

Accepted Manuscript

Title: How Might Language Affect Critical Thinking Performance?

Author: Emmanuel Manalo Chris Sheppard

PII: S1871-1871(16)30025-6

DOI: <http://dx.doi.org/doi:10.1016/j.tsc.2016.05.005>

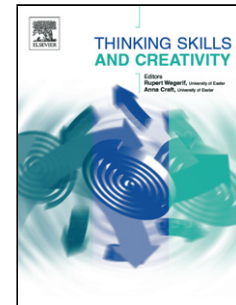
Reference: TSC 353

To appear in: *Thinking Skills and Creativity*

Received date: 27-7-2015

Revised date: 29-3-2016

Accepted date: 11-5-2016



Please cite this article as: Manalo, Emmanuel., & Sheppard, Chris., How Might Language Affect Critical Thinking Performance?. *Thinking Skills and Creativity* <http://dx.doi.org/10.1016/j.tsc.2016.05.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

How Might Language Affect Critical Thinking Performance?

Emmanuel Manalo

Graduate School of Education, Kyoto University, Kyoto, Japan

Chris Sheppard

Center for English Language Education, Faculty of Science and Engineering, Waseda University,
Tokyo, Japan

Address for correspondence: Emmanuel Manalo, Graduate School of Education, Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan.

Telephone: +81-90-6672-2075

Email: manalo.emmanuel.3z@kyoto-u.ac.jp

Highlights

- We examined whether language structure or proficiency affect critical thinking.
- Japanese students show critical evaluation better in Japanese than in English.
- Proficiencies in native- and second-languages correlate with critical evaluation in writing.
- But without critical evaluation instruction, no such correlations were found.
- Results indicate importance of language proficiency and provision of instruction.

Abstract

This study examined whether language structure or language proficiency might affect students' critical thinking performance. Previous research has claimed that many non-Western students struggle with the demands of demonstrating critical thought. Two language-related causes have been suggested: one concerning structural limitations in the non-Western students' first language, and the other concerning their second language proficiency. In Study 1 described here, reports written by 110 Japanese second year university students, who had received instruction in academic discourse for critical evaluation (which is one aspect of critical thinking), were analyzed for use of evaluative statements. No disadvantage was found for use of the Japanese language, which is considered as having a more indirect structure that may make critical evaluation more difficult. Measurements of language proficiency in English and Japanese, however, were found to correlate with production of evaluative statements in those respective languages suggesting that language proficiency could affect critical evaluation use. In Study 2, the same task was given to 43 first year students who had not yet received the same instruction. Analysis revealed similar patterns in their written work but at a lower level, suggesting that the second year students had benefitted from the skills instruction. Furthermore, unlike the second year students, the first year students evidenced no correlations between their language proficiency scores and their production of evaluative statements,

suggesting that proficiency on its own is inadequate: students need instruction on the specific language forms and structures to use to demonstrate critical thinking in their written work.

Keywords: critical thinking skills instruction, critical evaluation, cognitive cost, language proficiency, language structure

1 Introduction

In modern societies, the ability to evaluate the credibility of information that one encounters is an essential skill, especially considering the proliferation of unvetted information through the Internet and other forms of mass media (e.g., Glassner, Weinstock, & Neuman, 2005; Thomm & Bromme, 2011). However, educational development of critical thinking skills, which includes critical evaluation, is not a straightforward matter. There are, for example, disagreements about appropriate methods for critical thinking skills instruction and assessment, and various individual and situational factors are believed to influence the development and use of such skills (e.g., Davies, 2006; Halpern, 1998; Ku, 2009; Manalo, Kusumi, Koyasu, Michita, & Tanaka, 2013, 2015; ten Dam & Volman, 2004). One of the main areas of contention is whether cultural background contributes to differences in critical thinking performance (e.g., Ennis, 1998; Manalo et al., 2013, 2015). Some authors, for example, have portrayed Asian students as being deficient in critical thinking compared to Western students (e.g., Atkinson, 1997; Fox, 1994), and it has been found that many instructors at the tertiary level subscribe to such a view (e.g., Lee & Carrasquillo, 2006; Robertson, Line, Jones, & Thomas, 2000).

Two language-related explanations have been proposed for the apparent differences in critical thinking performance manifested by students from different cultural backgrounds. One of those explanations concerns the structure of the student's native language (L1). This explanation

posits that, due to their structure, some languages may present constraints in the ease with which certain thinking skills can be undertaken or expressed. This explanation is sometimes referred to as the “Sapir-Whorf hypothesis” or the “linguistic relativity hypothesis” (see, e.g., Au, 1983; Hill & Mannheim, 1992; Hockett, 1954; Whorf, 1956). An example of a claim of this kind is Bloom’s (1981) proposal that counterfactual thinking (i.e., thinking about what might have been, contrary to facts) may be more difficult in Chinese compared to English.

