TEXTBOOK ANALYSIS: VOCABULARY INSTRUCTION OF CHINESE AS A FOREIGN LANGUAGE

A Thesis

Presented to the

Faculty of

San Diego State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

In

Linguistics

by

Ran He

Summer 2010
SAN DIEGO STATE UNIVERSITY

The Undersigned Faculty Committee Approves the

Thesis of Ran He:

Textbook Analysis: Vocabulary Instruction of Chinese as a Foreign Language

Zheng-sheng Zhang, Chair
Department of Linguistics and Asian/Middle Eastern Languages

Deborah Poole
Department of Linguistics and Asian/Middle Eastern Languages

Guang Lei
Department Of Political Science

5/23/2010
Approval Date
This thesis focuses on the study of vocabulary instruction in Chinese. Due to the fact that Chinese is not a cognate language of Indo-European languages, word formation becomes one of the most challenging factors when learning Chinese vocabulary. The other cause that makes the Chinese vocabulary so difficult to acquire is the lack of assistance from the textbooks. This study points out the deficiencies of the vocabulary selection in three widely used series of textbooks at elementary and intermediate level, as well as the non-systematic vocabulary instruction methods found within the textbooks. One finding of this study is an unexpectedly low coverage of the most frequently used words as classified by the HSK (a national Chinese language proficiency test) in the textbooks at elementary and intermediate level. Meanwhile, the vocabulary selections in the textbooks also do not reflect the word distribution of natural Chinese language. Besides the quantitative analysis focusing on the vocabulary selections in each of the textbook series, the thesis also sheds light on the lack of connections among vocabulary items. These lack of connections include (1) low word and morpheme recycling among the vocabulary items as well as a lack of repetition of words and morphemes throughout the textbooks; and (2) the selected new words do not manage to complete semantic networks. Therefore, the vocabulary instruction methods of the textbooks are not satisfactory when it comes to teaching vocabulary in Chinese as a foreign language.

In order to improve CFL vocabulary instruction, two teaching techniques are advocated by this thesis, including the increasing of vocabulary recycling and improving the connections among new vocabulary items by making use of shared morphemes and integrated semantic networks. When it comes to vocabulary selections, a well-organized vocabulary list should be adapted throughout each textbook series. Meanwhile, the high-frequency words found in the HSK vocabulary list should be included more, especially at the elementary and intermediate level.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT ............................................................................................................................. iv</td>
</tr>
<tr>
<td>LIST OF TABLES .................................................................................................................. vii</td>
</tr>
<tr>
<td>LIST OF FIGURES ............................................................................................................... viii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS ......................................................................................................... ix</td>
</tr>
<tr>
<td>CHAPTER</td>
</tr>
<tr>
<td>1 INTRODUCTION .........................................................................................................1</td>
</tr>
<tr>
<td>The Difficulties of Learning Chinese Vocabulary ...................................................2</td>
</tr>
<tr>
<td>Semantic Networks ..................................................................................................6</td>
</tr>
<tr>
<td>Vocabulary Repetition/Recycling ............................................................................7</td>
</tr>
<tr>
<td>CFL Textbooks ........................................................................................................7</td>
</tr>
<tr>
<td>Research Questions ..................................................................................................8</td>
</tr>
<tr>
<td>Research Question 1: .........................................................................................8</td>
</tr>
<tr>
<td>Research Question 2: .........................................................................................8</td>
</tr>
<tr>
<td>2 METHODOLOGY ........................................................................................................9</td>
</tr>
<tr>
<td>Materials: The CFL Textbooks ................................................................................9</td>
</tr>
<tr>
<td>The Reference Vocabulary List .............................................................................10</td>
</tr>
<tr>
<td>Instruments and Procedures ...................................................................................12</td>
</tr>
<tr>
<td>3 DATA ANALYSIS ......................................................................................................14</td>
</tr>
<tr>
<td>Vocabulary Presentation ........................................................................................14</td>
</tr>
<tr>
<td>Vocabulary Exercises ............................................................................................16</td>
</tr>
<tr>
<td>The Quantitative Analysis of Vocabulary in Textbooks ........................................19</td>
</tr>
<tr>
<td>Comparison of Word Types Between the Textbook Vocabulary and the Hsk Vocabulary List ..................................................................................................................21</td>
</tr>
<tr>
<td>The Distribution of the Four Grades in the Textbooks ........................................21</td>
</tr>
<tr>
<td>The Distribution of the Four Grades Words at Each Level .............................22</td>
</tr>
<tr>
<td>Grade A Words in the Textbooks .................................................................24</td>
</tr>
<tr>
<td>Vocabulary Repetition/Recycling .................................................................25</td>
</tr>
</tbody>
</table>
Lack of Integration of Semantic Networks ............................................................27
Sequencing of Multiple Meanings .........................................................................27

