Do Language Structure or Language Proficiency Affect Critical Evaluation?

Emmanuel Manalo (emanuel.manalo@gmail.com)
Kyoko Watanabe (watanabe@emmanuel.manalo@gmail.com)
Chris Sheppard (chris@waseda.jp)
Faculty of Science and Engineering, Waseda University
3-1-0 Ohkubo, Shinjuku-ku, Tokyo 169-8555, JAPAN

Abstract
This study examined whether language structure or language proficiency might influence students’ use of evaluative language in written reports, and whether instruction might improve students’ use of evaluative language. Reports in Japanese and in English written by second year Japanese university students, who had received instruction in academic discourse pertaining to critical evaluation, were analyzed for use of evaluative statements. This revealed no disadvantage for use of the Japanese language, which is considered as having a more indirect structure that may make critical evaluation more difficult. English proficiency test scores, however, were found to correlate with production of evaluative statements in English, but not in Japanese, suggesting that inadequate second language proficiency could limit critical evaluation use. The second year students’ use of evaluative statements was also found higher than their first year counterparts’ (who had not yet received instruction), suggesting that such instruction is beneficial for skills development in both languages.

Keywords: critical evaluation; critical thinking; language structure; second language proficiency; cognitive cost

Introduction
The cultivation of students’ abilities to critically evaluate the soundness of knowledge claims and arguments is considered as one of the most important objectives of education (e.g., Glassner, Weinstock, & Neuman, 2005) and, with the proliferation of unvetted available information through the Internet and other forms of media in modern societies, the ability to determine credibility has become crucial (e.g., Thomm & Bromme, 2011). Developing students’ critical thinking skills (the broader set of skills to which critical evaluation belongs) is, however, not without its challenges (e.g., Halpern, 1998). There are various factors that have been claimed to affect people’s use of critical thinking, including some culture-related factors. Asian students, in particular, have often been portrayed as lacking in critical thinking skills compared to Western students (e.g., Atkinson, 1997; Fox, 1994), and many tertiary instructors have been found to subscribe to such a view (e.g., Lee & Carrasquillo, 2006; Robertson, Line, Jones, & Thomas, 2000).

One explanation that has been put forward for the apparent differences in critical thinking skills manifested by students from different cultural groups concerns the structure of their native language. This explanation posits that, due to their structure, some languages may present constraints in the ease with which certain thinking skills can be carried out or expressed. This explanation is sometimes referred to as the “Sapir-Whorf hypothesis” (see Au, 1983; Hockett, 1954), which suggests that languages differ in the relative ease with which they can be used to convey certain ideas. 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.

More recent 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 notion 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. For example, in Japanese, criticism is usually indirectly conveyed and is frequently preceded by an apology.

Language Structure or Language Proficiency?
Previous studies, however, had not clarified whether language structures could actually impose constraints in what users of the language can do. Although the earlier-mentioned study by Bloom (1981) claimed to have found evidence for this where counterfactual thinking in the Chinese language is concerned, subsequent investigations failed to replicate or support Bloom’s results (Au, 1983). Thus it remains unclear whether, for example, the structure of a language like Japanese would make it relatively more difficult to undertake tasks like critical evaluation (cf. Itakura and Tsui’s, 2011, findings), and hence make a person appear less competent in his or her critical thinking skills.

Concerning international students who have been reported as appearing less competent in critical thinking skills (cf. Lee & Carrasquillo, 2006; Robertson et al., 2000), there is another possible explanation that other authors have previously suggested (e.g., Floyd, 2011; Lun, Fischer, & Ward, 2010; Paton, 2005) but which had not been adequately tested. This explanation hinges on the fact that many international students have to use a second language (L2), like English, in their host environment. It suggests that, if a person is not so proficient in a language, he or she would generally manifest lower competence in carrying out tasks when using that language. Tasks that are likely to get affected include cognitive tasks like critical thinking.
This possible influence of language proficiency on critical thinking skills application can be explained in terms of cognitive cost (i.e., the mental resources cost associated with executing tasks). 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 (i.e., the cognitive cost would be higher). The application of critical thinking skills would likewise require the use of working memory resources. The resources available in working memory, however, are limited (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) 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.