Some observations of linguistic differences, such as “indirectness” being a feature more prevalent in some languages, particularly Asian languages (e.g., Kong, 2005), would appear to support the idea that language structure could affect the ease with which certain modes of thinking could be undertaken or expressed. In a study by Itakura and Tsui (2011), for example, evidence was found that book reviewers use different strategies to convey critical evaluation when writing in Japanese compared to English: in Japanese, criticism is usually indirectly conveyed and is frequently preceded by an apology. Previous studies, however, have failed to provide convincing evidence that language structures could actually impose constraints in what users of the language can communicate. Although the earlier-mentioned study by Bloom (1981) claimed to have found evidence for such constraints where counterfactual thinking in the Chinese language is concerned, subsequent investigations failed to replicate or support Bloom’s results (Au, 1983; Liu, 1985; Takano, 1989; Vorster & Schuring, 1989). Even where the weaker version of the linguistic relativity hypothesis is concerned (i.e., that language can merely influence people’s thought patterns, rather than constrain their thought processes), the findings of the few studies that claim to have found supporting evidence (e.g., Boroditsky, 2001) have not been replicated in subsequent independent investigations (e.g., Chen, 2007; January & Kako, 2007; Tse & Altarriba, 2008).

The other language-related explanation for the apparent culture-based differences in critical thinking performance concerns students’ second language proficiency (e.g., Floyd, 2011; Lun, Fischer, & Ward, 2010; Paton, 2005). This explanation is based on the observation that the students

who are usually reported as having lower levels of critical thinking competence are international students in English speaking countries. It suggests that, because most of those students have to use a second language (L2, which is usually English) in their host environment, and they may lack adequate proficiency in that L2, they would likely manifest lower competence when performing tasks that require use of that L2. Tasks that are likely to get affected include cognitive tasks like critical thinking.

One way of understanding this possible influence of language proficiency on critical thinking performance is in terms of cognitive cost (i.e., the mental resources cost associated with task performance). Language processing entails the use of cognitive resources in working memory (Baddeley, 1986, 1998), and lower proficiency in a language would require the use of more resources. Thinking critically would likewise require the use of working memory resources. There are, however, limited resources available in working memory (Baddeley, 1986, 1998) and, if a considerable amount of those resources has already been expended on utilizing a language in which proficiency is low, there may not be adequate resources remaining for the satisfactory execution of critical thinking.

The negative impact of the higher cognitive cost entailed in using a language in which proficiency is low, on the execution of other cognitive tasks, has been demonstrated in previous research. Takano and Noda (1993, 1995) showed that the use of a foreign language detrimentally affects performance in concurrently undertaken non-linguistic tasks like arithmetic calculation and mental imagery, and Manalo and Uesaka (2012, 2014) reported evidence indicating that students' lower proficiency in an L2 limits their ability to use diagrams when explaining information in that L2. Where critical thinking is concerned, both Lun et al. (2010) and Floyd (2011) reported indications that lower proficiency levels in English could detrimentally affect Asian students' performance in critical thinking tests administered in English. However, neither of those studies

used appropriate, objective measures of L2 proficiency to reliably confirm the connection between L2 proficiency and critical thinking skills performance.

1.2 Overview of the Present Study

The present study examined whether there might be evidence to support either or both (i) the language structure explanation, and/or (ii) the language proficiency explanation, in students' manifestation of critical thinking in their written work. The study was not intended to be a comprehensive test of the language structure hypothesis: it examined only whether, in the written work of Japanese university students, there might be observable differences in the presence of critical thinking qualities depending on the language used, Japanese or English. Critical thinking was operationalized as students' use of evaluative statements. This decision was based on the fact that use of evaluative statements comprises a salient expression of critical evaluation, which in turn is central to the notion of critical thinking application (e.g., Fisher & Scriven's, 1997, p. 21, definition of critical thinking as "skilled and active interpretation and *evaluation* of observations and communications, information and argumentation" – italics added).

In the present study, Japanese was deemed an appropriate language to examine because, like a number of other Asian languages, it employs patterns of expression that make it more indirect and inductive compared to English (e.g., Itakura & Tsui, 2011; Scollon & Wong-Scollon, 1991). Evaluation, however, requires precision and directness in conveying judgments about the quality or value of the subject being referred to. Thus, structural features of the Japanese language could make the production of evaluative language relatively more difficult. If so, it should be possible to detect lower rates of evaluative language use in the students' written work in Japanese compared to English.

It was equally important to consider whether using an L2 might detrimentally affect students' critical evaluation performance. If this were the case, one would expect that critical evaluation performance in the L1 would exceed that in the L2. One would also expect to find a relationship

between the critical evaluation performance and L2 proficiency. Thus, possible relationships between students' L2 proficiency and their production of evaluative statements were investigated. The question was whether L2 proficiency would manifest as a limiting factor because lower proficiency entails higher cognitive cost when using the L2, leaving insufficient resources in working memory for critical evaluation. If this explanation is supported, a relationship should be found between the students' L2 proficiency and their evaluative statements production in the L2, *but not in the L1*. A relationship in both the L1 and the L2 would suggest that general language or intellectual abilities – rather than L2 proficiency – affect critical evaluation performance. The reason is that language abilities, and intellectual abilities and performance, are generally considered as being related (e.g., Ackerman, 1986; Neisser et al., 1996). Thus, a student with higher language and intellectual abilities could be expected to score higher in measurements of L2 proficiency, and evidence better performance in tasks like critical evaluation – in both their L1 and L2.

