Effective Learning Strategies for the Recognition of Characters and Words by Learners of Chinese with Varying Proficiency in Different Learning Environments

Jing Wang & Robert B. Harris
Indian University-Purdue University, Indianapolis

Abstract

The present study gauges effective strategies for recognizing characters and words by learners of Chinese as a foreign language. Besides answering background questions, 203 participants completed a questionnaire, a vocabulary recognition test, and a vocabulary checklist. The vocabulary test was found to be valid and reliable, and six categories of learning strategies were generated from the questionnaire. With the two background variables of language proficiency and language environment (studying abroad versus studying at home), the six strategy categories explained 48% of the variance in vocabulary recognition. Specifically, "frequent meaningful interaction with characters and words" positively predicted vocabulary recognition, whereas "focusing on character orthography" and "focusing on character pronunciation" negatively predicted it.

The individual strategy of typing characters positively predicted vocabulary recognition, whereas using Pinyin to help remember pronunciation negatively predicted it. Pinyin can facilitate vocabulary recognition through typing characters yet may hinder it if excessively relying on Pinyin for pronunciation.

1. Introduction

The present study aims to gauge the effective application of learning strategies for the recognition of characters and words by learners of Chinese as a foreign language (CFL). This study is of great

relevance to the CFL field because characters and words assume critical roles in the instruction and learning of the Chinese language.

In Chinese, characters are the smallest free-standing written units, yet they do not correspond to words in a straightforward manner. Characters are separated by spaces, while words are not. A character may or may not be a word, whereas a word may be composed of single or multiple characters. Because Chinese syntax rules are simple (Kalgren, 1949), characters and words play a significant role in conveying meaning (Hoosain, 1991).

Yet character and word recognition is extremely difficult for CFL learners who are native speakers of Roman alphabetic languages. Specifically, Ke (1998a) reported that in an earlier study, Everson had found that beginning CFL learners "could only…correctly identify the meaning of 58.0 percent of the characters they were taught" (p. 92). Everson and Ke (1997) found that even intermediate students encountered great vocabulary difficulties when reading authentic news.

Hence, the identification of the effective application of strategies for character and word learning will enable students to monitor their learning processes and support instructors in developing effective pedagogical methods. It is relevant to base the present study on the existing literature on background variables, strategies for learning characters and words, and the gauging of vocabulary knowledge.

2. Literature Review

A review of background variables reveals important ones in strategy application, and a review of strategies for learning characters and words clarifies our current knowledge in this field. In addition, a literature review regarding the measurement of vocabulary knowledge provides guidance in designing vocabulary test for this study.

2.1 Background Variables

Learning strategies are related to diverse background variables. Summarizing the existing research on the application of learning

strategies, Nambiar (2009) listed "variables like proficiency, learning environment, ethnicity, age, gender, learning styles, motivation, and beliefs" (p. 137). She concluded from the existing research that "the level of proficiency of a learner does influence learning strategy use while environment helps determine how a learner learns and what strategies are employed to enhance learning" (p. 144). Hence, these two background variables are crucial to strategy application.

Learning environment may be related to different circumstances. For example, Churchill and Dufon (2006) pointed out that "there are perhaps few contexts as potentially rich and complex as study abroad" (p. 1). They abbreviated studying abroad as SA and studying at home as AH.

2.2 Strategies for Learning Characters and Words

2.2.1 Frequently used strategies

Various researchers have studied CFL learners' frequently used strategies for character learning. McGinnis (1999) found at the beginning level, "students generally resort to more idiosyncratic methodology, ranging from the very creative (personally concocted stories) to the very mechanical (rote repetition)" (p. 161). Yin (2003) conducted a survey on character learning strategies with CFL learners at different proficiency levels. Yin found that participants ranked difficulties in the following order (from high to low): writing, remembering pronunciation, and getting the meaning of characters. In addition, he found several commonly used strategies. Specifically, in learning how to write characters, the most frequently used learning strategies were copying characters and learning components. In learning character pronunciation, the most common strategies were using Pinyin (Romanization to assist in pronunciation) and reading aloud. In learning character meaning, the most frequently used strategies involved learning characters in context and using the knowledge of radicals (Yin, 2003). Shen (2005) synthesized others' research and summarized that "rote memorization, graphic cues, context cues, and knowledge of radicals are all used in learning characters" (p. 53). Defining "orthographic-knowledge-based learning strategies" as using "the three aspects of radical knowledge,

graphemics, semantics, and phonetics, as cues to encode character," Shen concluded that CFL learners frequently used these strategies (p. 61).

Researchers have also explored the effects of background variables on strategy application. First, studies have shown that proficiency influences the application of learning strategies. Shen (2005) found that with the improvement of language proficiency, "students have acquired more orthographic knowledge and the application of this knowledge in acquiring new vocabulary becomes more automatic" (p. 60). Second, Jin (2006) found that, at the beginning level, the application of learning strategies was influenced by CFL learners' first languages.

2.2.2 Perceptions of effective learning strategies

Ke (1998b) studied students' perceptions of the effectiveness of character learning strategies and the effects of perception on character learning. Gathering data from CFL learners who had just completed one year of study, Ke found that these learners perceived the following practices to be effective: learning semantic and phonetic radicals, writing characters, and learning characters in words. He also found that some perceptions predicted vocabulary knowledge. As for language background, Ke (1998b) did not find major differences between groups of heritage and non-heritage learners in their perception of effective strategies on character learning.

Wang and Leland (2011) studied first-year CFL learners' perceptions of character learning. They found that learning characters in isolation, in context, and in social context were perceived to support character recognition in different ways.

Although research has mainly focused on strategies for learning characters and words, limited studies have directly targeted the effective application of strategies. The main barrier lies in the lack of a valid and reliable vocabulary test to gauge vocabulary knowledge among students with different backgrounds.

