Language Proficiency, Reading Development, and Learning Context

Liu Li

Ball State University

Introduction

In the era of globalization, the number of students studying abroad is increasing dramatically. Accordingly, there is an increasing interest among researchers in students' language gains in the study-abroad (SA) context. However, a large number of these studies have aimed to understand the improvement of speaking skills (Brecht & Robinson, 1995; Freed, 1995a; Lafford 1995, 2004; Segolowiz & Freed 2004). Only a fraction of the studies explored the gains of literacy skills in the SA context, and the majority of these studies focused on writing skills (Serrano, 2010). To deepen our understanding of second language (L2) reading acquisition in different contexts, it is necessary to examine students' reading development in the SA context and at home (AH) context, as well as to find out the effect of context on reading development.

In addition, there is a lack of diversity within research regarding the students' L2 in the area of SA context (Serrano, 2010). Most of the research on learning context so far has examined the learning of Spanish. The rest of it looked at the learners of English, French or Japanese. There is little research on the learning of other languages, such as Chinese, in the SA context. Yet, China has become one of the most popular spots for international students due to its rapid economic development. Study abroad programs in China are rapidly growing along with domestic programs at many universities worldwide. For example, China is now one of the most popular SA destinations among US-based college students (Institute of International Education, 2012). Such a dearth of empirical research on these students calls for more attention to their language learning experiences in the SA context. It is hoped that the insights from this study would enrich various language programs in both study-abroad and domestic settings.

Literature review

It is commonly assumed that a student in a SA context would achieve higher levels of proficiency than in formal classroom settings where only instruction is provided, because it is believed that studying in a natural language environment involves naturalistic language learning through out-of-class contact with the L2. Thus, SA is considered valuable for the opportunities it provides for naturalistic learning. However, Ellis (1994) reviewed some studies and found this assumption questionable in two ways. First, naturalistic settings did not necessarily lead to higher proficiency. Instead, educational settings often brought about higher proficiency, especially higher grammatical competence. Second, even in naturalistic settings, the type of the contact was more influential than the amount of contact with the target language. Therefore, students did not always successfully improve every aspect of their language competence in naturalistic settings. While SA students may improve in such skills as listening and

speaking, or at least feel that they have improved in those areas, the same does not necessarily hold true for reading and writing skills and sociolinguistic competence (Iwasaki, 2007).

In addition, students' language proficiency and types of contact also need to be taken into consideration. For example, in a study of the effects of contact on the L2 proficiency of 40 undergraduate American students in a six-week study-abroad program in France, Freed (1990) reported that, for the lower-level students, increased interactive contact (e.g., speaking with native speakers) led to clear gains in the test scores on grammar and reading while it did not have the same effect for advanced-level students. In contrast, non-interactive contact (e.g., reading newspapers, watching television) benefited the advanced but not the lower-level students. Freed (1995a, 1998) and Coleman (1997a; 1997b) provided surveys of previous studies of the effects of study-abroad programs on L2 learning. Among their main findings, one noticeable phenomenon is that the higher the students' initial level of proficiency, the lower the gains in proficiency as a result of studying abroad.

In terms of linguistic dimensions, most of these studies of foreign language (FL) skills in a SA context so far have focused on gains in listening and speaking, whereas little attention has been given to literacy competency (Taillefer, 2005; Serrano, 2010). Only a few studies on reading development have been done in the past. Among these studies, Carroll (1967) discovered that time spent abroad could significantly predict L2 reading comprehension, as measured by the Modern Language Association reading tests. Freed (1990) discovered that more advanced SA French students could benefit linguistically from the so-called passive, "non-interactive" contact with the language, a category that included reading. Brecht, Davidson, & Ginsberg (1995) measured the reading proficiency gains made by students studying Russian abroad for a semester. Using ETS (define acronym) Reading and ACTR (define acronym) qualifying exams, they found that students made significant gains in reading during SA. Lapkin, Hart and Swain (1995) used both self-assessment questionnaires and proficiency tests in the various language skill areas including reading skills to determine the "linguistic impact" of a bilingual exchange program in Canada. They also found that SA students made significant gains in reading competence. Huebner (1995) made use of qualitative methods, such as diaries, interviews and observations, to compare beginning Japanese students in SA and AH programs. Huebner (1995) compared the gains of students who studied abroad to those of students who studied in the United States, and he provided findings that shed light in particular on reading development. Huebner (1995) found that students without any prior experience studying Japanese improved more during a summer intensive program in Japan than those who studied in an equivalent summer intensive program in the United States, both in speaking and reading. Reading comprehension assessed by the ETS Japanese proficiency test (Education Testing Service 1981, 1986) showed that the learners who studied in Japan had an advantage. He found that the widest difference in test scores between the AH and SA groups was in reading. Huebner related the advantage to the students' sense of urgent need to be literate in an environment where they found themselves to be largely illiterate.

Freed (1998) summarized the multidimensional work in the 1990s and reported that foreign language (FL) grammar proficiency was a strong predicator of improvement in speaking and listening as well as reading during the SA context. Freed considered that students' "linguistic identities extend beyond the expected acquisition of oral skills to a new self-realization in the social world of literacy" (p. 50).

These studies provided useful insight into reading development in different contexts. However, in many of these studies, reading was examined only as one aspect of students' general linguistic achievement. Only a handful of studies, to my knowledge, have specifically targeted FL reading. These studies were devoted entirely to the complex process of reading development in the SA context. Kline (1998) pointed out that earlier quantitative research into cognitive and linguistic aspects of reading comprehension had some limitations, which made it very difficult to draw conclusions about reading and the effect of SA context on reading development. Therefore, she took a qualitative, ethnographic approach to explore the process of literacy as social practice among American students studying in France. Her research resulted in an extensive description of learners' reading habits. With her emphasis on the cultural aspect of reading, she observed that several factors such as gender, group affiliations and host-family preferences could influence reading habits. However, she made no connection between reading habits and the development of reading comprehension. Therefore, her study still had the same problem she pointed out in quantitative studies.