4 CONCLUSION...........................................................................................................29
Research Question 1 ..............................................................................................29
Research Question 2 ..............................................................................................30

5 PEDAGOGICAL IMPLICATIONS...........................................................................33
Systematicity of Vocabulary Selections .................................................................33
The Use of Shared Morphemes in Different Compounds ....................................33
Making Connections among Words.........................................................................36
Vocabulary Recycling............................................................................................36

6 LIMITATIONS AND FURTHER RESEARCH....................................................38
REFERENCES .............................................................................................................39
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Students’ Mistakes of 人 from Li’ Study (2005)</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>中國漢語水平考試詞彙大綱 (Chinese Proficiency Test Vocabulary Guideline a Dictionary of Chinese Usage: 8000 words)</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>Comparison of Vocabulary List in Each Lesson for the Three Textbooks Series</td>
<td>15</td>
</tr>
<tr>
<td>4</td>
<td>Comparison of Vocabulary Index in Each Lesson for the Three Textbooks Series</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>Comparison of Vocabulary Exercise in Each Lesson for the Three Textbooks Series</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>Word and Morpheme TTR in CL and IC</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>Morpheme TTR in the Vocabulary Lists in the Textbooks</td>
<td>26</td>
</tr>
<tr>
<td>8</td>
<td>A Modified Vocabulary List of Lesson 11 Level 1 Part 2, Integrated Chinese</td>
<td>34</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

PAGE

Figure 1. Explanations of 人 in the Chinese proficiency test vocabulary guideline a dictionary of Chinese usage: 8000 words. .................................................................5

Figure 2. Total number of the vocabulary items (type) from the investigated textbooks. .................................................................................................................19

Figure 3. The number of the vocabulary items (type) distribution in 1st and 2nd year study of Chinese .................................................................20

Figure 4: An excerpt from Yu Wen, p.70. .................................................................................................................20

Figure 5. The distribution of the four levels of the vocabulary items in the three textbook series. ..................................................................................21

Figure 6. The distribution of four grades’ vocabulary in CL .................................................................22

Figure 7. The distribution of four grades’ vocabulary in IC ..................................................................................23

Figure 8. The distribution of four grades’ vocabulary in NPCR ...........................................................................23

Figure 9. Comparison of the three textbook series with the HSK vocabulary list ........................................24
ACKNOWLEDGEMENTS

This thesis is dedicated to my loving parents, 赫玉民 and 刘伏中, and my dearest husband, Eric Dalton. Also, I am greatly indebted to Dr. Zheng-sheng Zhang. Without his patient instruction, this thesis might not have been finished.
CHAPTER 1

INTRODUCTION

Learning a second language successfully involves students’ mastery of multiple linguistic skills; including phonology, morphology, vocabulary, grammar and pragmatics. Nevertheless, focusing on vocabulary instruction has not been a priority in the literature of second language acquisition (SLA) or teachers. Teachers tend to agree that grammar is the most challenging part of learning Chinese as a foreign language (CFL) but are of the opinion that vocabulary is easier to learn and thus not worth the investment of a large amount of in-class time (Carter & McCarthy, 1988). Zimmerman (1997) points out that the teaching and learning of vocabulary have been undervalued in the field of SLA. Consequently, the neglect of vocabulary instruction leads to the incorrect assumption that the only way for students to learn vocabulary words is to memorize each vocabulary item in the textbooks by themselves.