2.3 Vocabulary Knowledge

2.3.1 Evaluation of English vocabulary knowledge

In designing vocabulary tests for English words, it is important to consider frequency and context. Nation (1983) designed a test for English vocabulary in which "each section of the test consists of six words and three definitions" (p. 14). The choice of the tested words was guided by "the statistical distribution of vocabulary" (p. 17), while "the definitions in the test use words from a higher frequency level than the tested words" (p.15). Reed and Chapelle (2001) pointed out that "a vocabulary test should require learners to perform tasks under contextual constraints" (p. 1).

2.3.2 Evaluation of Chinese vocabulary knowledge

Researchers have designed their own tests to evaluate vocabulary knowledge for CFL learners. For vocabulary tests focusing on word recognition, researchers often have chosen learned words as tested words and asked students to provide meanings in their first language and to provide pronunciations in Pinyin. The tested words stood alone in studies by Ke (1998b) and by Shen and Ke (2007), while they were contextualized in sentences in a study by Shen (2003).

However, in order to test vocabulary knowledge for CFL learners with various native languages and with different learned vocabulary, a vocabulary test needs to be designed and then evaluated in terms of its validity and reliability.

2.3.3 Variables influencing vocabulary knowledge for CFL learners

Previous studies have analyzed the effects of background variables on vocabulary knowledge for CFL learners. First, language proficiency clearly influences vocabulary knowledge. Intermediate CFL learners still face vocabulary difficulties in the reading process (Everson & Ke, 1997; Lee-Thompson, 2008), whereas advanced learners are more competent in vocabulary knowledge (Everson & Ke, 1997). Second, in terms of learning environment, Hayden (1998)

found "a certain amount of gain in reading ability resulting from a study abroad experience" (p. 22).

Third, with regard to the language background of heritage and non-heritage learners, Ke (1998a) found that among first-year CFL learners, the two groups did not differ in their "performance on Chinese character recognition and production" (p. 91). Xiao (2006) studied heritage and non-heritage students in a "high beginning level" Chinese course (p. 49), and she found no significant differences between the two groups on "vocabulary quizzes, character writing, or reading comprehension…" (p. 52).

2.4 Research Question

Limited study has been carried out to measure effective strategies to learn characters and words by CFL learners. The present study addresses that gap. To simplify the research, the present study focuses on language proficiency and learning environment (i.e., SA and AH) as two background variables. In addition, since there is a significant difference between learners' abilities to recognize and produce characters (Ke, 1996), it is important to separate these two aspects of character learning and to focus on character recognition. The major research question is as follows: What are effective learning strategies for character and word recognition by CFL learners with varying proficiency levels in different learning environments? In order to identify generally effective learning strategies, this study includes students with varying proficiency levels studying in two environments: studying at home in the U.S. and studying abroad in China.

3. Materials and Methods

3.1 Instruments

3.1.1 Background informationz

Twelve background questions were designed to solicit background information, and they provided a general picture of the participants. The following background questions were directly related to the research question of this paper and are further analyzed in

subsequent sections. Self-reported language proficiency was elicited through questions on language proficiency and number of characters recognized. The students' current learning environments were elicited through the question, "In which university are you studying Chinese now?"

3.1.2 Questionnaire

A questionnaire was designed on the basis of findings from a study by Wang and Leland (2011). In particular, the questionnaire examined the learning of characters and words in isolation, in linguistic context (i.e., words, phrases, sentences, and paragraphs), and in social interaction with other people. Next, the questionnaire was modified based on a pilot study with three students. Then, the questionnaire was further modified based on feedback from several experts in the field of Chinese language instruction. Finally, the questionnaire incorporated some optimal methods for word learning (Liu, 2000). For example, Liu suggested the method of teaching words that are semantically related.

The questionnaire on character and word recognition consisted of 36 items. Please refer to Appendix A for the original background questions and the questionnaire.

Each questionnaire item used a Likert scale of five choices: "strongly agree," "agree," "neutral," "disagree," and "strongly disagree." Each of these items was assigned a score, with "strongly agree" equaling five points, "agree" equaling four points, "neutral" equaling three points, "disagree" equaling two points, and "strongly disagree" equaling one point. Hence scores less then three indicated that the item was not frequently used by participants.

3.1.3 Vocabulary test

Thirty multiple-choice questions were designed to test students' knowledge of characters and words. For each question, a statement was provided with one tested vocabulary item missing. Participants were required to choose the missing tested item from five choices: the correct answer, the three distracters (incorrect choices), and the last option "I do not know." The last option reduced the possibility

for students to randomly choose an answer if they did not know it. As to distracters, some sounded similar to the correct choice: 有[yǒu, have] and又[yòu, again]; some distracters looked similar: 比[bǐ, compare] and 此[cǐ, this]. The first ten questions were designed to test the vocabulary knowledge of beginners, the second ten questions to test that of intermediate learners, and the last ten questions to test that of advanced students. Please refer to Appendix B for the test.

The designing of the test items and distracters was guided to some extent by the wordlists from an official Chinese language proficiency test*: Hanyu Shuiping Kaoshi* (HSK). According to the HSK test outline (Testing center of Beijing Language University, 2006), there are four HSK Wordlists: Wordlist One has 1,033 words, Wordlist Two 2,018 words, Wordlist Three 2,202 words, and Wordlist Four 3,569 words. Wordlists One, Two, and Three target beginning and intermediate CFL learners, while Wordlist Four targets advanced learners. Many tested words and distracters for the first 20 questions were taken from HSK Wordlists One, Two, and Three, while some tested words and distracters for the last 10 questions were taken from HSK Wordlist Four.

The designing of the vocabulary test took contexts of language use into consideration. Generally speaking, beginning CFL learners use the language in daily communication, intermediate CFL learners read about extended aspects of daily topics and begin to tackle abstract concepts, and advanced learners handle the language in a variety of social, cultural, economic, and environmental contexts. Correspondingly, the first ten questions included everyday topics, such as self introduction, talking about family members, comparing day-to-day weather, estimating distances, going to hospitals, mailing letters, and giving directions; the second ten questions included sentences talking about majors, registering for courses, traveling, visiting museums, commenting on living standards, and talking about computers; and the last ten questions covered a wide range of topics on abstract issues.