More recently, Dewey (2004) compared the reading development of a group of students who studied in a nine-week intensive program in Japan and those who studied in an eleven-week intensive immersion program in the United States. He found the SA students had more confidence about their reading comprehension than the AH group, whereas the free-recall and the vocabulary tests indicated no difference between the two groups. While Huebner's (1995) findings suggest that being in Japan helps to motivate students at an elementary level, Dewey's (2004) findings indicate that such an advantage may not necessarily lead to substantial growth observable by some measures. Dewey's study on the comparison between the two contexts is not cross-sectional. Since proficiency may play a role in students' progress in different contexts (Freed, 1998), it is possible that if the students are at a certain proficiency level different from Dewey's, the results might turn out differently.

Taillefer (2005) conducted a large-scale cross-sectional study to examine the reading comprehension and reading strategy use of 177 European SA students with various first language (L1) backgrounds. The results showed that, at given levels of FL competency, the students' reading skills and strategy use differed significantly according to their particular academic literacy and sociolinguistic background. This study provided valuable information about the strategies used by the SA students. However, Taillefer made no comparison between SA and AH students in this study. It is hard to draw a conclusion that the use of these strategies was unique only among SA students due to the influence of the learning context.

Serrano (2010), in his review on SA research, pointed out that one of the main limitations detected in the design of many studies in the field was the lack of a control group who remained in their home country while a comparative group was staying abroad. Another problem he mentioned was the scarcity of studies that adopted a cross-sectional approach in their design.

Due to the limitations in previous studies, it is necessary to conduct a cross-sectional study with both the SA and AH groups for comparison if we want to have a more complete picture of the impact of contexts on students' reading development. In terms of methodology, it is important to combine quantitative and qualitative methods for a thorough exploration of reading development, because reading, as defined by Urquhart and Weir (1998), is not just a product but also a process of comprehension. Quantitative measures such as comprehension test scores are good indicators of reading as a product of comprehension, but they tell us little about the process of comprehension. In this case, qualitative methods, such as observation and interviews, will be helpful. More importantly, it has been

found that reading not only calls into play the varying roles of cognitive faculty but also some social factors around the readers (Grabe, 2009). A qualitative approach could be more effective in exploring the impact of various social factors involved in different learning contexts.

To find out whether reading is developed differently in SA and AH contexts, the current study adopted the view that reading is both a product and a process (Urquhart & Weir, 1998). While measuring learners' reading comprehension, this study also delved into the process of reading through examining the reading strategies used by L2 readers during the comprehension process. To this end, both the qualitative and quantitative approaches were employed. To examine the cognitive factors involved in reading, comprehension tests were used to find out the results of reading. At the same time, a think-aloud protocol, a strategy questionnaire, and observation were used to explore the process of reading. As to the social factors involved, interviews and the Language Contact Profile were used.

In this study, we focused on one important aspect of the reading process: the use of strategies. Reading is a complicated process that involves effective comprehension strategies to use. With more difficult texts, strategic processing abilities are needed to maintain a high level of comprehension for extended periods of time (Pressley, 2002). Strategies are defined as learning techniques, behaviors, problem-solving or learning skills which make learning more effective and efficient (Oxford & Crookkall, 1989). Reading strategies indicate how readers conceive of a task, how they make sense of what they read, and what they do when they don't understand. In short, strategies are processes used by the learner to enhance reading comprehension and overcome comprehension failures (Singhal, 2001). It has long been found that the use of strategies is closely related to readers' language proficiency. In the 1980s, researchers on L1 reading discovered that the use of various strategies was effective in improving students' reading comprehension (Baker & Brown, 1984; Brown, 1981; Palinscar & Brown, 1984). At the same time, studies in L2 reading found that less proficient students used fewer strategies and used them less effectively in their reading comprehension (Garner, 1987; Waxman & Padron, 1987; Brown, Armbruster & Baker, 1986). Anderson (1991) carried out an investigation on individual differences in strategy uses by adult L2 readers. In his study, although both high scoring and low scoring students used the same set of strategies, high scoring students seemed to be applying strategies more effectively and appropriately. Overall, these studies suggested that high proficiency students seemed to use strategies differently from those used by low proficiency-t students, and also appeared to apply them more effectively.

Think-aloud protocols and observation have been widely used to analyze reading processes. Because comprehension is normally a silent, hidden process, researchers cannot determine what is happening in a reader's minder by simple observation or by product-based assessment. Asking readers to report verbally is the most direct way to access this process. Meanwhile, think-aloud protocols and observation are alternative assessments for investigating reading behaviors or strategies. A major benefit of think-aloud protocols and observation is that there are no processing-reporting interval effects in think-aloud protocols and observation. Readers can report their thoughts while reading simultaneously, while researchers can observe the reading behavior quietly without any interference. Therefore, think-aloud protocols and observation can provide direct information about the readers' cognition during the reading process.

76

Since the use of strategies are related to learners' proficiency, this study takes into account the influence of learners' language proficiency while examining the use of reading strategies in two contexts: SA and AH.

Research Questions

The above literature review shows the need for cross-sectional research that compares L2 reading in the SA context with that in the AH context. Therefore, this study attempted to find the relation between language proficiency, reading development, and learning contexts through careful investigation of potential connections between contact with language and culture outside of class and reading strategy development.

To reach the aim of exploring reading proficiency and reading strategy development among learners of Chinese in two different contexts, the following questions were addressed:

1. Are there any differences in overall language proficiency between SA and AH students across the three different language proficiency levels?
2. Are there any differences in reading proficiency between SA and AH students across the three different language proficiency levels?
3. Are there any differences in the use of reading strategies between SA and AH students across the three different language proficiency levels?

