b*d* (Cohen's *d*).

Overall, all of the results of linguistic measures showed significant differences with varied effect sizes before and after the grammatical training. This meant that both grammatical knowledge and the ability to use that knowledge, defined as grammatical competence in the current study, were improved after the grammatical exercises. This answers the first question posted in this study: does the grammatical training exert positive effects on the students' development of grammatical competence?

Critical Writing Ability (CR)

As to criticality (CR), the two raters' scores per essay were simply averaged to provide a single data value for analysis. However, those essays, with a large discrepancy of more than two points between raters, were assessed by a third rater, also a native English speaker, and the mean of the three raters' assessments was then used as the value for analysis. Table 4 shows the descriptive statistics of the pre- and post-tests, in which the Cronbach's coefficient alpha was measured before the third rater's evaluation was included when such evaluations were necessary. The results of a paired *t*-test are shown in Table 5.

These results show that criticality, a central focus of the current research, was moderately

improved between the pre- and post-writing samples. One of the reasons for this rise in criticality might have been the participants' enhanced skill in building properly structured complex sentences, thus making it easier for them to convey their thoughts. This corroborates the arguments that there is a relationship between complexity and text quality (Beers & Nagy, 2009; Crowhurst, 1980; Gebhard, 1978; Ortega, 2003, 2015; Vyatkina, 2012). In this sense, it seems reasonable that some of the development of criticality paralleled the enhancement of grammatical competence. Meanwhile, from the raters' perspectives, the grammatically improved essays might have helped them concentrate on reading and evaluating those essays, which may in turn have increased their evaluation of criticality. In other words, for those essays with many grammatical errors, the "criticality" scale might not have been appropriate, as indicated by the relatively low reliability coefficient throughout the entire evaluation.

Nonetheless, the critical writing ability saw a significant difference with a medium effect size of .50 after grammatical training, which enhanced grammatical competence as indicated in the previous section.

Table 4

Descriptive Statistics of Pre- and Post-tests (Criticality) (N = 160)

		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>	<i>Skewness</i>	<i>Kurtosis</i>	α^a
CR	Pre	160	3.36	0.69	2	5	-0.19	-0.56	.72
	Post	160	3.71	0.72	2.5	5	0.20	-0.88	.79

Note. ^a α : Cronbach's coefficient alpha

Table 5

Results of Paired t-test between Pre- and Post-tests (Criticality) (N = 160)

	<i>Mean difference</i>	<i>t</i>	<i>df</i>	<i>p</i>	d^a
CR	0.35	8.52	159	< .001	0.50

Note. ^a*d* (Cohen's *d*).

In order to exemplify the development of criticality assessment, the writing examples of two students (with Students A and B whose criticality evaluation rose from 2.5 to 4.0, and 3.0 to 4.5 respectively) are shown below

Example of Student A (Pre-writing with criticality of 2.5)

I agree with paternity leave. I think (1) only mother does childcare is wrong, and (2) father works outside should care about childcare. Babies are children of father and mother.

(3) Because childcare is important for both of them. (4) Why only mother does childcare? I don't understand. All mothers must prepare meals, wash clothes, and go shopping. They have a lot of stress and may become a disease. (5) If paternity leave, their stress will be able to be decreasing. (6) The couple share the duties of parenting is best way. So I agree with paternity leave.

Example of Student A (Post-writing with criticality of 4.0)

I oppose the right to vote from 18 years old. (7) Because the greater part of 18 years old are students, they don't have knowledge of election (8) required. (9) If the government adopts the right to vote for the 18-year-old, many students learn about the election and (10) candidates to choose and so on from teachers. This means (11) that many students are influenced by teachers. In other cases, they decide the same candidates with friends, or famous (12) people they know very well. So the 18-year old will not be able to change society. That's (13) why I oppose the right to vote for the 18-year-old.

In the pre-writing above, it is apparent that the writer made mistakes in SV construction in sentences (1), (4), and (6), use of participle as in (2), and use of a dependent conjunction as in (3) and (5). Overall, it is notable that the entire passage mostly consists of simple instead of complex sentences. Meanwhile, in the post-writing, the writer successfully used dependent clauses in (7), (9), (11), and (13), participles as post-modifiers in (8) and (10), despite his unnatural use of "required". He also utilized an infinitive in (10), and a relative clause in (12) (deleted though).

Example of Student B (Pre-writing with criticality of 3.0)

I agree with paternity leave. In Japan, men must work every day. Now, many women work, too. I think (1) childcare not only women's work. Also men's childcare is good for children. For example, my uncle have two children. He didn't take paternity leave. So he didn't contact much with his children. (2) I think important to contact children since they are very small. Many Japanese companies give paternity leave only few days. But it is senseless. Japanese companies should give more paternity leave. They should change the companies.