For each multiple-choice question, the correct answer equaled one point, and the other choices equaled zero points. All the questions were totaled in the variable *test*.

3.1.4 Vocabulary checklist

Reed (1988) pointed out "a checklist test, then, has much to recommend it as a broad measure of vocabulary knowledge, especially if it incorporates a correction procedure for overrating" (p. 23). The present study used a checklist to ask students to gauge different aspects of vocabulary knowledge in order to reduce overrating. Specifically, for each word, students had four choices, with each having a different point value: "I know both the meaning and pronunciation of this word" (two points), "I only know the meaning of this word" (one point), "I only know the pronunciation of this word" (one point), and "I don't know this word" (zero points).

For the checklist of vocabulary knowledge, there were a total of 28 words with seven words chosen from each of the four HSK wordlists. Please refer to Appendix C for the checklist of vocabulary knowledge. The 28 words were quantified in the variable *checklist*.

3.1.5 Summary of all the instruments

The background questions, the questionnaire, the vocabulary test, and the checklist were put into a website with two additional open-ended questions: 1) "Do you have other comments concerning the recognition of Chinese characters and words?" and 2) "Have you taken HSK tests? If so, what level test have you taken and what is your score?" All together there were 108 questions, which will be referred to as the survey in the following paper. There were two versions of the survey, one with simplified characters and the other with traditional characters.

Nearly all the CFL learners from one university in the U.S. took a placement test developed by Brigham Young University. The total number of participants from that university was 68.

3.2 Participants

Altogether, 203 participants from two universities in the U.S. and one university in China completed the survey. The students studying at the two American universities were native English speakers (or they had native English proficiency) taking regular Chinese courses,

while the students studying at the Chinese university spoke different native languages and were enrolled in a regular semester.

In terms of Chinese language proficiency, participants reported their language proficiency as follows: 38% beginners, 46% intermediates, and 16% advanced learners. The percentage of participants at the beginning and intermediate levels closely corresponded to the self-reported number of characters they recognized: 35% recognized fewer than 500 characters, corresponding to beginners, while 44% recognized between 500 and 2,000 characters, corresponding to intermediates. Consequently, it is acceptable to use self-reported proficiency to indicate their language proficiency. As for the learning environment, 80% studied at home institutions, while 20% studied abroad in China.

Among the 203 participants, 184 did not experience a mismatch because they read the version of characters (either simplified or traditional) they were familiar with. However, nineteen participants were treated as mismatches because they mentioned in the comment box that they were familiar with one version of characters, yet read another version in the vocabulary test and the checklist. The data from the 203 participants were used in analyzing the questionnaire, while the data from the 184 participants were used in the analyses involving the vocabulary test and the checklist.

3.3 Data Analyses

3.3.1 Previous methods

Researchers such as Jafarpur (2002) and Golkar and Yamini (2007) correlated targeted tests with other established tests to determine their validity. As another validity check, they used "decision consistency" to see if the targeted tests successfully placed participants at the appropriate level (Jafarpur, 2002, p. 42). Golkar and Yamini (2007) also used t-tests to see if targeted tests differentiated aspects of vocabulary knowledge and different learners.

Shen (2005) used a principal component analysis (an exploratory factor analysis) "to determine the underlying factors characterizing the strategies used most frequently by learners of Chinese" (p. 56).

3.3.2 Analyses

The selection of various statistical methods in the present study was influenced by ones used in previous studies. The following statistical tests were used to analyze the data:

(1) ANOVA, correlational studies, decision consistency check, and Cronbach's Alpha gauged whether or not the vocabulary test was a valid and reliable instrument to evaluate students' vocabulary knowledge.

(2) A principal component analysis of the questionnaire provided the basis for different strategy categories.

(3) Multiple regression analyses determined how well the two background variables and strategies predicted vocabulary knowledge.

4. Results

4.1 Vocabulary Test

The following four sets of statistical analyses were conducted to measure the validity of the vocabulary test. First, a one-way ANOVA on *test* (the vocabulary test) used self-reported proficiency level as the independent variable. There was a significant effect of language proficiency on *test*, $F(2, 181) = 25.63$, $p = .000$. Moreover, there was a significant linear trend, $F(1, 181) = 40.10$, $p = .000$, indicating that as students' proficiency increased, their testing scores increased proportionally. Because there were different participants in each proficiency level, Hochberg's GT2 post hoc tests were conducted. Post hoc comparisons indicated that there were significant differences among participants at the three proficiency levels. Hence, the vocabulary test could be used to differentiate students with different proficiencies.

Second, a one-way ANOVA on *test* used the number of characters which students thought they could recognize as the independent variable. Since Levene's test was significant, Welch F

was reported. There was a significant effect of self-reported vocabulary size on *test*, Welch $F(4, 56.90) = 52.30$, $p = .000$. Statistical analysis indicated that the test significantly differentiated students with the number of characters they thought they could recognize. Hence, the vocabulary test could be used to differentiate students with different vocabulary knowledge. Third, *test* was significantly correlated with *checklist*, Pearson $r = .80$, $p = .000$.

Fourth, the following two analyses were performed on the subset of participants at one university who took the placement test and whose course levels were known. *Test* was significantly correlated with the placement test (Pearson $r = .58$, $p = .000$). In addition, decision consistency was used as another check for the validity of the test. Specifically, the top and bottom 15% (n=9) of the participants were differentiated by the test. One hundred percent of the bottom students were first-year students, while 78% of the top students were students at second year and above. Clearly, the test was able to differentiate between first- and above first-year students. In sum, the above statistical analyses verified the validity of the vocabulary test.

Statistical analysis was also conducted to measure the reliability of the vocabulary test. The reliability of all 30 test items in the test was also high (Cronbach's Alpha = .94), indicating the reliability of the test.