Methodology

Participants

The participants were 73 learners of Chinese between the ages of 19-22 from a prestigious university in the US. Their Chinese language proficiency was at three different levels: beginning, intermediate and advanced. At the each proficiency level, students were split into two groups to study Chinese either in China or in the US. The beginners had no or almost no previous experience of learning Chinese before. They would take Chinese 101 either in China or in the US. The intermediate learners had finished Chinese 102 at the university in the previous semester, and would study Chinese 201 either in Chinese during the summer, or at their home university in the US during the following fall semester. The advanced students were similar to the intermediate students. They had finished Chinese 202 in the US, and would proceed to Chinese 301 either in China or in the US. The demographic information for the students is shown in Table 1. The SA students spent eight weeks in an intensive language and culture program in China as the equivalent of the regular course offered by their university in the US during the fall semester. When they came back, they needed to pass a proficiency test, which would find out whether they had made satisfactory progress for the transfer of their credits back to the US. If they passed the proficiency test, they would be placed into the next level of the Chinese classes offered by their university in the US. All participants were native English speakers. The AH students were in three classes randomly chosen from the multiple sections of each proficiency level at the university. Students with previous SA experience in any Chinese-speaking regions were excluded from the AH groups. Heritage students of Chinese were also excluded from all the groups.

Table 1. SA and AH group demographic information

Group demographics	SA	AH
Number of students at the beginning level	9	15
Number of students at the intermediate level	15	13
Number of students at the advanced level	11	10
Average age	20.7	21.4

Instruments

A holistic understanding of reading should include both the reading process and its results. Therefore, both quantitative and qualitative measures were used. On the one hand, quantitative measures such as reading comprehension tests and strategy questionnaires were used to accurately assess reading comprehension—the product of reading. On the other hand, qualitative methods such as think-aloud protocols, observation, and interviews were used to look into the learners' reading development process.

Chinese language proficiency test: The tests were chosen from a test pool collaboratively designed by the instructors at the university. The purpose of the test pool was to assess the students' overall proficiency of Chinese after they finished a certain level of class. These tests had been used as the final tests for Chinese language classes at different levels. There were six proficiency levels: Beginning 1 for Chinese 101, Beginning 2 for Chinese 102, Intermediate 1 for Chinese 201, Intermediate 2 for Chinese 202, Advanced 1 for Chinese 301 and Advanced 2 for Chinese 302. Passing the test indicated successful completion of a semester's learning at a certain level at the university or its equivalent. Each of these tests included five sections: listening, grammar, reading, translation and writing. The highest possible score for each test was 100 points.

Reading comprehension test: The quantitative measure used in the study was reading comprehension tests developed from part of the reading section of simulated HSK (Chinese Proficiency Test for foreign learners of Chinese). There were seven levels in HSK. Level 1-2 was chosen as the post-test for beginners. Level 1-3 was used as the pre-test and Level 2-1 as the post-test for intermediate learners. Level 2-2 was used as the pre-test and Level 3-1 as the post-test for advanced leaners. Before being administered to the students, the tests were evaluated and approved as fitting into students' current proficiency level by three instructors who had been teaching Chinese for five years at the university. The test consisted of two passages, followed by five multiple-choice questions for each. The highest possible score was ten points. The test items were designed to test readers' understanding of general ideas, detail information and inference ability respectively. The time limit for this part was 15 minutes.

Think-aloud: There was one level-appropriate passage in Chinese for the think-aloud protocol at each proficiency level. No time limit was set for the think-aloud. When the students were ready, they could begin think-aloud in any language they felt comfortable with.

Observation: Observation of participants' reading behavior was conducted while the participants were doing the reading comprehension test and think-aloud. Notes were taken down as records.

Reading strategies questionnaire: Strategies approach to reading was assessed by a self-report

questionnaire developed by Taillefer (2005). The questionnaire was designed to offer an immediate retrospective image of the readers' text processing. Students answered "yes" or "no" to questions covering five categories of general reading strategies, calling on their conscious (or unconscious) use of behavioral moves and their thinking about meaning and about themselves as readers. These categories correspond to the following "higher-level syntactic and semantic process" (Nassaji, 2003, cited from Taillefer, 2005): textual content (Q1-7), reader response (Q8-10), concrete techniques (Q11-17), task perception (Q18-21), and state of mind while reading (Q22-24). Eleven questions (Q25-35) then explored more language-oriented local problem-solving techniques, or "lower-level word recognition and graphophonic processes" (Nassaji, 2003, cited from Taillefer, 2005).

Semi-constructive retrospective interview: A semi-constructive retrospective interview was conducted with each participant after they finished the think-aloud. The interview focused on their reading strategies and reading development in terms of the context.

The first two questions of the semi-constructive interview were the same for both the AH and SA students:

1. What do you think of the test you did just now? Was it difficult?
2. Among the four skills: speaking, listening, reading and writing, which do you think has improved most during your study the past two semesters? Which do you think has improved least? Were there any reasons for such a situation?

The next three questions for AH students were:

1. Have you tried deliberately to improve your reading proficiency? Did you try any specific strategies to improve your reading proficiency the last two semesters?
2. What reading strategies do you use for Chinese reading? Are they different from reading in other languages?
3. Was your study in the last two semesters beneficial or even crucial to your reading development?

The next three questions were for SA students:

1. How did your stay in China influence your Chinese reading proficiency, reading habits and reading strategies?
2. What reading strategies do you use for Chinese reading? Are they different from reading in other languages? What specific reading strategies did you develop during your study in China?
3. Was your study in China beneficial to your reading in the long term? What do you think of your study in China in terms of reading development?

Language Contact Profile (LCP) and Reading Language Contact Profile (RLCP): For decades, researchers have formulated various questionnaires and interview approaches in order to gain insight into the quantity and quality of target language use and contact. A good example of it is the Language Contact Profile (LCP), which has been used and adapted by various researchers. It is "a multifaceted questionnaire that examines various aspects of a student's language history and use" (Segalowitz & Freed, 2004 p. 179). LCP is a questionnaire formulated to assess second language contact for students in

language programs to reveal details of various learning contexts including study abroad programs. In 1990, Freed "refined and expanded what she came to call the Language Contact Profile (LCP) in her first project on the effects of language learning in a study abroad context" (Freed, Dewey, Segalowitz & Halter, 2004, p. 350).

In this study, the participants' contact with language and culture was measured using the Language Contact Profile (Freed, 1995b, Freed, Dewey, Segalowitz & Halter, 2004), which could provide invaluable information concerning the kinds of activities students choose to engage in during SA. The Reading Language Contact Profile (RLCP) (Dewey, 2002) was used to measure the students' language contacts in terms of reading. The RLCP is a self-report of language use focusing largely on reading activities and assessment of the amount of time spent outside of class. The RLCP was adapted to include the following categories: (a) reading time (b) interacting with others (c) learning about Chinese Culture.