Vocabulary is central to second language learning for a few reasons. First of all, knowledge of vocabulary is crucial to successful communication as vocabulary plays an important role in delivering meaning. While the lack of grammatical knowledge hinders successful communication, it is the absence of vocabulary that impedes complete communication (Wilkins, 1972). That is to say, the meaning of sentences can be understood based on the context, even if there are grammar mistakes; however, the wrong vocabulary choice could make a sentence completely incomprehensible. Secondly, students understand the importance of vocabulary and are eager to learn new vocabulary items (Leki & Carson, 1994). Based on James’s (1996) survey results, students are particularly interested in receiving vocabulary instruction. Lastly, vocabulary acquisition affects other language skills’ development. Sternberg (1987) has argued that "one's level of vocabulary is highly predictive, if not determinative, of one's level of reading comprehension"(p. 90). Studies of spoken English (Schonell, Meddleton, & Shaw, 1956) show that a vocabulary of around 2,000 word families needs to be mastered to provide 95% coverage of informal conversations. In addition to vocabulary’s influence on reading skills, Moss and Marslen-Wilson (1993) acknowledge that vocabulary is the key element in SLA listening.
comprehension and according to Nation (2001), vocabulary choice is a strong indicator of whether a writer has adopted the conventions of the relevant discourse community. Also, Leki and Carson (1994) found that second language learners regard vocabulary as the major element affecting their writing quality. Last, Thompson (2008) points out that vocabulary learning has significant influences on reading skills of American learners of CFL. In her research, reading difficulties were mainly caused by vocabulary issues above all other factors that she examined. These other factors included orthography, grammar and background knowledge.

Besides the importance of vocabulary acquisition in SLA as well as in Chinese as a foreign language, the difficulties inherent in CFL vocabulary learning and the importance of vocabulary instruction in Chinese are also accentuated. In particular, the confusion between morpheme and word in vocabulary instruction is pointed out in the next section. In order to gain insights relating to the CFL vocabulary teaching, the CFL vocabulary instruction approaches are also discussed within this study with respect to the significance of morphemes, semantic networks and vocabulary recycling. Last but not least, the thesis points out the significance of CFL textbooks in vocabulary acquisition. Meanwhile, the inadequacies of vocabulary instruction in textbooks, such as non-systematic vocabulary choice and insufficient vocabulary presentation and practice, are discussed in this study as well.

THE DIFFICULTIES OF LEARNING CHINESE VOCABULARY

From the perspective of vocabulary acquisition, Chinese can be a very intimidating language for learners who study it as a foreign language. First of all, it is difficult to infer the pronunciation from the words. That is, at the beginning stage, CFL students have to memorize the pronunciation of each word; since the writing system does not provide much of a hint to matching the sounds and the written symbols. Not until much later on are students able to use the knowledge of phonetic components to guess a word’s pronunciation. Even so, the phonetic components are not completely informative. According to Zhou (2003), the fraction of the words that have effective phonetic components is approximately 30% of all compound characters. The data does not count the tone differences. If tonal differences were considered in Zhou’s research, the effectiveness of phonetic clues would have been even
More importantly, due to the fact that Chinese is not a cognate language of Indo-European languages, Chinese is more difficult for students of English to learn than other Indo-European languages. According to Liskin-Gasparro (1982), English learners of Chinese take three or four times as long to achieve professional level proficiency when compared to learning a cognate language, such as French or Italian. Take the word “government” for example. The English word “government” is a cognate of the French “gouvernement” and Italian “governo”; while in Chinese, “government” is “zheng fu” (“politics” and “mansion”), which do not respond to “govern” and “ment” respectively either in pronunciation or in meaning. The cognates of the word “government” are similar in both spelling and pronunciation thus facilitating comprehension and retention of vocabulary. Compared to learning a cognate language of English, to acquire Chinese vocabulary is arduous for CFL students whose first languages are Indo-European languages.

Besides the lack of cognates of Indo-European languages which makes Chinese difficult to access, the other obstacles relating to CFL vocabulary pedagogy center around the basic unit of vocabulary: the word 詞 or the morpheme 語素, which has influence on students’ learning of the ionized morphemes, the slippery boundary between words and phrases. It is true that word is the most important and most salient type of speech unit between the phoneme and the connected speech (Chao, 1968) and is the most frequently used unit of vocabulary in language teaching. In Chinese, on the other hand, it is the morpheme that plays a crucial role in constructing words. However, the morpheme is not given as much attention in CFL vocabulary teaching. The reason why the morpheme is neglected may be that CFL vocabulary instruction very much emulates English as a foreign language (EFL) vocabulary instruction in which words are vocabulary items, not morphemes. However, Chinese is a language that is very different from English from the perspective of word formation.