Example of Student B (Post-writing with criticality of 4.5)

I have some reasons (3) to agree with the right to vote for the 18-year-old. First, young (4) people who are over 18 are interested in politics. They read newspaper, watch news at TV every day, and ask teacher questions about politics. This is (5) why I agree (6) that they can join politics. Second, (7) because voting rate is occupied by old people now, Japanese politics doesn't change. (8) If young people votes, a lot of new opinions would be added, and Japan might change. Finally, (9) if young people were interested in politics, adult would be interested in politics, too. For these reasons. I think (10) that young people's voting is a good measure.

In this pre-writing sample, as well as the first example, most of the sentences were short and simple with nearly no utilization of post-modifiers. Attempts to use dependent "that" clauses were seen in (1) and (2), but were not successful. In the post-writing sample, appropriate use of dependent clauses was seen in (5), (6), (7), (8), (9), and (10), and post-modifiers in (3) with an infinitive, and (4) with a relative clause.

As the criticality assessment was based on subjective judgement, even when utilizing a

rubric, it largely depends upon how strict the raters are about grammatical usage, word choice, and other errors. However, the results showed a certain degree of overall development of criticality in their writing samples after grammatical training undertaken to enhance grammatical competence. This answers our second research question: does criticality in writing enhance in parallel with the development of grammatical competence?

Conclusions

First, as for the longitudinal development of grammatical competence after grammatical training, grammatical knowledge (indicated by scores on three grammatical tests) showed a significant difference with medium effect sizes. Regarding the ability to apply grammatical knowledge in writing, the results demonstrated significant difference with large effect sizes for complexity, and a small effect size for accuracy. These results seem to underpin the positive effects of grammatical training on students' grammatical competence through integrating both their grammatical knowledge and the ability to use it for output, although it is difficult to rule out the influence of other grammar learning stimuli.

Meanwhile, for the development of criticality, the results showed a significant difference with a medium-sized effect before and after the grammatical training. This suggests that improved grammatical competence in their writing samples may possibly enable them to be more critical from the readers' subjective perspectives, as is underscored by the arguments supporting a relationship between linguistic factors and levels of text quality (Crowhurst, 1980; Gebhard, 1978; Ortega, 2003, 2015; Vyatkina, 2012).

Despite those conclusions, several limitations were identified with regard to the methods and procedures of the writing assessments. One concerns the effects on all the statistical data of other class activities than the grammatical training and writing applied for this study. It is undeniable that language learning occurs through various media and instruction. It also remains a concern whether the prompts were equally fair and attentive to all participants. The next limitation relates to the assignment scheme of the writing prompts. In this study, all students wrote on the same title at the same time due to the author's intention to assign writing on the most recent topic. However, from the perspectives of fair assessment, the counter-balanced scheme might have been a better option to generate more dependable and practical outcomes. Regarding the use of dictionaries while writing, the unexpectedly varied level of proficiency in dictionary use may have obscured genuine evaluation of the participants' writing ability itself. As writing proficiency and dictionary use skill seem interdependent, it would be worth reconsidering and regulating the rules of dictionary use. Lastly, regarding the writing format, participants' claims that the use of handwriting was inconvenient seems justified, as they had difficulties rewriting sentences and paragraphs, and this may have hindered their willingness to improve their essays. Reconsideration of all these aspects might help ensure a more dependable, thus generating a more reliable outcome that is useful for designing classroom activities in the future.

Notwithstanding its limitations, regarding the supposition that the ability to use critical writing to express one's views has become increasingly important in society, the results of this study might serve to reconfirm that grammar is a necessary infrastructure for practical writing. This might contribute to the reconsideration of English education in Japan, which has

increasingly focused on communicative approaches rather than grammar learning.

As a final remark, the study might prompt the renovation, even if limited, of English education at Japanese universities, based on the renewed recognition of the importance of grammar to allow Japanese EFL learners to efficiently engage in written communication in society. It is hoped that more university students can acquire practical English writing proficiency reinforced with grammatical competence, which would help them advance as members of the social community with pride and confidence.

Acknowledgements

I would like to thank the three anonymous reviewers for their helpful comments and suggestions in revising this article.

References

- Bachman, L. F. (1990). Constructing measures and measuring constructs. In B. Harley, P. Allen, J. Cummins and M. Swain (Eds.), *The Development of second language proficiency* (pp. 26–38). Cambridge: Cambridge University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Barnett, R. (1997). *Higher education: A critical business*. Buckingham: Open University Press.
- Beason, L. (2001). Ethos and error: How business people react to errors. *College Composition and Communication*, 53(1), 33–64.
- Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre?. *Reading and Writing*, 22(2), 185–200.
- Bonzo, J. D. (2008). To assign a topic or not: Observing fluency and complexity in intermediate foreign language writing. *Foreign Language Annuals*, 41(4), 722–735.
- Bruton, A. (2007). Vocabulary learning from dictionary referencing and language feedback in EFL translational writing. *Language Teaching Research*, 11(4), 413–431.
- California's last nuclear power plant to shut down. (2016, July 12). Retrieved from <https://www3.nhk.or.jp/nhkworld/en/news/>
- Campbell, R., & Wales, R. (1970). The study of language acquisition. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 242–260). Harmondsworth: Penguin Books.
- Canagarajah, A. S. (2002). *Critical academic writing and multilingual students*. Ann Arbor: University of Michigan Press.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards, J. C. and R. W. Schmidt (Eds.), *Language and communication* (pp. 14–40). New York: Routledge.
- Celce-Murcia, M. (1991). Grammar pedagogy in second and foreign language teaching. *TESOL Quarterly*, 25(3), 459–480.
- Celce-Murcia, M. (2002). Why it makes sense to teach grammar in context and through discourse. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 119–134). Mahwah, NJ: Lawrence Erlbaum.