4.2 Principal Component Analysis of the Questionnaire

Shen (2005) used mean score to identity "commonly used strategies" in her study (p. 56). In this study, the mean score was calculated for each item. A mean score of less than 3.00 indicated that the item was not frequently used by participants; consequently, such items were excluded from the subsequent analysis. Hence, the following items were deleted from the questionnaire: 19, 39, 41, 43, 44, and 48.

A principal component analysis of the questionnaire was conducted. Since the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) was .75, the sample was adequate for this study. Because all items had an anti-image correlation of more than .5, none

was removed from the analysis. Six factors[1] with a loading of more than .4 emerged. Because item 20 had loadings of less than .4, it was removed from analysis.

The six factors were termed as categories. Category One included items 16, 17, 18, 31; Category Two included 14, 37, 38, 40, 42, 45, 46, 47; Category Three included 29, 30, 32, 33, 34, 35; Category Four included 25, 26, 27, 28; Category Five included 13, 15, 24, 36; Category Six included 21, 22, 23.

4.2.1 Reliability

A reliability test was carried out on each factor. Item 28 was deleted from Category Four because its corrected item-total correlation was below .3 and deleting this item could improve the reliability of the category.

For all 28 items, Cronbach's Alpha α equaled .84. Table 1 lists the alpha values for each category.

Table 1: Cronbach's Alpha for Each Category

Category	Cronbach's Alpha
Category One	α = .85
Category Two	α = .76
Category Three	α = .72
Category Four	α = .70
Category Five	α = .60
Category Six	α = .59

Six more variables were created based on the six categories. The value of each variable was the total of points for items included in each category. Specifically, variable *Category One* was the total of points for items 16, 17, 18, 31; *Category Two* was the total of points for items14, 37, 38, 40, 42, 45, 46, 47; *Category Three* was the total of points for items 29, 30, 32, 33, 34, 35; *Category Four* was the total of points for items 25, 26, 27; *Category Five* was the total of points for

[1] In order to ensure at least three items for each factor, the SPSS was programmed to output six factors.

items 13, 15, 24, 36; *Category Six* was the total of points for items 21, 22, 23.

4.2.2 Six strategy categories

There were six categories. Strategy items under these six categories were summarized into six strategy categories.

Category One Strategies: Focusing on Character Orthography

16. My instructors have explained stroke and stroke order.
17. My instructors have explained character structures (e.g., left-right, up-down).
18. My instructors often explain character components (radicals).
31. My instructors have explained word structures (i.e., how characters are combined into words).

The above items belong to Category One. All these four items were significantly negatively correlated with *test*.

Category Two Strategies: Frequent Meaningful Interaction with Characters and Words

14. I often type characters.
37. I often listen to Chinese.
38. I often read in Chinese.
40. I often write in Chinese.
42. I use characters daily.
45. I actively participate in class activities.
46. I often practice speaking Chinese with my peers.
47. I often talk with native speakers in Chinese.

The above items belong to Category Two. Whereas items 37, 38, 40, and 45 were not significantly correlated with *test*, items 14, 42, 46, and 47 were significantly correlated with *test*. Item 14 had the highest correlation with *test*, Pearson $r = .37$, $p = .000$.

Category Three Strategies: Learning Characters and Words in Context

29. I often learn word meanings in sentences.
30. I often learn word usages.

32. When I encounter a new double-character word, I try to guess its meaning from the composite characters.
33. I often compare and contrast related words (e.g. synonyms, antonyms, hypernyms/hyponyms, and homophones).
34. I often learn word collocations (i.e., words often occurring together).
35. I often do exercises to use characters in context.

The above items belong to Category Three. None of these items within this category was significantly correlated with *test*.

Category Four Strategies: Looking up Unknown Characters

25. When I encounter a new character, I try to find out its pronunciation.
26. I know how to use a dictionary (either a paper or an electronic dictionary).
27. I often look up characters in a dictionary (either a paper or an electronic dictionary).

The above items belong to Category Four. Within this category, whereas items 25 and 26 were not significantly correlated with *test*, item 27 was significantly correlated with *test*.

Category Five Strategies: Conventional Ways of Learning Characters and Words

13. I often copy new characters.
15. I often study characters from a vocabulary list.
24. When I encounter a new character, I try to find out its meaning.
36. I often read aloud Chinese sentences and phrases.

The above items belong to Category Five. Within this category, whereas items 24 and 36 were not significantly correlated with *test*, items 13 and 15 were significantly negatively correlated with *test*.

Category Six Strategies: Focusing on Character Pronunciation

21. I often use Pinyin to help me remember the pronunciation of characters.
22. I often listen to the pronunciation of characters.
23. I often read aloud characters.

The above items belong to Category Six. All three items were significantly negatively correlated with *test*. Item 21 had the highest negative correlation with *test*, Pearson $r = -.39$, $p = .000$.

4.3 Prediction

4.3.1 Predicting vocabulary knowledge by strategy categories

Correlation analyses indicated that self-reported proficiency was significantly correlated with *test* (Spearman rho = .53; $p = .000$) and that learning environment was significantly correlated with *test* (Spearman rho = .39; $p = .000$). The latter statistical result indicated that students studying abroad in China got better test scores than those studying at home in the U.S. These two background variables were included in the subsequent regression analyses.

In a hierarchical regression, the dependent variable was the test, and the independent variables in the first step were participants' language proficiency (two dummy variables) and learning environment, and independent variables in the second step were *Categories One*, *Two, Three, Four, Five,* and *Six*.

For each independent variable, it was necessary to have 10 to 15 cases (Field, 2005). Since there were 9 predictors, 90 to 135 cases were needed. The sample of 184 participants was large enough to carry out the multiple regression analysis.

With $F = 19.94$, $p < .000$, the model as a whole was significant. Overall, participants' language proficiency, learning environment, and six categories of strategies explained 48% of the variance in the vocabulary test. As shown in Table Two, the predictors of vocabulary knowledge were proficiency, learning environment, *Category One* (negative), *Category Two, and Category Six* (negative). Specifically,

students studying abroad in China got better test scores than those studying at home in the U.S.