Furthermore, it is conceivable that this kind of relationship between language proficiency and critical evaluation also exists in the L1. In other words, that the more proficient or skilled a student might be in using his or her native language, the less cognitive resources such use might demand, and the more resources might be made available for critical evaluation – which could then be better executed. Thus, in the present study, the possible existence of this relationship in the L1 was also investigated.

When considering the development of critical evaluation skills through instruction, if the language structure explanation holds, one would expect that skills instruction provided in the L2 would *not* result in any discernible skills improvement in the L1 (i.e., since the L1 has inherent structural constraints on the use or expression of those skills). In contrast, if the language proficiency explanation holds, one could expect that skills instruction provided in the L2 – where proficiency is lower – would result in at least some discernable transfer to the L1, where proficiency is higher and more cognitive processing resources are likely to be available for the application of

those skills. Critical evaluation skills are thinking skills and its acquisition should not be bound to the language of instruction. The important question is whether constraints or limitations in the language to use could detrimentally affect critical evaluation performance.

The research reported here comprised two related studies. In Study 1, evaluative statements that second year Japanese university students produced in Japanese (their L1) and in English (their L2) were examined. These students had received instruction on academic discourse. Thus, they were not naïve as to the requirements of expressing evaluative language, and any differences in the writing they produced in L1 and L2 could be attributed to either the inherent structure of the language they were using or their proficiency in using that language. In Study 2, the same writing task was given to first year students who had received little instruction on academic discourse, and nothing explicit on the production of evaluative language. The purpose of this second study was to find out if the characteristics of L1 and L2 written work produced by the first year students, compared to their second year counterparts, differed. Hence, the aim was to find out whether the additional instruction that had been received by the more advanced second year students might have made a difference.

2 Study 1

Study 1 tested two hypotheses. The first was that students' production of evaluative statements in Japanese and in English would differ. The second was that the students' proficiency in a particular language would be related to the amount of evaluative statements they produce in that language.

2.1 Method

2.1.1 Participants

The participants were 110 Japanese university students in their second year of study in science and engineering disciplines. For these students, Japanese is L1 and English is L2. They were taking a compulsory English communication skills development course that covers oral and written

academic discourse using a task-based learning approach. The students came from four different classes in that course.

2.1.2 *Materials and Procedure*

As part of the communication skills course, the students were provided class instruction, textbook explanations and examples (Anthony, Rose, & Sheppard, 2010), and practice in the use of language appropriate for critical evaluation, including ranking and debating different reasons and other forms of alternatives (e.g., clearly stating the premises, and then drawing conclusions). These materials and the instruction were all provided in English.

For the purposes of the present investigation, the students were additionally provided with a single page Japanese translation of the part of the textbook dealing with how to make valid arguments. They were also supplied brief (one page) written examples (one in English and one in Japanese) of how alternative reasons could be ranked according to judgments about their relative importance. The example texts conveyed someone's opinion about the most important reason for learning the English language, among four possible reasons. The texts provided examples of evaluative statements and provision of support for claims, although those were not labeled or overtly identified in any way in the texts. The equivalence and appropriate use of language in the English and Japanese versions were checked by several bilingual teachers of the course. Although all materials provided in the course are usually in English, the Japanese versions were supplied in this case to avoid possible disadvantage to the students' production of evaluative language in Japanese (i.e., without the Japanese versions, it could be argued that the students might have simply been unfamiliar with the equivalent Japanese expressions for critical evaluation).

During two consecutive 90-minute class sessions of the course, the students were introduced to the Titanic and Space Shuttle Challenger disasters, including four basic causes that have been proposed for the occurrence of each of those disasters. During the class sessions, the students

participated in guided exercises to explore and discuss the disasters and their corresponding possible causes.

For homework, the students were asked to write two brief reports to explain what they considered to be the most important cause of each of the disasters. To avoid any possible misunderstandings about the requirements of the homework task, written instructions were provided in Japanese. The students were randomly assigned to write one report in English and the other in Japanese (i.e., if they were asked to write the Titanic report in English, they had to write the Challenger report in Japanese, and vice versa).

2.1.3 *Analyses*

The total number of sentences [Total], and the number of evaluative sentences (i.e., sentences where some evaluation of the relative value of the topic is made) [Evaluative], were counted and scored in the analysis of the students' written work.