Procedure

A pre-test and a post-test for both language proficiency and reading proficiency were administered, except to the beginning level group, which did not have previous learning experience. For intermediate and advanced learners, the scores of their final examination at the end of the previous semester were taken as their pre-test scores. The SA students would need to take a proficiency test in order to transfer the credits after their return. The proficiency tests were the same as the final test used for the AH students after they finished the fall semester. Therefore, the scores of the proficiency tests from the SA students and the final examinations of the fall semester from the AH students were used as the post-test scores.

A similar procedure was applied to the administration of the reading tests. Among the SA groups, the intermediate and advanced students took a pre-test one week before they left for China. All three SA groups took the post-test during the first week of the fall semester. Among the AH students, the intermediate and advanced students took the pre-test in the first week of the fall semester. All three AH groups took the post-tests in the last week of the fall semester.

A think-aloud protocol followed the reading post-tests. The participants verbally reported their cognitive processes while reading. They received the think-aloud instruction and then conducted a short practice section. During the training session, the participants were first shown a brief demonstration of the technique and then they were given a practice passage. The practice passage was different from the test passage and much shorter. The participants produced verbal reports of their thinking in any language they felt comfortable with during the reading process. They were instructed to verbalize whatever they were thinking as their thoughts naturally came to mind while reading, such as comments on language, content, task or whatever they associated with them. After the introduction of the technique and some practice, the students received the passage for reading. When they were ready, they made a signal and began to think aloud while reading. The whole process was carefully observed. Notes and videos were taken for recording the students' reading behaviors.

After the think-aloud, a semi-constructive retrospective interview was conducted. The focus of the interview was to find out which strategies the students used in reading, how they acquired those strategies, how their reading in L2 developed, and the impact of learning context on their reading development.

Data analysis

Coding was conducted after each test and interview. Language proficiency tests and reading comprehension tests were scored.

Since the use of strategies was a main concern of this study, the coding schemes of the think-aloud were generated based on the strategy questionnaire. In other words, we used the strategy-type coding schemes for protocol analyses. There were two categories of strategies for coding: general strategies and local strategies. When any of the strategies was identified in the think-aloud, a point was added to the appropriate category.

The reading strategies questionnaires were scored. Students' *yes* or *no* answers were transformed into positive or null values.

Interviews were examined to find out more about the students' reading strategies and development and their relationship to different learning contexts.

Each item in the LCP and RLCP was also assessed. The amount of time spent outside of class was calculated in each of the following categories: (a) reading time; (b) interacting with others in Chinese; and (c) learning about Chinese. The overall reading time was calculated in hours. Interaction in Chinese and culture learning was assessed by a 4-point scale, with 0 being *Never*, 1 being *Seldom*, 2 being *Sometimes*, 3 being *Often*, and 4 being *Very Often*. Scores for each category were calculated.

Results

Chinese proficiency:

Table 2 provides the descriptive statistics of the comprehensive language proficiency pre-test scores. The SA and AH groups were very similar at both the intermediate and advanced levels. In order to find out whether there was a significant difference between the SA group and AH group across the two proficiency levels, a 2x2 between-subject analysis of variance (ANOVA) was performed on the test scores as a function of proficiency levels (intermediate vs. advanced) and learning contexts (SA vs. AH). The results show that there was no significant difference, $F(2, 47) = 45.893$, $p = 3.986$.

Table 2. Mean scores of language proficiency pre-test

	Intermediate	Advanced
SA		
M	86.63	85.33
SD	(9.96)	(11.31)
AH		
M	86.14	85.92
SD	(12.16)	(10.07)

Table 3 shows the scores of the comprehensive language proficiency post-test. Overall, the SA groups were slightly higher than the AH groups. A 2x3 ANOVA was run in a way similar to what was done in the pre-tests. The independent variable was the post-test scores, and the dependent variables were learning contexts and proficiency levels. The results indicate that overall, the scores of the SA groups were significantly higher than the AH groups, $F(3, 69) = 8.781$, $p = .004$. In order to find out the pattern of

differences on the post-test scores between the two learning contexts across the different proficiency levels, simple main effect of contexts was performed for each proficiency level. There was no significant difference on the scores between the SA and AH groups among beginners, $F(1, 24) = 2.342$, $p = .101$. The scores of the SA groups were significantly higher than the AH groups at the intermediate and advanced level respectively, $F(1, 27) = 31.634$, $p < .001$; $F(1, 20) = 12.804$, $p = .003$.

Table 3. Mean score of language proficiency post-test

	Beginning	Intermediate	Advanced
SA			
M	93.89	90.20	88.23
SD	(14.19)	(21.45)	(12.12)
AH			
M	93.14	85.88	85.71
SD	(11.57)	(13.24)	(14.10)

Reading

Table 4 shows the mean scores of the reading pre-test. The scores indicate that the SA and AH groups were very similar at both the intermediate and advanced levels. In order to find out whether there was a significant difference between the SA group and AH group across the two proficiency levels, we ran a 2x2 ANOVA on the test scores as a function of the proficiency levels (intermediate and advanced) and the learning contexts (SA and AH). The results show that there was no significant difference, $F(2, 47) = 14.201$, $p = 5. 387$.

Table 4. Mean score of reading pre-test

	Intermediate	Advanced
SA		
M	9.51	9.12
SD	(19.11)	(14.09)
AH		
M	9.39	9.19
SD	(10.50)	(13.77)

Table 5 shows the mean scores of the post-test on reading. Through a 2x3 ANOVA, we found that the scores of the SA groups were significantly higher than the AH groups, $F(3, 69) = 11.137$, $p = .003$. In order to find out the pattern of differences on the post-test scores among the different proficiency levels between the two learning contexts, simple main effect of contexts was performed for each proficiency level. There was no significant difference on the scores between the two learning contexts for the beginners, $F(1, 24) = 64.238$, p = 42.093. The scores of the SA groups were significantly higher than the AH groups among the intermediate students, $F(1, 27) = 41.382$, $p <. 001$. There was no significant difference among the advanced learners, $F(1, 20) = 12.804$, $p = 2. 103$.