Morphologically complex words are formed through the morphological process of inflection, derivation, and compounding. Take English, for example. The main word formations used by English are inflection and derivation, while Chinese uses compounding most commonly. An inflectional change occurs in English, for instance, when “-s” or “-ing” is added to a word to signify plurality or aspect, whereas an example of derivational change
takes place when adding “-able” to the end of a verb to create an adjective, such as in play and playable.

Unlike English, where inflections and derivations are the main word formation method, Chinese makes use of a limited number of inflectional morphemes, such as the numeral classifier 們, as well as a limited amount of derivational processes, for example 的 (changing a verb into a noun) and 化 (changing a noun or an adjective into a verb).

Compounding is the formation of new words by combining two or more stem morphemes. In Chinese, compounds encompass the majority of words and as such, compounding is the main word formation method (Chen, Hao, Geva, Zhu, & Shu, 2009) and a highly productive way of creating new words. It has been shown that compound words make up over 70% of all words used in Chinese (Institute of Language Teaching and Research, 1986), or approximately 75% of the words according to Chen et al. (2009).

Many morphemes are used recurrently in constructing compound words. On average, a single Chinese morpheme appears in about 17 compound words (Yin, 1984; Yuan & Huang, 1998). The productivity of the compounding system can be explained by semantic transparency of the morphemes (McBride-Chang et al., 2008). Most Chinese compound words are highly semantically transparent, which means that the meaning of each constituent morpheme transparently contribute to the meaning of a compound.

Take the words relating to meat 肉 for instance. Rather than three separate words for “pork”, ”beef”, and ”mutton” in English, in Chinese, these same three words are all connected by one morpheme 肉 (meat): 豬(pig)肉, 牛(cow)肉, 羊(mutton)肉. In addition to the examples given above, there are more compound words that relate to meat which share the same morpheme 肉, such as 肥肉(fatty meat), 肉餅(patties) and 肉湯 (broth). As one can see, all the examples show the productivity and semantic transparency of morphemes in Chinese.

However, even the compounds can be semantically transparent CFL vocabulary cannot be simply taught by using morphemes as the basic unit. Li’s (2005) research demonstrates that words in Chinese are not always constructed by putting related morphemes together. A typical set of mistakes made by CFL learners, the misuse of 人 (human being), are shown as follows.
According to the dictionary entries of 人 taken from *the Chinese Proficiency Test Vocabulary Guideline a Dictionary of Chinese Usage: 8000 words* (shown in Figure 1), entry 3 is used when 人 functions as the morpheme referring to specific kinds of people; while for the other entries, 人 is regarded as a word.

![Dictionary entry for 人](image)

Translation:

1. Human being
2. Adults
3. Specific kind of people
4. Other people
5. Personality

Refer in particular to a person

Figure 1. Explanations of 人 in the Chinese proficiency test vocabulary guideline a dictionary of Chinese usage: 8000 words. 

As shown in Table 1, the first mistake, which involves the word farmer (農民), shows that farmer possesses a meaning of person. Adding the morpheme 人 to the word farmer reflects the students are not aware of the meanings of components. The second misuse
Table 1. Students’ Mistakes of 人 from Li’ Study (2005)

<table>
<thead>
<tr>
<th>Students’ mistakes</th>
<th>Gloss</th>
<th>Correct forms</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>農民人</em></td>
<td>*farmer+person</td>
<td>農民</td>
<td>Farmer</td>
</tr>
<tr>
<td><em>正直人</em></td>
<td>*honest+person</td>
<td>正直的人</td>
<td>Honest person</td>
</tr>
<tr>
<td><em>騙人</em></td>
<td>*cheat+person</td>
<td>騙子</td>
<td>Liar</td>
</tr>
</tbody>
</table>

indicates that the students are not aware of the meanings of components. The second misuse indicates that students do not distinguish word 人 from morpheme 人. The last mistake shows that 人 is not the only morpheme expressing “human being” as 子 can also fulfill a similar role. As can be seen in Table 1, the interpretation of words are more complicated than adding up the constituent morphemes in them, and wholly relying on morphemes’ meanings may result in the inaccurate or the false understandings of words.