Operational criteria were drawn up for determining what data counted as “evaluative sentences”. Examples of such requirements included: the sentence must explicitly say something about the worth or value of the subject, and that worth or value must be in comparison to something else. Conditional statements that explicitly convey a relative evaluation of the subject were counted. The following examples, in contrast, did not count: the use of simple adjectives or adverbs to describe something, prescriptive statements not explicitly expressing a relative evaluation or judgment, and conditional statements in general. Inter-rater reliability was checked by asking an independent coder to score a randomly selected sample of 25% of the data. Reliability coefficients obtained (Cronbach's alphas) were deemed to be satisfactory: they were .922 in English, and .960 in Japanese.

Language proficiency was measured (i) by using the students' TOEIC-IP (Test of English for International Communication – Institution Program) scores, and (ii) in terms of the complexity of the sentences the students produced in English and in Japanese (explained below). TOEIC is a

norm-referenced test of English listening comprehension and reading skills, widely used as a measure of students' English language proficiency levels in Japan (<http://www.ets.org/toeic>). Although there is a speaking and writing test version, it is not as widely used as the listening and reading test version, which is what the TOEIC-IP contains. Students in the faculty where the students came from are required to sit the TOEIC-IP test at regular intervals during their period of enrolment, and their scores on that test are made available to their teachers. The TOEIC test has been used in numerous previous studies as a measure of L2 proficiency (e.g., Dornyei & Katona, 1992; Eggly, Musial, & Smulowitz, 1999), and various studies have reported that it has at least adequate reliability and validity (e.g., In'nami & Koizumi, 2011; Perkins, 1987; Powers, Kim, & Wang, 2008). The TOEIC is also considered a good measure of non-native English speakers' ability to use English in real life situations rather than just in strictly academic contexts (Eggly et al., 1999). For these reasons, the authors of the present study decided to use the students' TOEIC-IP scores as one measurement of their English proficiency.

The other measure of proficiency used – complexity – was selected because, along with accuracy and fluency, it is generally considered as one of the key dimensions of proficiency in language production – both oral and written (e.g., Ellis, 2009; Ishikawa, 1995; Ortega, 2003; Skehan & Foster, 1999; Wolfe-Quintero, Inagaki, & Kim, 1998). However, unlike accuracy and fluency, for which it is difficult to come up with meaningful and equivalent/comparable measurements for English and Japanese, the authors of the present study were able to determine a viable way to measure complexity. Usually, analysis of production complexity in English entails calculating the ratio of words over sentences or terminable units (T-units; e.g., Ishikawa, 1995; Ortega, 2003; Richards, Platt, & Platt, 1992). However, what counts as a word in Japanese is not exactly the same as in English, and thus, such a method for calculating complexity was deemed inappropriate. Instead, the number of verbs that the students used were counted and divided by their total number of sentences to generate a measure of complexity. Verbs were chosen as they are

comparable in English and Japanese, and the use of more verbs can be considered a good indicator of more complex sentence structures (i.e., simple sentences usually contain only one verb, and inclusion of more verbs within one sentence usually requires the use of a more complex structures). It should also be noted that even though this measurement of complexity and the measurement of the students' use of critical evaluation were derived from the same source (i.e., the writing that the students produced), these measurements were not dependent on each other. In other words, students can produce complex writing without being critically evaluative, and vice versa: they can write evaluative sentences that are low in complexity.

Analyses of variance were conducted to compare the students' production of sentences (Total, and Evaluative) in English and in Japanese. Correlational analyses were carried out to examine possible relationships with the students' most recent TOEIC test scores and their complexity scores in English and in Japanese.

2.2 Results

Table 1 shows the means, and standard deviations (in brackets), obtained for Total and Evaluative categories of sentences in the students' written work in English and in Japanese.

Repeated measures 2 x 2 ANOVAs demonstrated that there were no significant effects due to the task (i.e., the Titanic compared to the Challenger reports), hence demonstrating task equivalence. The analysis however revealed significant effects due to language in both the total number of sentences written [Total], $F(1, 108) = 11.47, p = .001, \eta_p^2 = .020$; and the number of evaluative sentences [Evaluative], $F(1, 108) = 4.81, p = .030, \eta_p^2 = .006$.

These results indicate that the students wrote significantly more sentences in English compared to Japanese, but produced more of the target evaluative language in Japanese. The ANOVA results also revealed that the students wrote more verbs in Japanese (mean = 48.52, $SD = 14.01$) than in English (mean = 38.77, $SD = 12.54$), $F(1, 109) = 84.56, p < .001, \eta_p^2 = .119$. The language produced was also significantly more complex in Japanese as measured by the number of

verbs per sentence ($F(1, 109) = 100.8, p < .001, \eta_p^2 = .269$), indicating that the reason for the production of fewer sentences in Japanese was because the students' greater proficiency in that language enabled them to produce more complex sentences.

Because the total number of sentences that the students produced differed considerably between the two languages, the proportions of evaluative sentences (i.e., Evaluative as a proportion of Total) were also calculated and compared according to the language used. The comparison revealed a significant difference: $F(1, 109) = 20.32, p < .001, \eta_p^2 = .039$. This result indicates that the proportion of evaluative sentences was higher in the reports that the students wrote in Japanese compared to those they wrote in English.