Table 5. Mean score of reading post-test

	Beginning	Intermediate	Advanced
SA			
M	8.70	9.32	9.01
SD	(14.19)	(21.45)	(12.12)
AH			
M	8.61	8.58	8.79
SD	(17.57)	(13.24)	(14.10)

Think-aloud

Table 6 displays the mean scores of the think-aloud protocol. At the intermediate and advanced level, the SA groups scored higher than the AH groups. A 2x3 ANOVA was performed on the think-aloud scores as the function of learning contexts and proficiency levels. The scores of the SA groups were significantly higher than the AH groups, $F(3, 69) = 11.824$, p < .001. In order to find out the pattern of differences on the scores among different proficiency levels between the two learning contexts, simple main effect of contexts was performed for each proficiency level. There was no significant difference on the scores between the learning contexts among the beginners, $F(1, 24) = 82.051$, p = 14.932. The scores of the SA groups were significantly higher than the AH groups at the intermediate and advanced level respectively, $F(1, 27) = 6.723$, $p < .001$; $F(1, 20) = 8.832$, $p < .001$.

Table 6. Mean scores of think-aloud

Category	Beginner	Intermediate	Advanced
SA			
M	5.42	9.11	12.18
SD	(20.6)	(19.2)	(22.3)
AH			
M	5.60	7.77	10.60
SD	(13.7)	(12.0)	(15.6)

Reading strategy questionnaire

Table 7 shows the mean scores of the self-reported reading strategies of the groups. The scores of the groups between the two learning contexts were close at the beginning and advanced levels. However, there was a big gap at the intermediate level.

Table 7. Mean scores of the reading strategy questionnaire

	Beginning	Intermediate	Advanced
SA			
M	11.67	23.82	27.24
SD	(23.87)	(25.74)	(15.44)
AH			
M	12.48	18.13	26.57
SD	(12.09)	(17.25)	(18.52)

To find out whether there was a significant difference between the SA and AH context, a 2x3 ANOVA was performed on the strategy scores as a function of contexts and proficiency levels. The results indicated that the SA groups were significantly higher than the AH groups, $F(3, 69) = 6.462$, $p = .002$. In order to find out the pattern of differences on the scores among the different proficiency levels between the two learning contexts, simple main effect of contexts was performed for each proficiency level. There was no significant difference on the scores between the learning contexts among the beginners, $F(1, 24) = 31.095$, $p = 14.171$. The scores of the SA groups were significantly higher than the AH groups among the intermediate and advanced groups respectively, $F(1, 27) = 4.483$, $p< .001$, $F(1, 20) = 5.092$, $p = .011$. There was no significant difference among the advanced learners, $F(1, 20) = 37.482$, $p =9.394$.

Observation

Observation of the students' reading procedures showed no difference between the SA and AH groups. All the students were found to make use of the following methods to facilitate comprehension: (1) underlying key words or unknown words, and (2) reading some parts of the texts again if time was allowed. However, the SA groups showed a stronger tendency to communicate with the test supervisor for the explanation for the unknown words, although it was not stated in the test instructions that they could seek such help. Twenty-three students of the SA group asked the supervisor to explain the unknown words, whereas only two students in the AH group did so.

The SA students also showed a stronger tendency to use their smart phones to consult online dictionaries for the new vocabulary. Seven SA students asked the test supervisor whether they could use the apps in their phones, whereas no one in the AH group made the same request.

Interview

In the semi-constructive interview after the test and the think-aloud, all the students responded that the reading materials were difficult. The students who scored the lowest points in each group confessed that they could not make much sense of the reading, since there were too many unknown words for them. The AH students reported that reading skills were improved most in their studies, compared with listening, speaking and writing. They reported that they had deliberately spent a large amount of time developing reading proficiency, since they did not have many opportunities to practice speaking and listening outside of the classroom. However, the methods they relied on for reading development were simply reading the textbooks and memorizing the characters. Occasionally they listened to Chinese songs and read the lyrics and subtitles. As to the SA students, they reported that reading was the second least developed skill among the four skills, only ahead of writing. They confessed they did not pay much attention to reading, since they believed that reading could be improved in the AH situation. Therefore, they devoted most of their spare time to listening and speaking, which had been significantly enhanced after the eight-week study abroad period. They were much assured of their listening and speaking abilities but quite uncertain about their reading proficiency. They considered that the study abroad experience did help them in reading, but not as much as in listening and speaking.

Many of the students in both contexts reported using online tools for reading. All of the SA students had at least one app in their phone for vocabulary and translation. They reported that they used the apps and the Internet quite frequently, especially in China. On the other hand, only one third of the AH students downloaded and used at least one app to improve their reading.

Language Contact

Tables 8-10 present the quantitative data summarized from LCP and RLCP for measures of language and culture contacts.

Table 8 shows the data from the beginning learners. There were some differences between the SA and AH groups. In terms of reading time (textbook and test time excluded), the AH group spent less than one hour reading, whereas the SA group spent about one hour on average reading. With regard to the learning of the target culture, the AH group scaled themselves at 0.75 on average, which is between

Never and *Seldom*. The SA group, however, scaled themselves at 1.89 on average, between *Seldom* and *Sometimes*. As to the interaction in Chinese, the AH group reported a point of 0.25, which is rather low compared with the SA group's report of 1.33. With regard to the reading on texts and tests, the AH students spent 45 hours on textbooks, and 16 hours on tests. The SA students spent about 40 hours on textbooks and 12 hours on tests and quizzes.