Table 2: Coefficients

		B	Std. Error	Beta
Step 1				
	(Constant)	7.13	.77	
	r_AH_SA	9.03	1.10	.46**
	Intermediates v. beginners	6.93	.99	.43**
	Advanced v. beginners	10.98	1.39	.48**
Step 2				
	(Constant)	9.61	4.43	
	r_AH_SA	8.44	1.21	.43**
	Intermediates v. beginners	4.77	1.08	.29**
	Advanced v. beginners	7.50	1.57	.33**
	Category One	-.29	.14	-.13*
	Category Two	.22	.11	.14*
	Category Three	-.01	.16	-.01
	Category Four	.38	.25	.10
	Category Five	.12	.27	.03
	Category Six	-.82	.25	-.21*

* indicates that $p <= .05$; ** indicates that $p <= .001$

4.3.2 Predicting vocabulary knowledge by individual questionnaire items

The 28 questionnaire items from the six categories were correlated with *test* to see which items were significantly correlated. Results indicated that items 14, 16, 17, 21, 22, and 42 were highly ($r > .25$; $r < -.25$) and significantly ($p < .05$) correlated with *test*. Specifically, items 14 and 42 were positively correlated with *test*, while items 16, 17, 21, and 22 were negatively correlated with *test*.

In a hierarchical regression, the dependent variable was the test, and the independent variables in the first step were participants' language proficiency (two dummy variables) and learning environment, and independent variables in the second step were items 14, 16, 17, 21, 22, and 42.

With $F = 20.71$, $p < .000$, the model as a whole was significant. Together with language proficiency and learning environment (e.g., at home or abroad in China), item 14 "I often type characters" positively predicted vocabulary recognition, whereas item 21 " often use Pinyin to help me remember the pronunciation of characters" negatively predicted vocabulary recognition.

5. Discussion

Statistical analyses found that the vocabulary test was valid and reliable and yielded six strategy categories. In regard to the strategy categories, the regression analysis on the vocabulary test indicated that language proficiency, learning environment, and three strategy categories (i.e., Categories One, Two, and Six) predicted vocabulary recognition. In addition to the two individual background variables, two strategies (i.e., "I often type characters" and "I often use Pinyin to help me remember the pronunciation of characters") predicted vocabulary recognition.

As pointed out by Anderson (2005), "there are no good or bad strategies; there is good or bad application of strategies" (p. 762). Although the excessive use of certain categories of strategies (such as Categories One and Category Six) negatively predicted vocabulary recognition, it might still be relevant to use all six categories in the learning process. These six categories are discussed in detail in the following sections.

5.1 Focusing on Orthography

Category One strategies focus on how characters and words are constructed. These strategies indicate that learners have consciously registered knowledge of strokes, character components, and character structures, as well as morphology. Surprisingly, Category One strategies negatively predicted vocabulary recognition.

One possible explanation is that Category One strategies focus on orthography, which is the way characters are written. Nation (2001) categorized a word's form into written and spoken forms and pointed out that "the strength of the connection between the form and its meaning will determine how readily the learner can retrieve

the meaning when seeing or hearing the word form" (p. 48). That is, strong ties among orthography, meaning, and pronunciation ensure the quick recognition of characters. However, focusing on orthography may deter character recognition.

It is still relevant and important to introduce orthography and morphology in the teaching process. Because written English words differ greatly from Chinese characters, CFL learners need to acquire knowledge of strokes and radicals. In this way, CFL learners can systematically acquire the knowledge of orthography.

Yet, in the meantime, it is important to caution students against excessively relying on orthography to decipher characters and words in the reading process. They need to form strong ties among orthography, meaning, and pronunciation in order to recognize characters and words automatically.

5.2 Frequent Meaningful Interaction with Characters and Words

Including using characters and words in reading, writing, typing, listening, and speaking, Category Two strategies focus on frequent interaction with characters and words for meaning construction. This category predicted vocabulary recognition. The finding is in line with interaction theory in second language acquisition. While interaction theory focuses on people's oral interaction to negotiate meaning (Gass, 2003), Category Two strategies also include interaction with written characters and words.

5.3 Learning Characters and Words in Context

Category Three strategies focus on learning characters and words in context. Interestingly, even though it does not predict vocabulary recognition, this category embodies strategies language instructors usually encourage students to use for vocabulary learning. One possible explanation is that although Category Three strategies that do not predict character recognition, they may be conducive for vocabulary production.

As pointed out by Swain and Lapkin (1995), "these learners sometimes engage in grammatical analysis which, though not essential to comprehension, is essential to accurate production" (p. 384). Different from Roman alphabetic languages, "A Chinese grammar becomes in actual fact very meager; mainly rules for the relative position of the words in the sentence and, in addition, the functions of a number of auxiliary grammatical words" (Karlgren, 1949, p. 68). Hence, learning characters and words in context may help students use them, which may have a similar function to grammatical analysis in Roman alphabetic languages.

5.4 Looking up Unknown Characters

Category Four strategies focus on looking up unknown characters. This category did not predict vocabulary recognition. The results indicate that looking up unknown characters is not sufficient for character and word recognition. It may be because CFL learners need to integrate the looked-up vocabulary into their existing vocabulary in order to improve character and word recognition.

It may still be relevant to encourage students to look up unknown characters because this category enables students to acquire vocabulary on their own. It is also important to teach them how to integrate the new characters and words into their existing vocabulary.

5.5 Conventional Ways of Learning Characters and Words

Category Five strategies focus on learning characters in conventional ways, such as copying characters and studying a vocabulary list. For example, in the learning process, CFL learners often need to copy characters many times in a character workbook. Researchers have found that CFL learners perceived that copying characters is conducive to character learning (Ke, 1998b; Wang & Leland, 2011).