For the correlational analysis, the correlations between the students' TOEIC scores and their writing complexity scores in English and in Japanese, and the evaluative sentences they produced in the respective languages (actual and proportions of total), were examined. The results are shown in Table 2. The students' TOEIC scores correlated significantly with the number of evaluative sentences they produced in English. The correlation between the students' TOEIC scores and the number of evaluative sentences they produced in Japanese was not significant. Where the complexity scores were concerned, the complexity of the students' writing in English significantly correlated with both the number and proportion of evaluative sentences they produced in English. The complexity of their writing in Japanese significantly correlated with the proportion of evaluative sentences they produced in Japanese. All cross language correlations between writing complexity and numbers/proportions of evaluative sentences were not significant. Taken together, these results lend support to the argument that language proficiency influences students' ability to evidence critical evaluation in their writing.

A significant correlation was also found between the proportions of evaluative sentences the students produced in English and in Japanese, $r = .72, p < .001, R^2$ (effect size) = .518. This result

suggests that students who were more critically evaluative in one language also demonstrated greater critical evaluation in the other language.

2.3 *Discussion*

Differences were found in both actual numbers and proportions of evaluative sentences that the students produced in English and in Japanese. The direction of the differences, however, was opposite to the language structure-based prediction: higher proportions of evaluative sentences were found in Japanese instead of English. This result suggests that the students were better at producing evaluative language in their L1.

The significant relationship found between the students' language proficiency and their critical evaluation performance lends support to the notion that language proficiency could explain apparent cultural differences in students' critical thinking performance. When using an L2 (such as English) in which they may lack adequate proficiency (e.g., less automated processing, inadequate linguistic knowledge), Asian students in Western countries may not be able to demonstrate critical thinking to the same degree as their native English-speaker counterparts because of the higher cognitive processing resource requirements of that L2, which depletes resources available in working memory for critical thinking. The important finding in the present study is that proficiency in both languages (English and Japanese) significantly correlated with critical evaluation performance in the respective language.

There still remains the possibility that the significant correlations between the two languages in the proportions of the target language produced, are a result of other factors, which are correlated to both first and second language proficiency, and which might affect the production of evaluative language – irrespective of the language being used. In other words, in addition to language proficiency, people's critical evaluation performance may also be influenced by some other individual factors – such as general intelligence, working memory capacity, or communication

skills – that are independent of the language being used for expression. It would be useful to identify and examine those other factors in future research.

As noted earlier, the student participants in this first study had already received instruction in academic discourse that included the use of evaluative language. Therefore, an important next question to address was, “To what extent had that instruction affected the relative production of evaluative language in English and in Japanese?” – which was pursued in Study 2.

3 Study 2

Study 2 tested three hypotheses. The first was that first year students would manifest lower use of the target evaluative language compared to the second year students from whom data were gathered in Study 1. The second was that the differences in evaluative language use would be consistent across English and Japanese. In other words, the prediction was that, like the second year students, the first year students would evidence greater use of evaluative statements in their L1 (Japanese) compared to their L2 (English). The third hypothesis was that the first year students’ proficiency in both languages would not be related to the proportions of evaluative language they produce. This hypothesis was based on the assumption that, without receiving instruction about how to undertake and convey critical evaluation, students’ language proficiency would not make a difference to their production of the target evaluative language as they would have very limited knowledge about the structures necessary for producing the required language.

3.1 Method

3.1.1 Participants

Forty-three Japanese university students who were in their first year of studies in the same science and engineering faculty as the students in Study 1 were added to the participants of Study 1, for a total of 153 participants. The additional students came from two classes of a compulsory first year English communication skills course which deals with various aspects of oral and written academic

discourse, but nothing explicit about evaluative language (which is not covered until the second year course).

3.1.2 *Materials, Procedure, and Analysis*

For one of their homework assignments, the students were given brief reading materials (in English and in Japanese) about the Titanic and Space Shuttle Challenger disasters, including the proposed causes of those disasters. These materials were drawn from the textbook used in the second year course. The Japanese translations were provided to these first year participants to ensure that their subsequent writing performance would not have been compromised by possible difficulties in understanding the English versions. The content of those materials were not covered in class.

The homework task that the students had to do was the same as that given to the second year students: to produce two brief reports to explain what they considered to be the most important cause of each of the disasters, after reading the materials provided. Like the second year students, they were randomly assigned to write one report in English and the other in Japanese. Also, like the second year students, they were provided with the one-page examples (one in English and one in Japanese) of how alternative reasons (for learning the English language) could be ranked according to judgments about their relative importance. The crucial difference was that the first year students were not provided class instruction and exercises on the use of academic discourse specifically pertaining to evaluative language.

The written reports that the students produced were analyzed and scored in the same manner described in the first study. The first and second year students' data were then compared using repeated measures ANOVAs. Correlational analyses were carried out to examine possible relationships between the students' language proficiency scores and the proportions of evaluative language they produced.

3.2 Results

Table 3 shows the means, and standard deviations (in brackets), obtained for the Total and Evaluative categories of sentences in the first year students' written work in English and in Japanese.