Table 8. Descriptive statistics for measure of language and culture contact for AH and SA groups at the beginning level

Category	AH	SA
Total class hours	60 hours	150 hours
Reading Chinese (except textbooks)	<1 hour	1 hour
Learning about the culture	0.75	1.89
Interacting with others in Chinese	0.25	1.33
Reading textbooks	45 hours	40 hour
Reading tests/quizzes	16 hours	12 hours

Table 9 shows the data for the intermediate learners. In terms of reading time, the AH group spent an average of 1.4 hours reading materials, whereas the SA group spent 20 hours on average reading. With regard to the learning of the target culture, the AH group scaled themselves at 0.69 on average, which is between *Never* and *Seldom* The SA group, however, scaled themselves at 2.85 on average, between *Sometimes* and *Often*. As to the interaction in Chinese, the AH group reported a point of 0.26, which is rather low compared with the SA group's report of 3.03 between *Often* and *Highly Frequently*. With regard to the reading on texts and tests, the AH students spent 45 hours on textbooks and 16 hours on tests. The SA students spent only 40 hours on textbooks and 12 hours for tests.

Table 9. Descriptive statistics for measure of language and culture contact for AH and SA groups at the intermediate level

Category	AH	SA
Total class hours	60 hours	150 hours
Reading Chinese (except textbooks)	1.4 hours	20 hours
Learning about the culture	0.69	2.85
Interacting with others in Chinese	0.26	3.03
Reading textbooks	45 hours	40 hour
Reading tests/quizzes	16 hours	12 hours

Table 10 shows the data for the advanced learners. In terms of reading time (textbook-reading time and test-taking time excluded), the AH group spent an average of 1.6 hours, whereas the SA group spent 25 hours on average. With regard to the learning of the target culture, the AH group scaled themselves at 0.71 on average, which is between *Never* and *Seldom*. The SA group, however, scaled themselves at 3.29 on average between *Often* and *Highly Frequently*. As to the interaction in Chinese, the AH group reported a point of 0.31, which is rather low compared with the SA group's report of 3.34. With regard to the reading on texts and tests, the AH student spent 60 hours on textbooks and 16 hours on tests. The SA student spent about 45 hours on textbooks and 12 hours on tests and quizzes.

Table 10. Descriptive statistics for measure of language and culture contact for AH and SA groups at the advanced level

Category	AH	SA
Total class hours	60 hours	150 hours

Reading Chinese (except textbooks)	1.6 hours	20 hours
Learning about the culture	0.71	3.29
Interacting with others in Chinese	0.31	3.34
Reading textbooks	60 hours	45 hour
Reading tests/quizzes	16 hours	12 hours

Overall, the differences in language contacts between the SA and AH groups were huge only at the intermediate and advanced level, but not at the beginning level. Generally speaking, at the intermediate and advanced level, the SA students spent far more time on reading other materials that were not their textbooks. They also had far more opportunities for language and culture contacts. The LCP and RLCP revealed that most of the print input the SA students received were incidental and spontaneous messages from the environment. For example, they often read text messages in the cell phones, advertisements through various channels and media, notices on the university webpage and menus in school canteens and restaurants outside. In this sense, the SA setting provided opportunities not only for oral communication, but also for the improvement of literacy skills.

Discussion

The main purposes of the study were to answer the following questions: (1) Are there differences in overall language proficiency between the SA and AH students across three different language proficiency levels? (2) Are there differences in reading proficiency between the SA and AH students across three different language proficiency levels? (3) Are there differences in the use of reading strategies between the SA and AH students across three different language proficiency levels?

To the first research question, we found that overall, the language proficiency of the SA groups was significantly higher than the AH groups across the three different proficiency levels. Closer examination revealed that there was no significant difference among the beginners. However, the advantage of the SA context over the AH context was significant among the intermediate and advanced students. This is not consistent with some previous studies. Freed (1993, 1995a, 1998) and Coleman (1997b) found that the higher the students' initial level of proficiency, the lower the gains in proficiency as a result of studying abroad. A possible explanation for the different results found in this study is that the type of language may have an impact on the students' initial gain at different proficiency levels. In Freed and Coleman's studies, the learners' target languages were European languages, which were close to their first language. However, in this study, the English-speaking students learned a typologically different language—Chinese. Chinese is notoriously difficult for English speakers. Mandarin competence takes 2,200 class hours, with half of that time spent in a country where it's spoken, according to the U.S. State Department's Foreign Service Institute, whereas Spanish can be learned in 600 to 750 class hours. The extra difficulty of learning Chinese may hinder beginners from taking advantage of the SA context. For example, students may need more time to deal with two aspects of Mandarin Chinese which are particularly difficult for beginners: tones in oral communication and characters in literacy acquisition. Students taking Chinese 101 usually need to spend at least 6-8 weeks in a regular semester to get a sense of how tones and characters work. The instructors' guidance is essential at this stage. In an SA context, although beginners have more opportunities for language and culture contacts outside of class, the linguistic information they receive from the outside is far beyond their comprehension. Therefore, it is

impossible for them to digest and absorb the rich input. The intermediate students have already overcome the tone and character barriers to some extent. They can converse to local people with some simple sentences, and understand many texts outside of the classroom. Thus the natural language contact outside of class greatly benefits the SA group of students at the intermediate level. The SA context also brought significant benefits to the advanced students, although the benefit may not appear as obvious as for the intermediate students. This may be due to the fact that the reading and writing sections of the tests for the advanced learners focus primarily on the academic and literature content. The SA advanced students may have made great progress in listening and speaking as well. However, the language and culture contact outside of the classroom may not help them much in the test, because L2 academic and literary reading was mostly learned through formal instruction and classroom discussion. Overall, the developmental pattern of the intermediate and the advanced level was consistent with Freed and Coleman's conclusion: less competent learners benefit more than more competent learners in the SA context.

We found huge differences in language contacts between the SA and AH at the intermediate and advanced level, but not much difference at the beginning level. The SA students at the intermediate and advanced level spent far more time on reading materials other than textbooks. They also had far more opportunities for language and culture contacts. However, the beginners were not able to make best of the opportunities. Generally speaking, Chinese was especially hard for English speakers due to the typological distance between the two languages. Usually it takes longer time for the beginners to embark on the journey of learning. The language input, verbal or in print, may be far beyond the beginners' comprehension level. It was still too tough for beginners to make sense of the language input, let alone produce appropriate responses. Meaningful language contacts outside of the classroom may be minimal. Therefore, the benefits of the SA context may not work for the beginners.