This category did not predict vocabulary recognition. The reason might be that the conventional ways of learning do not sufficiently set up strong ties between the form and meaning of vocabulary.

It may still be relevant to let students practice characters and words in conventional ways. Yet it is also important to inform students that the conventional ways alone are not sufficient to achieve automatic character and word recognition. The effective use of conventional methods depends on what additional activities a teacher requires in the teaching process.

5.6 Focusing on Character Pronunciation

Category Six strategies target character pronunciation. The reason that CFL learners pay special attention to character pronunciation is because it is difficult for them. Yin (2003) found that his participants perceived writing and remembering character pronunciation as difficult. Similarly, Shen (2010) found that, "both sounds and shapes were considered to be more difficult to learn than meanings" (p. 51).

Category Six strategies negatively predicted vocabulary recognition. The explanation may be that by focusing on pronunciation only, CFL learners do not set up strong ties among orthography, meaning, and pronunciation.

It may still be relevant to encourage students to employ Category Six strategies to practice pronunciation. However, they need to be further pushed to efficiently connect form with meaning for automatic character recognition.

5.7 Pinyin

As for individual strategy items, item 14 (I often type characters) positively predicted vocabulary recognition, whereas item 21 (I often use Pinyin to help me remember the pronunciation of characters) negatively predicted vocabulary recognition. The above findings indicate the unique function of Pinyin as a crucial tool in vocabulary learning.

Pinyin plays both the facilitative and hindrance roles in recognizing characters and words. On the one hand, knowledge of Pinyin is important because it enables participants to type characters, which is a predictor for vocabulary recognition. Specifically, in typing

a character, CFL learners need to complete the following two steps: inputting Pinyin and then choosing the appropriate character from a list with the same pronunciation (regardless of tones) on the basis of its orthography and meaning. This statistical finding indicates from a different perspective the importance of combining the knowledge of character pronunciation (demonstrated through Pinyin) with orthography and meaning. On the other hand, "I often use Pinyin to help me remember the pronunciation of characters" negatively predicts vocabulary recognition. Studies of native Chinese children indicated that "Chinese children do rely on phonetics for phonological cues in reading Chinese characters" (Ho & Bryant, 1997a, p. 287) and "learning to read Chinese progresses from a visual phase to a phonological phase" (Ho & Bryant, 1997b, p. 950). However, relying on Pinyin to read out character pronunciation does not help CFL learners to derive phonological cues from characters. Instead, after students learn character pronunciation through Pinyin, it might be important for them to become accustomed to pronouncing characters without referring to Pinyin.

6. Conclusion

Combining students with various language proficiencies in two learning environments, the present study indicates that successful application of learning strategies for the recognition of characters and words is predicted by learners' ability to connect forms with meanings effectively, which is achieved by frequently interacting with characters and words for meaning construction. In contrast, excessively relying either on orthography or pronunciation hinders character recognition. As for learning characters and words in context, this category may be conducive to production, yet does not predict character and word recognition. In regards to looking up unknown characters and employing conventional ways of character learning, these two categories may be necessary, yet not sufficient, for character recognition.

The study also identifies the unique role of Pinyin as a tool. Findings from this study suggest that the use of Pinyin to type characters facilitates vocabulary recognition, while overreliance on Pinyin to pronounce characters hinders the recognition process.

With regard to background variables, the present study finds that learning environment predicts vocabulary recognition. Specifically, studying abroad significantly improves vocabulary recognition more than studying at home in the U.S. The study also confirms that proficiency levels predict vocabulary recognition.

The present study has several pedagogical implications. First, one important role of instructors is to ensure frequent, meaningful interactions with characters and words by CFL learners. This is especially crucial for beginning and intermediate students whose limited language proficiency may impede their frequent interactions with characters and words. Instructors, therefore, bear the burden of designing meaningful exercises to ensure students' regular interactions with both characters and words. In addition, instructors can use computer technology to help beginning and intermediate CFL learners read Chinese texts (Wang, 2011; Wang & Upton, 2012). Second, language instructors need to guide students so that they will not excessively focus on limited aspects in the learning process. Third, instructors should choose different exercises to strengthen different achievements. Specifically, if they plan to improve students' character recognition, they should not solely focus on learning characters and words in context. Here a fine distinction needs to be made. Learning characters and words in context does not equal frequent meaningful interactions. Fourth, instructors should make CFL learners aware of the insufficiencies of looking up unknown characters and of employing conventional methods of learning characters and words. Fifth, on the one hand, instructors need to help learners become proficient in using Pinyin to type characters; on the other hand, instructors need to gradually help students form the habit of pronouncing characters directly without relying on Pinyin. That is, after students initially learn the pronunciation of characters through Pinyin, they should not refer to Pinyin every time they pronounce the characters.

Future Directions

More data need to be gathered regarding the influence of other background variables on the effective application of learning strategies to acquire characters and words. In addition, it would be

useful to focus on specific proficiency groups to identify their respective learning strategies. Furthermore, it would be beneficial to identify various effective strategies used by students in different learning environments.

Acknowledgements

Jing Wang would like to thank the IU School of Liberal Arts at Indiana University- Purdue University, Indianapolis for granting her the Summer Research and Creative Activity Grant in 2010 to support this study.

References

Anderson, N. J. (2005). L2 learning strategies. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 757-771). Mahwah, NJ: Erlbaum.

Churchill, E., & DuFon, M. A. (2006). Evolving threads in study abroad research. In M.A. DuFon & E. Churchill (Eds.), *Language learners in study abroad contexts* (pp. 1-27). Clevedon, UK: Multilingual Matters.

Everson, M. E., & Ke. C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 32*(1), 1-20.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage.

Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of second language acquisition* (pp. 224-255). Malden, MA: Blackwell Pub.

Golkar, M., & Yamini, M. (2007). Vocabulary, proficiency and reading comprehension. *The Reading Matrix, 7*(3), 88–112.

Hayden, J. J. (1998). The influence of a semester abroad on reading proficiency: A descriptive study. *Journal of the Chinese Language Teachers Association 33*(3), 13- 24.