**Semantic Networks**

Semantic network refers to the interconnecting groups of words within the same semantic field.

Mackey (1965) defines the concept of a semantic field as being:

made of basic key-words, which command an army of others. The semantic area may be regarded as a network of hundreds of associations, each word of which is capable of being the center of a web of associations radiating in all directions. (p. 76)

The pedagogical value of utilizing connections among words lies in the fact that the words in the network are connected so closely that once one word is activated, other words within the same network are activated as well. Crow and Quigley (1985) use a semantic approach to vocabulary acquisition and argue that long-term retention of vocabulary learning that has been organized into some type of cognitive categories is superior to retention of randomly presented words. Also, Nation (2001) suggests that understanding semantic networks is useful for understanding the meanings of words. In addition, Fu (2005) obtains similar results, which suggest that the participants of his study were using semantic networks as part of their vocabulary learning strategies. Examples of the semantic networks strategies employed include “creating mental linkages”, “grouping and associating words by radicals” and “reviewing and linking with already known materials”.

Word association is not only limited to semantic connections among words but also through shared morphemes, parts of speech and so on. For example, Mo (2004) analyzes the association among words which are required for *Hanyu Shuipin Kaoshi* (HSK, the National Chinese Proficiency Test) by a number of criteria which include morpheme position (i.e. front and back), sequence (i.e. number, seasons), synonyms, antonyms, homogony, semantic meanings, hierarchy (i.e. fruit) and function words (i.e. measure words). She points out that by organizing words and expressions as mentioned above, it helps students memorize not only the words but also the groups of words related to those words. Furthermore, Nation (2001) points out that understanding relations among words helps students comprehend the words’ full meanings and also helps students build upon these meanings when applied to the appropriate contexts. As a result, the associations among words enhance the effectiveness of CFL students’ vocabulary acquisition.

**VOCABULARY REPETITION/RECYCLING**

Greene (1992) notes that “The saying ‘practice makes perfect’ illustrates the accepted fact that repetition is critical for learning. All other things being equal, our memory for information will depend on the number of times that we have encountered or studied it” (p.132). From the perspectives of vocabulary learning, the more times learners encounter vocabulary items, the better chance they have of learning them. According to Barcroft (2004), learners begin to acquire new words by having the words presented to them frequently and repetitively in the input. Blake and Majors (1995) also point out that vocabulary learned through recycled words fosters long-term learning rather than short-term retention. According to Nation (2001), vocabulary is not about knowing words, but also knowing them so well that they can be fluently used.

**CFL TEXTBOOKS**

Textbooks are the primary resource for students when learning vocabulary, besides from the teachers. Winke & Abbuhl (2007) indicate that one of the most widely used vocabulary learning strategies students employ is “reading, reviewing or studying from the book” (p.704). Therefore, CFL textbooks play a significant role in students’ vocabulary acquisition and all the issues of vocabulary instruction mentioned above reflect in CFL textbooks.
RESEARCH QUESTIONS

This study is guided by the following research questions:

Research Question 1:
How are textbooks different from each other from the perspective of vocabulary presentation and vocabulary exercises?

Research Question 2:
Is the presentation of vocabulary in the CFL textbooks systematic?

In order to answer the research questions, the present study will first look at the quantity of vocabulary distribution in both the selected textbooks as well as the HSK vocabulary list. This will be followed by a comparison of the two groups of vocabulary items. Also, the results of word/morpheme TTR will be analyzed in order to investigate the amount of recycling of word/morpheme.
CHAPTER 2

METHODOLOGY

This chapter begins with a description of the materials chosen for this analysis. A standardized vocabulary list (The HSK Vocabulary List) is introduced against which the vocabulary lists from three series of CFL textbooks are compared. Lastly, the methods used for obtaining the Type/Token ratio of words and morphemes, their frequency, and the resulting data’s respective applications within this study will be discussed.