As to the second and third research questions, we found beneficial impact of the SA context on reading proficiency and the use of reading strategies. In this study, we viewed reading as both a product and a process. The reading tests provided us with the information about the learners' reading as a product of comprehension. It is found that there was a significant advantage of the SA context on reading comprehension at the intermediate level. There was no statistically significant difference in learning context among the beginners or the advanced learners. The combined effects of context and proficiency may account for these results. Based on the data from LCP and RLCP, the SA students had more opportunities to contact the target language in various reading activities, such as reading notices, sending text messages in Chinese, and emailing Chinese friends, which compensated for the lack of reading time on textbooks to a large extent. The reading scopes of the two groups were also different. The AH students focused primarily on textbooks, whereas the SA students read a wide range of authentic materials for real communicative purposes. These language contacts and reading activities were a good match for the intermediate students' reading proficiency. Therefore, it was no surprise to see that the intermediate students significantly benefited from the language contacts in the SA context. However, reading notices or sending text messages was still too difficult for the beginners who were struggling with Pinyin (the Chinese phonetic system) or characters. Meanwhile, the casual literacy contacts outside of the classroom were too easy for the advanced students who started to deal with short academic writing or literary essays.

The data from RLCP showed that the AH students spent more time in reading textbooks and test

papers. They also considered that reading had been improved much more than other skills in their studies. The SA students, compared with the AH students who read more in class, had much less time for self-reading during their 8-week intensive training in China. As a result, the AH learners' performance was as good as their SA counterparts' at the beginning and advanced levels.

Viewing reading as a process, we examined the comprehension process in details. The think-aloud protocol and the strategy questionnaires showed significant differences between the contexts at the intermediate level. This means that although learners in both learning contexts have made reasonable progress in reading, the SA group at the intermediate level benefitted most from the learning context.

In the interviews, we found that AH students had more confidence in reading, contrary to Dewey's (2004) findings in which the SA students felt more confident of reading in L2. The SA students in the present study did not report much confidence in the semi-constructive interview. On the contrary, the AH students reported that they felt that their skills in reading was better than other skills. The lack of confidence among the SA students may be due to the relatively small proportion of time the SA students spent on reading during their study in China. However, objective measurement did not support their evaluation of themselves. At the beginning and advanced level, the SA students performed equally well as the AH students, in spite of their low confidence. In fact, the SA students even significantly outperformed the AH students at the intermediate level. From the RLCP of the SA and AH groups, we found that SA students had large quantities of contact with Chinese in various ways, which may improve their reading ability. A variety of exposure to the target culture also provided the students with an increased background knowledge, which would lead to better top-down processes. (Koda, 2005).

When it comes to the use of reading strategies, there was no significant difference for the two contexts at the beginning level in either the think-aloud or questionnaire. This may imply that due to the difficulty of literacy learning in Chinese, the beginners of Chinese from an alphabet literacy background had to spend a prolonged period of time to understand the basics of Chinese characters in the classroom. Before they could acquire basic literacy skills in the classroom, the naturalistic learning environment outside of the class provided by an SA context may not benefit them much. Most of the beginners still struggled at the very first step of the literacy learning—word recognition at this level. There was very limited room for them to develop complicated strategy use in reading passages.

Both the think-aloud protocol and the questionnaires show that the SA students at the intermediate level employed significantly more strategies than their AH counterparts. This may mean there is a threshold for learners of Chinese. After learners acquire certain amount of literacy knowledge, they will be able to take the advantage of the study-abroad context, and get benefit from various language contacts with native speakers. In addition, the literacy environment outside of the class provided by the study-abroad context is a good match for intermediate students. For example, students can exchange text messages with their newly-made Chinese friends. Such an activity may be too difficult for beginners, and too easy for advanced students, but well-suited for intermediate learners. The environment outside of the classroom is full of literacy materials suitable for intermediate learners, such as various signs, billboard notices, and advertisements.

The SA students at the advanced level also showed a slight but not statistically significant advantage in strategy use in the think-aloud and questionnaires. This may imply that the benefit of the SA context on reading may diminish as the learners' progress continues to the advanced level. Advanced learners

focus on more academic or literary works. Classroom instruction and discussion play a bigger role for the progress in these areas.

To sum up, context did not play a role for beginners on either overall language proficiency, nor on reading development. However, the SA context was beneficial for intermediate learners on both language proficiency and reading. For the advanced learners, the promotional effects of the SA context only occurred in overall language proficiency test, but not in reading. There seemed a threshold effect for beginners in both general language skills and in reading specifically, and a plateau effect for advanced learners in reading specifically.

The findings are pedagogically informative. We know from the results that the SA context does play a beneficial role in learning, but learners' language proficiency interferes. To make the best of the SA context, it is suggested that learners acquire a certain amount of knowledge of Chinese before they leave. The SA curriculum or program designers in China should also prepare more level appropriate language contacts for beginners and advanced learners, so that they can overcome the threshold effect and plateau effect in learning respectively.

Conclusion

This study made use of a variety of instruments to collect both quantitative and qualitative data so that we can gain a better understanding of how reading develops and of what factors influence that development in AH and SA contexts. Multiple objective measures, plus some qualitative methods, such as observation and research interviews, were employed for investigation. Together, these methods provided a clearer picture of how contexts influence reading development of learners of Chinese at different proficiency levels.

With these multiple instruments, we found that there was a significant difference on reading proficiency and use of reading strategies between the AH group and SA group at the intermediate level, although the differences were less pronounced at the beginning and advanced level. We also found that the AH students spent more time in classroom learning, whereas the SA students at the intermediate and advanced level had more exposure to reading materials and culture, which compensated for their deficiency in classroom learning. However, the beginners seemed unable to make best of the SA context due to their limited language ability.

The findings may provide some useful information for designing SA programs in China or other areas where the target language is typologically different from the learners' first language.

References

Baker, L. & Brown, A. (1984). Metacognitive skills and reading. In D. Pearson (Ed.), *Handbook of Reading Research.* (pp. 353-394). New York: Longman.

Brecht, R., & Robinson, J. (1995). On the value of formal instruction in study abroad: Student reactions in context. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 317-334). Philadelphia: John Benjamins.