Ho, C. S.-H., & Bryant, P. (1997a). Learning to read Chinese beyond the logographic phase. *Reading Research Quarterly, 32*(3), 276-289.

Ho, C. S.-H., & Bryant, P. (1997b). Phonological skills are important in learning to read Chinese. *Developmental Psychology, 33*(6), 946-951.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese.* Hillsdale, NJ: Erlbaum.

Jafarpur, A. (2002). A comparative study of a C-Test and a cloze test. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und praktische Anwendungen* (Vol. 4, pp. 31-55). Bochum, Germany: AKS-Verlag.

Jin, H. G. (2006). Multimedia effects and Chinese character processing: An empirical study of CFL learners from three different orthographic backgrounds. *Journal of the Chinese Language Teachers Association, 41*(3), 35-56.

Karlgren, B. (1949). *The Chinese language: An essay on its nature and history.* New York: The Ronald Press Company.

Ke, C. (1996). An empirical study on the relationship between Chinese character recognition and production. *The Modern Language Journal, 80*(3), 340-349.

Ke, C. (1998a). Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals, 31*(1), 91-100.

Ke, C. (1998b). Effects of strategies on the learning of Chinese characters among foreign language students. *Journal of the Chinese Language Teachers Association, 33*(2), 93- 112.

Lee-Thompson, L.-C. (2008), An Investigation of reading strategies applied by American learners of Chinese as a foreign language. *Foreign Language Annals, 41*(4), 702–721.

Liu, X. (2000). *对外汉语教育学引论* [Introduction to the pedagogy of teaching Chinese as a foreign language]. Beijing, China: Beijing Language University Press**北京语言大学出版社**.

McGinnis, S. (1999). Student goals and approaches. In M. Chu (Ed.), *Mapping the course of the Chinese language field* (Chinese Language Teachers Association Monograph Series,

Monograph 3, pp. 151-188). Kalamazoo, MI: Chinese Language Teachers Association.

Nambiar, R. (2009). Learning strategy research—Where are we now? *The Reading Matrix, 9*(2), 132-149.

Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines, 5,* 12-25.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* NY: Cambridge University Press.

Reed, J. (1988). Measuring the vocabulary knowledge of second language learners. *RELC Journal, 19*(2), 12-25.

Reed, J., & Chapelle, C. A. (2001). A framework for second language vocabulary assessment. *Language Testing, 18*(1), 1-32.

Shen, H. H. (2003). A comparison of written Chinese achievement among heritage learners in homogeneous and heterogeneous groups. *Foreign Language Annals, 36*(2), 258-266.

Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non- native speakers of Chinese. *System, 33*, 49-68.

Shen, H. H. (2010). Analysis of radical knowledge development among beginning CFL learners. In M. E. Everson & H. H. Shen (Eds.), *Research among learners of Chinese as a foreign language* (Chinese Language Teachers Association Monograph Series, Vol. IV, pp. 45-65). Honolulu, HI: University of Hawai'i at Manoa.

Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *The Modern Language Journal, 91*(1), 97-111.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371- 391.

Testing Center of Beijing Language University (2006). *中国汉语水平考试大纲* [Chinese proficiency test outline] (2nd ed.). Beijing, China: Beijing Language University Press 北京语言大学出版社.

Wang, J. (2011). The use of e-dictionary to read e-text by intermediate and advanced learners of Chinese. *Computer Assisted Language Learning*, i*First* article, 1-13. DOI:10.1080/09588221.2011.631144.

Wang, J., & Leland, C. (2011). Beginning students' perceptions of effective activities for Chinese character recognition. *Reading in a Foreign Language, 23*(2), 208-224.

Wang, J., & Upton, T. A. (2012). The impact of using a pop-up dictionary on the reading processes of beginning learners of Chinese. *Journal of the Chinese Language Teachers Association, 47*(1), 23-41.

Xiao, Y. (2006). Heritage learners in the Chinese language classroom: Home background. *Heritage Language Journal, 4*(1), 47-56.

Yin, J. H. (2003). **美国大学生**记忆汉字时使用的方法--问卷调查报告 [A survey of methods used by American college students in memorizing Chinese characters]. *Journal of the Chinese Language Teachers Association, 38*(3), 69-90.

Appendix A. Background Questions and Questionnaire for Chinese Character and Word Recognition

Background Questions
1. In which university are you studying Chinese now?
2. What is your native language? • English • Japanese • Korean • Other European Languages • Other Languages
3. Did you grow up speaking Chinese at home? • No • Yes (always or sometimes)
4. What do you think your Chinese proficiency might be? • Elementary • Intermediate • Advanced
5. How many characters do you recognize? • Less than 500 • Between 500 and 1,000 • Between 1,000 and 2,000 • Between 2,000 and 3,000 • More than 3,000
6. Do you read simplified or traditional characters?

• Simplified characters • Traditional characters • Both simplified and traditional characters
7. How long have you been studying Chinese? • Less than a year • Between one and two years • Between two and three years • More than three years
8. What is your gender? • Male • Female
9. What is your age? • Under 18 • Between 18 and 30 • Between 30 and 40 • Over 40
10. What is the highest education you have received? • High school • Some college • College (Bachelor's degree) • Some graduate work • Graduate degree
11. What is your race or ethnic background? • Asian and Pacific Islander • Black • Hispanic • Native American • White
12. What is your family income range? • 0 to 25,000 • 25,000 to 50,000 • 50,000 to 75,000 • 75,000 to 100,000 • Above 100,000
Questionnaire
Character Recognition in Isolation
13. I often copy new characters.
14. I often type characters.