MATERIALS: THE CFL TEXTBOOKS

The data in this study is drawn from three sets of CFL textbooks: Integrated Chinese (Level 1, Part 1 and Part 2; Level 2); Chinese Link, (Elementary Chinese and Intermediate Chinese); and New Practical Chinese Reader (Book 1 and Book 2). Vocabulary items investigated in this study are taken from both the core vocabulary and the supplementary vocabulary of these texts.

These textbooks are used by students during their first and second years of Chinese language classes taken at the college level. These three sets of textbooks are so-called “comprehensive textbooks” which aim to teach the four language skills: listening, speaking, reading and writing.

Integrated Chinese (中文聽讀寫), hereafter known as (IC), written by Yuehua Liu and Tao-chung Yao, is one of the most widely used CFL textbooks in the United States. The edition investigated in this paper is the first edition, which was published in 1997. As the name of the textbook indicates, the intent of this book’s teaching philosophy is to “integrate the best of the grammar-based and communicative methodologies.”(Ning, 2001, p.45). The settings of IC and the topics in each lesson are based on students’ lives in the United States and in Mainland China. The settings and situations revolve around campus and family life, social issues, aspects of Chinese culture, and so forth. Each lesson follows IC’s standard layout which includes Dialogue/ Narrative, the vocabulary list for Dialogue/Narrative, Grammar Notes, Pattern Drills and Look and Say practice exercises. At the end of each
textbook, there is an index of vocabulary sorted by Pinyin.

_Chinese Link_ (中文天地，CL) by Wu, Yu, Zhang, and Tian (2006) is another set of CFL textbooks on the market. Compared to IC, CL includes both the Mandarin and Taiwan vocabulary items. “捷運；軟體；送簡訊”, for instance, are used in Taiwan rather than in Mainland China.

<table>
<thead>
<tr>
<th>Mandarin</th>
<th>Taiwanese</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>地鐵</td>
<td>捷運</td>
<td>MRT</td>
</tr>
<tr>
<td>軟件</td>
<td>軟體</td>
<td>software</td>
</tr>
<tr>
<td>發簡訊</td>
<td>送簡訊</td>
<td>send text messages</td>
</tr>
</tbody>
</table>

The layout of the CL series of textbooks is organized by the following order: vocabulary items, Language Link (including sentence patterns etc.), Dialogues/narratives, Language Notes, Grammar, Idiom Story and Culture Link. Also, there are separate vocabulary indexes at the end of each textbook, which are organized by lesson number, English translation, character and Pinyin.

Last but not least, _New Practical Chinese Reader_ (新實用漢語課本，NPCR), written by Zhang et al. (2004) is one of the recommended CFL textbooks by Hanban (Chinese Language Council International), which is a non-governmental and non-profit public institution affiliated with the Chinese Ministry of Education. This series of textbooks is used by students who study Chinese in Mainland China as well as in universities and language institutes in the United States. Its predecessor, Practical Chinese Reader (實用漢語課本), is one of the earliest TCFL textbooks that received widespread recognition and has been republished in many editions since 1981.

**The Reference Vocabulary List**

Besides the three series of textbooks (_Integrated Chinese, Level 1 and Level 2; Chinese Link, Elementary Chinese and Intermediate Chinese; and New Practical Chinese Reader, Book 1 and Book 2_), this study also uses a reference vocabulary list, 中國漢語水平考試詞彙大綱 (Chinese Proficiency Test Vocabulary Guideline: 8000 words), as shown in Table 2. This reference list reflects the word distribution in the natural
Table 2. Chinese Proficiency Test Vocabulary Guideline a Dictionary of Chinese Usage: 8000 words

<table>
<thead>
<tr>
<th>Grade</th>
<th>Number of vocabulary items</th>
<th>甲級詞 (Grade A)</th>
<th>乙級詞 (Grade B)</th>
<th>丙級詞 (Grade C)</th>
<th>丁級詞 (Grade D)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1033</td>
<td>2018</td>
<td>2202</td>
<td>3569</td>
<td>8822</td>
</tr>
</tbody>
</table>

language mostly and thus is used to measure the systematicality and the grading of vocabulary in the CFL textbooks.