Analyses of variance revealed significant effects due to year of enrolment (first year compared to second year) in the students' scores for the total number of sentences [Total], $F(1, 152) = 22.90, p < .001, \eta_p^2 = .113$; and the number of evaluative sentences [Evaluative], $F(1, 152) = 27.20, p < .001, \eta_p^2 = .130$. A significant language effects were also found for Total, $F(1, 152) = 26.67, p < .001, \eta_p^2 = .026$. No significant interaction effects between language and year were found; nor were any significant effects found due to the task (Titanic versus Challenger).

These results indicate that, compared to the second year students, the first year students wrote fewer sentences in total for their reports, and fewer evaluative statements. Like the second year students, the first year students wrote significantly more sentences in English than in Japanese. Analysis of the complexity of the sentences they produced revealed that the total number of verbs they used was likewise higher in Japanese compared to English, $F(1, 43) = 14.80, p < .001, \eta_p^2 = 0.032$; and the index of complexity (total number of verbs divided by total number of sentences) was higher for what they wrote in Japanese compared to what they wrote in English, $F(1, 43) = 75.58, p < .001, \eta_p^2 = .409$. This finding indicates that the first year students, like their second year counterparts, were producing more complex sentences in Japanese compared to English.

The results of the correlational analysis are shown in Table 4. None of the correlations were significant – suggesting that, without instruction, the students' proficiency in both languages generally did not make a difference to their production of the target evaluative language.

3.3 Discussion

The results of Study 2 showed that the second year students wrote more sentences in their reports, and produced more of the target evaluative language, compared to the first year students. This

finding suggests that instruction on appropriate language to use – which had been provided to the second year students – can improve students' abilities in manifesting critical evaluation in their written work. Although as noted the instruction was provided almost entirely in English, the significant language effects found were all in favor of the Japanese language, which suggests that there is transfer across the languages in skills acquisition. In other words, skills taught and learned in English also produce improvements in the production of evaluative language in Japanese.

The results of the correlational analysis underline the importance of providing instruction to cultivate student competence in critical evaluation. No significant correlations were found between the proportions of evaluative language the students produced and their proficiency scores in English and in Japanese. This therefore suggests that general proficiency on its own is not adequate: for students to evidence critical evaluation in their written work, it is necessary to also provide them with instruction on the specific language forms and structures to use.

4 General Discussion

The findings of this study provide evidence that, at least for Japanese students, using the Japanese language (their L1) presents no disadvantage compared to English (their L2) in the production of evaluative language (i.e., the Japanese language structure is not a limiting factor). How Japanese students' evaluative language use might compare to that of students whose first language is structured differently (e.g., native English speakers responding to the same tasks), or students who are fully bilingual in Japanese and English, would need to be examined in future research. However, in the present study, there appeared to be no obvious deficits in evaluative language production in Japanese among the second year students who had received instruction in the necessary academic discourse.

There is evidence in the present study, however, that language proficiency can influence the production of evaluative language. Although the conceptualization of critical thinking is likely to be

independent of the language being used, the use of a language that one is not so proficient in requires greater cognitive processing resources and thus limits the remaining resources that could be utilized for the expression of critical thinking. In this study, the significant correlations between the students' English proficiency scores and the amount of evaluative language they evidenced in their English writing indicate that performance varied with L2 proficiency. This provides useful evidence to corroborate previously made claims (e.g., Floyd, 2011; Lun et al., 2010; Paton, 2005) that some of the shortcomings in critical thinking skills manifested by international students can be attributed to their having to use an L2 in which they may not be as proficient compared to their native speaker counterparts. Although those significant correlations were not high, results were also obtained in the present study showing that the students' Japanese writing complexity significantly correlated with the proportion of evaluative language they produced in Japanese. This finding provides further support for the argument that language proficiency can influence the production of evaluative language, this time in the L1.

4.1 Conclusion

The finding that L2 proficiency is a potential limiting factor in students' use of critical thinking skills suggests that educational strategies for improving non-native speaker students' proficiencies in English (or the language used in the host country) would be beneficial. Successful implementation of such strategies would likely reduce the perceived deficiencies in Asian and other foreign students' critical thinking competencies (cf. Lee & Carrasquillo, 2006; Robertson et al., 2000). The findings of this study also indicate that appropriate classroom instruction is essential for facilitating the development of students' abilities to demonstrate critical thinking competencies in the work they produce. The second year students evidenced similar writing profiles to those of first year students; however, having received instructions in academic discourse relevant to critical evaluation, they also produced more of the target evaluative language. They did this in both languages, L1 and L2, even though academic discourse instruction was primarily provided in the L2 – suggesting some

transfer of skills across languages. Furthermore, the lack of significant correlations between the first year students' L1 and L2 proficiency and their use of the target evaluative language underline the necessity of providing critical evaluation instruction to students. Without receiving explicit instruction, most students – irrespective of their language proficiency – would not likely know how to effectively demonstrate critical evaluation in the work they produce.