15. I often study characters from a vocabulary list.
16. My instructors have explained stroke and stroke order.
17. My instructors have explained character structures (e.g., left-right, up-down).
18. My instructors often explain character components (radicals).
19. When I learn new characters, my instructors ask me to identify components.
20. I often compare and contrast characters which look similar.
21. I often use Pinyin to help me remember the pronunciation of characters.
22. I often listen to the pronunciation of characters.
23. I often read aloud characters.
24. When I encounter a new character, I try to find out its meaning.
25. When I encounter a new character, I try to find out its pronunciation.
26. I know how to use a dictionary (either a paper or an electronic dictionary).
27. I often look up characters in a dictionary (either a paper or an electronic dictionary).
28. I often use small, portable electronic devices (e.g. pocket e-dictionary, i-phone) to learn new characters.
Character Recognition in Context
29. I often learn word meanings in sentences.
30. I often learn word usages.
31. My instructors have explained word structures (i.e., how characters are combined into words).
32. When I encounter a new double-character word, I try to guess its meaning from the composite characters.
33. I often compare and contrast related words (e.g., synonyms, antonyms, hypernyms/hyponyms, and homophones).
34. I often learn word collocations (i.e., words often occurring together).
35. I often do exercises to use characters in context.
36. I often read aloud Chinese sentences and phrases.
37. I often listen to Chinese.
38. I often read in Chinese.
39. I often read subtitles in Chinese when I watch TV or DVDs.
40. I often write in Chinese.
41. I often sing Chinese songs in karaoke.
42. I use characters daily.
Character Recognition in Social Interaction
43. I often do written homework assignments with my peers.

44. I often practice reading Chinese texts with my peers.
45. I actively participate in class activities.
46. I often practice speaking Chinese with my peers.
47. I often talk with native speakers in Chinese.
48. I often write to native speakers in Chinese.

Appendix B. Vocabulary Test

Please choose one best answer for each sentence **without** using a dictionary. If you don't know the answer, please choose item E (**不知道**, meaning I don't know).

1. 你好! 我____学生, 你呢? A. 也 B. 吗 C. 不 D. 是 E. 不知道	2. 我 家____五个人: 妈妈, 爸爸, 哥哥, 姐姐和我。 A. 有 B. 又 C. 友 D. 上 E. 不知道
3. 我们____学中文。 A. 大 B. 都 C. 者 D. 那 E. 不知道	4. 我昨天给你____了一个电话。 A. 喝 B. 打 C. 到 D. 帮 E. 不知道
5. 今天____昨天冷。 A. 比 B. 花 C. 此 D. 话 E. 不知道	6. 你家____机场远不远? A. 买 B. 坐 C. 实 D. 离 E. 不知道
7. 我肚子____死了, 我得马上去医院。 A. 饿 B. 疼 C. 累 D. 忙 E. 不知道	8.小李常常帮助别人, 我____她的印象很好。 A. 双 B. 时 C. 对 D. 又 E. 不知道

9. 请问, ____一封平信多少钱? A. 读 B. 开 C. 放 D. 寄 E. 不知道	10. 一直往前走, 到了第2个红绿灯, 往左 一____就到了。 A. 说 B. 拐 C. 挑 D. 真 E. 不知道
11. 我学中文, 你的____是什么? A. 教学 B. 文学 C. 学分 D. 专业 E. 不知道	12. 你是新生吧, 你办了注册 ____了吗? A. 学分 B. 宿舍 C. 学校 D. 手续 E. 不知道
13.下了飞机以后, 我怎么也找不到我托运的____。 A. 机票 B. 机场 C. 行李 D. 里程 E. 不知道	14. 你____过北京博物馆吗？ A. **参**观 B. **旅行** C. **开放** D. **学**习 E. 不知道
15.这几年中国人的生活____有了很大的提高。 A. 变化 B. 收入 C. 水平 D. 问题 E. 不知道	16.____下大雨, 我也要去看这个电影。 A. 即使 B. 即可 C. 既然 D. 既是 E. 不知道
17. ____在这里等她, 不如去找她。 A. 未必 B. 相对 C. 至于 D. 与其 E. 不知道	18. 这个城市犯罪率比较低, 在这里生活很____。 A. 便利 B. 薄弱 C. 不利 D. 安定 E. 不知道

19.发明了计算机, 人类社会就____了一个崭新的时期。 A. 跨进 B. 操纵 C. 加紧 D. 排列 E. 不知道	20.目前许多国家都____促进经济发展。 A. 曲折 B. 坚固 C. 切实 D. 舒畅 E. 不知道
21.家长应该帮助孩子____爱读书的好习惯。 A. 养成 B. 朗诵 C. 奖励 D. 缓和 E. 不知道	22.我们必须____地面对各种困难。 A. 急躁 B. 谅解 C. 冷静 D. 缓慢 E. 不知道
23.随着环境的恶化, 动物的生存____越来越小。 A. 差别 B. 状态 C. 空间 D. 压力 E. 不知道	24.这个故事的开头很平常，可是____却出人意料。 A. 理解 B. 结局 C. 结算 D. 作者 E. 不知道
25.这个贫困地区矿产丰富，具有巨大的发展____。 A. 能量 B. 能力 C. 潜力 D. 潜在 E. 不知道	26.宏观经济调控____影响股市行情。 A. 所以 B. 于是 C. 然而 D. 势必 E. 不知道
27.为这点小事当众批评他____太小题大作了。 A. 难免 B. 未免 C. 不免 D. 以免 E. 不知道	28.小数点以后的数字____不计。 A. 忽略 B. 忽视 C. 忽然 D. 忽闪 E. 不知道

29.他是一名优秀的运动员：动作____，心理素质良好。 A. 快速 B. 敏捷 C. 敏锐 D. 高速 E. 不知道	30.今年的特大自然灾害给当地居民带来了无法____的损失。 A. 弥合 B. 弥补 C. 补救 D. 补缺 E. 不知道

Appendix C. A Checklist of Vocabulary Knowledge

Set One	Set Two	Set Three	Set Four
才	巧	猛	皆
快	鲜	顺	界
觉得	以及	仍旧	颁布
正在	概括	容器	储藏
相信	测验	谨慎	融洽
作业	季节	统计	连同
方向	粮食	科普	慈善