- Chomsky, N. (1967). *Aspects of the theory of syntax* (4th ed.). Cambridge: The M.I.T. Press.
- Chujo, K., Nishigaki, C., Nishibori, A., and Oghigian, K. (2008). An Effective Way to Teach Grammar to EFL Beginners Using Data-Driven Learning. *College of Industrial Technology, Nihon University Research Report B*, 41, 15–33.
- Crowhurst, M. (1980). Syntactic complexity and teachers' quality ratings of narrations and arguments. *Research in the Teaching of English*, 14(3), 223–231.
- East, M. (2006). The impact of bilingual dictionaries on lexical sophistication and lexical accuracy in tests of L2 writing proficiency: A quantitative analysis. *Assessing Writing*, 11(3), 179–197.
- Elder, C., & Davies, A. (1998). Performance on ESL examinations: Is there a language distance effect?. *Language and Education*, 12(1), 1–17.
- Gebhard, A. O. (1978). Writing quality and syntax: A transformational analysis of three prose samples. *Research in the Teaching of English*, 12(3), 211–231.
- Hewings, M. (2013). *Advanced grammar in use: a self-study reference and practice book for advanced learners of English* (3rd ed.). Cambridge: Cambridge University Press.
- Hinkel, E. (2002). Teaching grammar in writing classes: Tenses and cohesion. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 181–198). Mahwah, NJ: Lawrence Erlbaum.
- Hinkel, E. (2011). What research on second language writing tells us and what it doesn't. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning Vol. 2* (pp. 523–538). New York: Routledge.
- Hinkel, E. (2013). Research findings on teaching grammar for academic writing. *English Teaching*, 68(4). 3–21.
- Houghton, S. & Yamada, E. (2012). *Developing criticality in practice through foreign language education (ISFLL Vol. 8)*. Bern: Peter Lang.
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy, and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA* (pp. 1–20). Amsterdam: John Benjamins.
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Harmondsworth: Penguin Books.
- Infante, D. A., Rancer, A. S. & Wigley, C. J. (2011). In defense of the argumentativeness and verbal aggressiveness scales. *Communication Quarterly*, 59(2), 145–154.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 579–589.
- Muranoi, H. (2006). *SLA Research and Second Language Learning and Teaching*, Tokyo: Taishukan-shoten.
- Murphy, R. (2012). *English grammar in use: a self-study reference and practice book for intermediate learners of English* (4th ed.). Cambridge: Cambridge University Press.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. New York: Routledge.
- Ortega, L., (2003). Syntactic complexity measures and their relationship to L2 proficiency: A

- research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518.
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing*, 29, 82–94.
- US Congressman urges Obama to visit Hiroshima. (2016, April 18). Retrieved from <https://www3.nhk.or.jp/nhkworld/en/news/>
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, 96(4), 576–598.
- Wallace, M., & Wray, A. (2011). *Critical reading and writing for postgraduates* (2nd ed.). London: Sage Publications.
- Weigle, S. C. (1994). Effects of training on raters of ESL compositions. *Language Testing*, 11(2), 197–223.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wolfe-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.

Appendix
Scaling rubrics for criticality assessment

Scale	Description
6	<p>Major viewpoints are addressed clearly and effectively. Clearly appropriate details are used to support an essay or illustrate ideas. The writing is well-organized and well developed. Conclusions are demonstrated clearly and strongly based on the key concepts.</p>
5	<p>Major viewpoints are appropriately addressed. Appropriate details are used to support an essay or illustrate ideas. The writing is generally well organized and developed. Conclusions are demonstrated based on the key concepts.</p>
4	<p>Major viewpoints are generally addressed. Some details are used to support an essay or illustrate ideas. The writing is adequately organized and developed. Conclusions are reasonably demonstrated.</p>
3	<p>Major viewpoints are superficially identified. Inappropriate or insufficient details are used to support or illustrate generalizations. The writing is inadequately organized or developed. Conclusions are acceptable.</p>
2	<p>Major viewpoints are not identified clearly enough. Little or no detail, or irrelevant specifics are used The writing is disorganized or underdeveloped. Conclusions stray from the key concepts.</p>
1	<p>Major viewpoints are not identified at all. No details or specifics are used. The writing is undeveloped. No conclusions are demonstrated.</p>