Overview of the Present Study

The main purpose of the present study was to examine the possible influences of language structure, and proficiency in L2, on students’ manifestation of critical thinking in their writing. 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 being used, Japanese or English. Critical thinking was operationalized as students’ use of evaluative statements. Such use was chosen for investigation because it comprises a salient expression of critical evaluation, which in turn is central to the notion of critical thinking application (cf. 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.

As this study was focusing on students’ written work in both L1 and L2, it was equally important to consider whether using an L2 may detrimentally affect students’ critical evaluation performance. Thus, possible relationships between students’ TOEIC test scores (Test of English for International Communication, 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) and their production of evaluative statements were investigated. The question here 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’ TOEIC scores and their evaluative statements production in the L2, but not in the L1. A relationship in the L1 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 the TOEIC test, and evidence better performance in tasks like critical evaluation – in both their L1 and L2.

The research conducted 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 (particularly the L2).

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 – and hence, whether the additional instruction that had been received by the more advanced second year students might have made a difference.

Study 1

The first study was carried out to test the hypothesis that students’ production of evaluative statements in Japanese
and in English would differ. A second hypothesis was also tested: that, if L2 proficiency is a limiting factor in students’ critical evaluation performance, their TOEIC scores would be related to their evaluative statements production in L2, but not in L1. Lower use of evaluative statements in the students’ L2 work should also be observable if this L2 proficiency hypothesis applies.

Method
Participants The participants were 111 Japanese university students in their second year of study in science and engineering disciplines. For these students, Japanese is L1 and English is L2. These students were taking a compulsory English communication skills development course that covers oral and written academic discourse in task-based discussion and research development. The students came from four different classes in that course.

The students were required to sit the TOEIC test at regular intervals during their period of enrolment, and their scores on that test were available to their course teachers.

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 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 90-minute class sessions of the communication skills 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).

Analyses The following were counted and scored in the analysis of the students written work:

a) Number of sentences [Total];
b) Number of evaluative sentences (i.e., sentences where some evaluation of the relative value of the topic is made) [Evaluative];
c) Number of evaluative sentences specifically about the causes of the disaster (i.e., sentences where some evaluation is made about the relative importance of the causes given for the occurrence of the disaster) [Causes];
d) Number of evaluative sentences that are supported by reason or evidence of some kind [Supported].

Operational criteria were drawn up for determining what data counted under each of these categories. For example, where “evaluative sentences” were concerned, the following were required: 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 (e.g., .922 and .940 in English and .960 and .963 in Japanese for the “Evaluative” and “Causes” scores, respectively).

Analyses of variance were conducted to compare the students’ scores in each of the categories noted above in English and in Japanese. Correlational analyses were carried out to examine possible relationships with the students’ most recent TOEIC test scores.

Results
Table 1 shows the means, and standard deviations (in brackets), obtained under each category for the students’ written work in English and in Japanese.

No significant effects were found due to the task (i.e., the Titanic compared to the Challenger reports). The analysis however revealed significant effects due to language in the total number of sentences written [Total], $F(1, 110) = 11.51$,
students’ productivity and use of evidence in writing, irrespective of the language being used.

As noted earlier, the student participants in this first study had already received instruction in academic discourse that includes 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 the second study.

![Figure 1: Evaluative, evaluative about causes, and evaluative supported sentences, as proportions of the total number of sentences written, in English and in Japanese.](image)

**Table 1: Mean report scores according to language used**

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Evaluat.</th>
<th>Causes</th>
<th>Support.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>20.35</td>
<td>3.48</td>
<td>3.38</td>
<td>2.08</td>
</tr>
<tr>
<td></td>
<td>(5.58)</td>
<td>(1.77)</td>
<td>(1.77)</td>
<td>(1.42)</td>
</tr>
<tr>
<td>Japanese</td>
<td>18.72</td>
<td>3.76</td>
<td>3.67</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>(5.91)</td>
<td>(1.80)</td>
<td>(1.80)</td>
<td>(1.38)</td>
</tr>
</tbody>
</table>

Because the total number of sentences that the students wrote in English and in Japanese differed, the proportions (i.e., Evaluative, Causes, and Supported sentences as proportions of Total) were also calculated and compared according to the language used. The comparisons revealed significant differences in each case: for Evaluative, $F(1, 110) = 20.17$, $p < .001$, $\eta^2_p = .155$; for Causes, $F(1, 110) = 20.29$, $p < .001$, $\eta^2_p = .156$; and for Supported, $F(1, 110) = 24.90$, $p < .001$, $\eta^2_p = .185$. These results, depicted in Figure 1, indicate that the proportions of Evaluative, Causes, and Supported sentences were higher in the reports that the students wrote in Japanese compared to those they wrote in English.