Brown, A. (1981). Metacognition in reading and writing: The development and facilitation of selective attention strategies for learning from texts. In M. Kamil (Ed.), *Directions in reading research and*

instruction. (pp. 21-43). Washington D.C. National Reading Conference.

Brown, A., Armbruster, B. & Baker, L. (1986). The role of metacognition in reading and studying . In J. Orsany (Ed.) *Reading comprehension: From research to practice.* (pp. 49-75). Hillsdale, NJ : Lawrence Erlbaun.

Carrol, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals, 1*, 131-151.

Coleman, J. (1997a). Supporting language students during their year abroad. *The Linguist*, 36, 2-5.

Coleman, J. (1997b). Residence abroad within language study, *Language Teaching*, *30*, 1-20.

Dewey, D. (2002). *The effects of study context and environment in the acquisition of reading by students of Japanese as a second language during study-abroad and intensive domestic immersion.* Unpublished doctoral dissertation. Carnegie Mellon University, Pittsburgh, PA.

Dewey, D. P. (2004). A comparison of reading development by learners of Japanese in intensive domestic immersion and study abroad contexts. *Studies in Second Language Acquisition*, 26, 303–328.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Freed, B. F. (1990). Language learning in a study abroad context: The effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency. In J. Alatics. (Ed.), *Linguistics, language teaching and language acquisition: The interdependence of theory, practice and research* (Georgetown Round Table on Linguistics). Washington DC: Georgetown University Press.

Freed, B. 1993. Assessing the linguistic impact of study abroad: What we currently know - what we need to learn. *Journal of Asian Pacific Communication 4*:4, 151-166.

Freed, B. F. (1995a). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Philadelphia: John Benjamin's.

Freed, B. F. (1995b). Language learning and study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp.3-33). Amsterdam: Benjamins.

Freed, B. F. (1998). An overview of issue and research in language learning in a study abroad setting. *Frontiers: the Interdisciplinary Journal of Study Abroad, 4*, 31-62.

Freed, B. Dewey, D. Segalowitz, N. & R. Halter. (2004). The language contact profile. *Studies in Second Language Acquisition*, 26, 2, 349-356.

Garner, R. (1987). Metacognition and reading comprehension. Norwood, NJ: Ablex.

Huebner, T. (19950. The Effects of Overseas Language Programs: Report on a Case Study of an Intensive Japanese Course. In B.F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context.* Amsterdam and Philadelphia: John Benjamins.

Kline, R. R. (1998). Literacy and language learning in a study abroad context. *Frontiers, the Interdisciplinary Journal of Study Abroad, 4*, 139–166.

Koda, K. (2005). *Insight into second language reading: A cross-linguistic approach.* Cambridge: Cambridge University Press.

Lafford, B. (1995). Getting into, through and out of a survival situation: A comparison of communicative strategies used by students studying Spanish abroad and at home. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp.123-148). Philadelphia: John Benjamin's.

Lafford, B. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition 26*, 201-225.

Lapkin, S., D. Hart, and M. Swain. 1995. A Canadian Interprovincial Exchange: Evaluating the

Linguistic Impact of a Three-Month Stay in Quebec. In B.F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*. Amsterdam and Philadelphia: John Benjamin's.

Oxford, R.L., & Crookall, D. (1989). Research on language learning strategies: Methods, findings, and instructional issues. *Modern Language Journal*, 73, 404-419.

Palinscar, A. and Brown, A. (1984). Reciprocal teaching: A means to a meaningful end. In J. Osborn and P.T. Wilson (Eds.), *Reading education: Foundations for a literate America*, 310. Cambridge, Massachusetts: Lexington.

Pressley, M. (2002). Comprehension strategy instruction: A turn-of-the-century status report. In C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices*. (pp. 11-27). New York: Gulford Press.

Segolowits, N., & Freed, B. (2004). Context, contact, and cognition in oral proficiency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173-199.

Serrano, S. L. (2010). Learning languages in study abroad and at home contexts: A critical review of comparative studies. *Porta Linguarum, 13*, 149-163.

Singhal, M (2001). Reading proficiency, reading strategies, metacognitive awareness and L2 readers. *The Reading Matrix*, 1(1).

Taillefer, G. (2005). Foreign language reading and study abroad: Cross-culture and cross-linguistic questions. *Modern Language Journal*, *89*(4), 503-528.

Urquhart, S. & Weir, C. (1998). *Reading in a second language: Process, product and practice*. New York: Longman.

Waxman, H.C & Padron, Y. (1987). The effect of ESL students' perceptions of their cognitive strategies on reading achievement. Presented at *the Annual Meeting of the Southwest Educational Research Association*, Dallas.

Appendix: Strategy Questionnaire

During reading, did you:
1. Consciously link information in one sentence with information from the preceding one?
2. "Guess" what was coming?
3. At any point correct or change an idea formed earlier in your reading?
4. Keep ideas in your head while reading?
5. Differentiate important points from details?
6. Identify any organization of ideas?
7. Learn something new?
8. React intellectually to information in the text?
9. Interpret the texts (make inferences, draw conclusions)?
10. React "emotionally" to ideas in the texts?
11. Try to push ahead when blocked by a comprehension difficulty, possibly going back to the problem later?
12. Try to remember specific parts of the texts?
13. Notice the style of the text?
14. Did you read the text more than once?
15. Reread parts of the text?
16. Reformulate parts of the texts?

17. Check or evaluate your comprehension?
18. Realize that it was necessary to know the pronunciation of each word to understand the texts?
19. Feel it was necessary to understand every word?
20. Aim first for general understanding?
21. Feel you were an efficient reader?
22. How did you feel while reading?
23. If you had found these texts in an English newspaper, journal, or magazine, would you have read them?
24. Did they interest you?
25. Try to guess the meaning of the word or expression?
26. Skip the difficulty in question deciding that it wasn't very important?
27. Compare the word or expression with something similar in your native language?
28. Look for clues in the context?
29. Analyze a word in itself (prefix, root, and suffix)?
30. Grammatically analyze a difficulty within the sentence?
31. Consciously use punctuation, capitals?
32. Translate anything?
33. Want to use a bilingual dictionary?
34. A dictionary?
35. Pronounce the word or the expression?