List of Funder and Grant Number

This research was supported by a grant-in-aid (XXXXXXXX) received from the Japan Society for the Promotion of Science.

Conflict of Interest

There is no conflict of interest known to the authors.

Acknowledgements

The authors would like to thank Xxxxx XXXXXXXX and XXXXXXXX Xxxx for their invaluable assistance during various stages of this study.

References

- Ackerman, P. L. (1986). Individual differences in information processing: An investigation of intellectual abilities and task performance during practice. *Intelligence, 10*, 101–139.
- Anthony, L., Rose, R., & Sheppard, C. (2010). *Concept building and discussion: Foundations*. Tokyo: DTP Publishing.
- Atkinson, D. (1997). A critical approach to critical thinking in TESOL. *TESOL Quarterly, 31*, 9–37.
- Au, T. K. F. (1983). Chinese and English counterfactuals: The Sapir-Whorf hypothesis revisited. *Cognition, 15*, 155–187.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (1998). *Human memory*. Boston: Allyn & Bacon.
- Bloom, A. (1981). *The linguistic shaping of thought*. Hillsdale, N J: Erlbaum.
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22.
- Chen, J. Y. (2007). Do Chinese and English speakers think about time differently? Failure of replicating Boroditsky (2001). *Cognition, 104*, 427–436.
- Davies, W. M. (2006). An 'infusion' approach to critical thinking: Moore on the critical thinking debate. *Higher Education Research & Development, 25*, 179–193.
- Dornyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing, 9*, 187–206.
- Eggy, S., Musial, J., & Smulowitz, J. (1999). The relationship between English language proficiency and success as a medical resident. *English for Specific Purposes, 18*, 201–208.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*, 474–509.
- Ennis, R. H. (1998). Is critical thinking culturally biased? *Teaching Philosophy, 21*, 15–33.

- Fisher, A., & Scriven, M. (1997). *Critical thinking: Its definition and assessment*. Norwich, UK: Centre for Research in Critical Thinking.
- Floyd, C. B. (2011). Critical thinking in a second language. *Higher Education Research and Development, 30*, 289–302.
- Fox, H. (1994). *Listening to the world*. Urbana, IL: National Council of Teachers of English.
- Glassner, A., Weinstock, M., & Neuman, Y. (2005). Pupils' evaluation and generation of evidence and explanation in argumentation. *British Journal of Educational Psychology, 75*, 105–118.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*, 449–455.
- Hill, J. H., & Mannheim, B. (1992). Language and world view. *Annual Review of Anthropology, 21*, 381–406.
- Hockett, C. E. (1954). Chinese versus English: An exploration of the Whorfian thesis. In H. Hoijer (Ed.), *Language in Culture* (pp. 106–123). Chicago: University of Chicago Press.
- In'nami, Y., & Koizumi, R. (2011). Factor structure of the revised TOEIC test: A multiple-sample analysis. *Language Testing, 29*, 131–152.
- Ishikawa, S. (1995). Objective measurement of low-proficiency EFL narrative writing. *Journal of Second Language Writing, 4*, 51–69.
- Itakura, H., & Tsui, A. B. M. (2011). Evaluation in academic discourse: Managing criticism in Japanese and English book reviews. *Journal of Pragmatics, 43*, 1366–1379.
- January, D., & Kako, E. (2007). Re-evaluating evidence for the linguistic relativity hypothesis: Response to Boroditsky (2001). *Cognition, 104*, 417–426.
- Kong, K. C. C. (2005). Linguistic resources as evaluators in English and Chinese research articles. *Multilingua, 24*, 275–308.
- Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity, 4*, 70–76.

- Lee, K. S., & Carrasquillo, A. (2006). Korean college students in United States: Perceptions of professors and students. *College Student Journal*, *40*, 442–456.
- Liu, L. G. (1985). Reasoning counterfactually in Chinese: Are there any obstacles? *Cognition*, *21*, 239–270.
- Lun, V. M. C., Fischer, R., & Ward, C. (2010). Exploring cultural differences in critical thinking: Is it about my thinking style or the language I speak? *Learning and Individual Differences*, *20*, 604–616.
- Manalo, E., Kusumi, T., Koyasu, M., Michita, Y., & Tanaka, Y. (2013). To what extent do culture-related factors influence university students' critical thinking use? *Thinking Skills and Creativity*, *10*, 121–132.
- Manalo, E., Kusumi, T., Koyasu, M., Michita, Y., & Tanaka, Y. (2015). Do students from different cultures think differently about critical and other thinking skills? In M. Davies, & R. Barnett (Eds.), *Palgrave handbook of critical thinking in higher education* (pp. 299–316). New York: Palgrave Macmillan.
- Manalo, E., & Uesaka, Y. (2012). Elucidating the mechanism of spontaneous diagram use in explanations: How cognitive processing of text and diagrammatic representations is influenced by individual and task-related factors. In P. Cox, P. Rodgers, & B. Plimmer (Eds), *Diagrammatic representation and inference. Lecture Notes in Artificial Intelligence (LNAI) 7352* (pp. 35–50). Berlin Heidelberg: Springer-Verlag.
- Manalo, E., & Uesaka, Y. (2014). Students' spontaneous use of diagrams in written communication: Understanding variations according to purpose and cognitive cost entailed. In T. Dwyer, H. Purchase, & A. Delaney (Eds.), *Diagrammatic representation and inference. Lecture Notes in Artificial Intelligence (LNAI) 8578* (pp. 78–92). Berlin, Heidelberg: Springer-Verlag.