The results of the correlational analysis are shown in Table 2. In the students’ written work in English, TOEIC scores correlated significantly with all categories of scores obtained. However, in Japanese, TOEIC scores significantly correlated only with Total and Supported sentences.

**Table 2: Correlation coefficients between students’ TOEIC scores and categories of their report scores, according to the language used (effect sizes shown in brackets)**

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Evaluat.</th>
<th>Causes</th>
<th>Support.</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>.22*</td>
<td>.22*</td>
<td>.22*</td>
<td>.23*</td>
</tr>
<tr>
<td></td>
<td>(.047)</td>
<td>(.049)</td>
<td>(.048)</td>
<td>(.051)</td>
</tr>
<tr>
<td>Japanese</td>
<td>.27**</td>
<td>.18</td>
<td>.15</td>
<td>.25**</td>
</tr>
<tr>
<td></td>
<td>(.075)</td>
<td>(.032)</td>
<td>(.023)</td>
<td>(.062)</td>
</tr>
</tbody>
</table>

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

**Study 2**

The purpose of the second study was to examine whether first year students manifest lower use of the target evaluative language compared to the second year students, and whether any such differences might be consistent across English and Japanese.

**Method**

**Participants** The participants were 44 Japanese university students who were in their first year of studies in the same science and engineering faculty as the students in Study 1. The 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).

**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.

Results

Analyses of variance revealed significant effects due to year
of enrolment (first year compared to second year) in the
students’ scores for: Total, $F(1, 153) = 23.37, p < .001, \eta^2 = .133$; Evaluative, $F(1, 153) = 27.79, p < .001, \eta_p^2 = .154$; for Causes, $F(1, 153) = 27.15, p < .001, \eta p^2 = .151$; and for Supported, $F(1, 153) = 6.99, p = .009, \eta_p^2 = .044$. Significant language effects were also found for Total, $F(1, 153) = 26.84, p < .001, \eta_p^2 = .149$; Causes, $F(1, 153) = 4.31, p = .04, \eta_p^2 = .027$; and Supported, $F(1, 153) = 14.03, p < .001, \eta_p^2 = .084$. 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. They also produced fewer evaluative
statements (evaluative sentences, evaluative sentences about
causes, evaluative sentences that are supported). These
differences in the students’ production of evaluative language are depicted in Figure 2. Significant language
differences were found in the total number of sentences,
number of evaluative sentences about causes, and number of
evaluative sentences with support that the students wrote: in
each case, the students produced more in Japanese
compared to English.

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.

![Figure 2: Mean numbers of evaluative, evaluative about
causes, and evaluative supported sentences produced in
English and Japanese reports by the first year (English 1,
Japanese 1) and second year (English 2, Japanese 2)
students.](image)

General Discussion

The findings of this study provide clear 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 be a limiting factor in the
production of evaluative language. The significant
correlations between the students’ TOEIC scores and their
production of evaluative sentences in English (their L2) –
but not in Japanese (their L1) – indicate that performance
varied with L2 proficiency. This provides useful evidence to
corroborate previously made claims (e.g., Floyd, 2011; Lun

2971
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.

The finding about L2 proficiency being a potential limiting factor in students’ use of the target critical evaluation language suggests that, to address the perceived deficiencies in Asian and other foreign students’ critical thinking skills, educational strategies that would improve their proficiencies in English (or whatever language is used in the host country) would be helpful.

The findings of this study also show that appropriate classroom instruction promotes university students’ development of skills in critical evaluation. 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.

**Acknowledgments**

This research was supported by a grant-in-aid (A11550900) received from the Japan Society for the Promotion of Science. The authors would like to thank Takashi Kusumi for his support.

**References**