- Neisser, U., Boodoo, G., Bouchard, T. J. Jr., Boykin, A. W., Brody, N., Ceci, S. J., ... Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*, 492–518.
- Paton, M. (2005). Is critical analysis foreign to Chinese students? In E. Manalo, & G. Wong-Toi (Eds.), *Communication skills in university education: The international dimension* (pp. 1–11). Auckland: Pearson Education.
- Perkins, K. (1987). Review of the ‘Test of English for International Communication’. In J. C. Alderson, K. J., Krahnke, & C. W. Stansfield (Eds.), *Reviews of English language proficiency tests* (pp. 81–83). Washington, DC: TESOL.
- Powers, D. E., Kim, H.-J., Weng, V. Z., & VanWinkle, W. (2009). *The redesigned TOEIC (Listening and Reading) Test: Relations to test-taker perceptions of proficiency in English*. Retrieved March 28, 2016, from <http://dx.doi.org/10.1002/j.2333-8504.2009.tb02175.x>
- Richards, J., Platt, J., & Platt, H. (1992). *Longman dictionary of language teaching and applied linguistics*. Harlow, UK: Longman.
- Robertson, M., Line, M., Jones, S., & Thomas, S. (2000). International students, learning environments and perceptions: A case study using the Delphi technique. *Higher Education Research & Development*, *19*, 89–102.
- Scollon, R., & Wong-Scollon, S. (1991). Topic confusion in English-Asian discourse. *World Englishes* *10*, 113–125.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*, 93–120.
- Takano, Y. (1989). Methodological problems in cross-cultural studies of linguistic relativity. *Cognition*, *31*, 141–162.

- Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-cultural Psychology, 24*, 445–462.
- Takano, Y., & Noda, A. (1995). Interlanguage dissimilarity enhances the decline of thinking ability during foreign language processing. *Language Learning, 45*, 657–681.
- Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction, 14*, 359–379.
- Thomm, E., & Bromme, R. (2011). “It should at least seem scientific!” Textual features of “scientificness” and their impact on lay assessments of online information. *Science Education, 96*, 187–211.
- Tse, C.-S., & Altarriba, J. (2008). Evidence against linguistic relativity in Chinese and English: A case study of spatial and temporal metaphors. *Journal of Cognition and Culture, 8*, 335–357.
- Vorster, J., & Schuring, G. (1989). Language and thought: Developmental perspectives on counterfactual conditionals. *South African Journal of Psychology, 19*, 34–38.
- Whorf, B. L. (1956). The relation of habitual thought and behavior to language. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 134–159). Cambridge, MA: MIT Press.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. Technical report no. 17. Honolulu, Hawai'i. University of Hawai'i, Second Language Teaching and Curriculum Center.

Table 1

Means (with Standard Deviations in brackets) of Total and Evaluative Sentences that the Second Year Students Produced in English and in Japanese

	Total	Evaluative
English	20.31 (5.59)	3.46 (1.76)
Japanese	18.68 (5.92)	3.75 (1.80)

Table 2

Correlation Coefficients Between Second Year Students' TOEIC and Writing Complexity Scores, and Numbers and Proportions of Evaluative Sentences They Produced, According to the Language Used (Effect Sizes, R^2 Shown in Brackets)

Scores	English Evaluative		Japanese Evaluative	
	Sentences		Sentences	
	Number	Proportion	Number	Proportion
TOEIC	.23* (.053)	.14 (.020)	.18 (.032)	-.06 (.004)
English Complexity	.19* (.036)	.24* (.058)	.15 (.023)	.17 (.029)
Japanese Complexity	.02 (.000)	.09 (.008)	-.07 (.005)	.33*** (.109)

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 3

Means (with Standard Deviations in brackets) of Total and Evaluative Sentences that the First Year Students Produced in English and in Japanese

	Total	Evaluative
English	15.70 (8.20)	2.09 (1.71)
Japanese	12.84 (6.28)	2.26 (1.24)

Table 4

Correlation Coefficients Between First Year Students' TOEIC and Writing Complexity Scores, and Proportions of Evaluative Sentences They Produced, According to the Language Used (Effect Sizes, R^2 Shown in Brackets)

Scores	Proportions of Evaluative Sentences	
	English	Japanese
TOEIC	.03 (.000)	– .02 (.000)
English Complexity	.16 (.026)	– .18 (.032)
Japanese Complexity	.01 (.000)	.14 (.020)