This vocabulary list is designed for HSK (Chinese Proficiency Test), the so-called “Chinese TOEFL”. It is referred to as the HSK vocabulary list in this thesis henceforth. The HSK vocabulary list includes the vocabulary items that elementary, intermediate and advanced level CFL students are expected to know to pass the HSK proficiency test. The HSK list is designed not only for CFL students’ preparation for the test, but also for CFL teachers to use as reference.

As shown in Table 2, there are 8822 high-frequency words chosen for the HSK vocabulary list. The words are divided into four levels: 甲, 乙, 丙 and 丁 (Grade A, B, C and D). Grade A is the most frequently used group of words while the Grade D group is comparatively the least used. The words are collected from different genres. Besides frequency, learner needs are also taken into consideration when categorizing the words into their respective grades. Take 亦 (classical Chinese; also) and 議會 (congress) for example. Ma (2008) points out that in the Language Corpus from the Center for Chinese Linguistics Peking University, there are 29,194 entries for 亦; but only 10,312 entries for 議會. Also, in 《現代漢語頻率詞典》 (Institute of Language Teaching and Research, 1986) the token number of 亦 is 35 while the token number of 議會 is 22. Because 亦 is not used in contemporary Chinese any longer and thus not that useful for second language learners, even though the frequency of 亦 is higher than 議會, 議會 is categorized into Grade C while 亦 is assigned under Grade D.
INSTRUMENTS AND PROCEDURES

The measurement of simple frequency and type-token ratio (TTR) count are used in the statistical analysis. Tokens are the total number of words in a text while types only include the total number of different words (Retherford, 2000). The TTR is originally designed to measure the variety of one’s word use in written and spoken language. However, in this research, the TTR is used to measure the repetition/recycling of words and morphemes in the vocabulary system. The value of TTR is to divide the number of different words by the total number of words.

\[
\text{TTR} = \frac{\text{Type}}{\text{Tokens}}.
\]

The higher the TTR is, the less the morpheme/word is repeated/recycled and students therefore have a less chance to be exposed to the vocabulary item.

The Chinese Annotation Tool (CAT), written by Erik Petterson, is utilized to segment sentences taken from the investigated textbooks into words for the purpose of measuring word TTR. The link of CAT is http://www.mandarintools.com/ and the mirror on San Diego State University is http://www-rohan.sdsu.edu/dept/chinese/.

As there is a printing convention for Chinese characters in which no blanks are left between words, first and foremost one needs to segment sentences into words in order to obtain the TTR from the Chinese texts.

In the data used for TTR, the overall word tokens in IC (token number: 11,769) are greater in number than the ones in CL (token number: 8,696). Thus, the texts from IC are truncated (controlled IC token number: 8,924) to guarantee the token counts are constant. The token counts are kept constant for all texts to ensure the comparability of the TTR, because the larger the token counts, the lower the TTR. In other words, with more tokens there will be more repetitions.

Also, to obtain the number of morphemes for words, the Excel “Text to Columns” function under “Data” was used. After selecting the vocabulary items, one chooses “Data” and then “Text to Columns”. Under the option of “original data type”, “fixed width” should be selected and then break lines can be created among morphemes in the vocabulary items. These break lines separated the words into their individual morphemes. Thus, after the abovementioned steps, the morphemes of vocabulary items are distributed into individual cells. The function of “COUNTA” is used again to count the token number of morphemes. In
sum, the number of morphemes was calculated by using “Covert text to Columns wizard” and “COUNTA”.

Each vocabulary item from the list of investigated textbooks is typed into an Excel file, as is the HSK vocabulary list from the vocabulary guide. The total number count of each word/morpheme in the investigated textbooks is calculated by the Excel function “COUNTA”. For the purpose of comparison between two lists of words, a utility that can locate the repeating word entries between the two columns in Excel files is used. The other source of the current data is the texts. The texts of CL is typed into a Word file; while the texts of IC is gained from http://www-rohan.sdsu.edu/dept/chinese/ by Dr. Zheng-sheng Zhang. Besides the abovementioned data collecting methods, all the charts are made by using the “Chart” function.

Besides the functions from Excel, the freeware Textstat was used to obtain the number of word/morpheme types and word/morpheme tokens. Textstat is written by Matthias Hüning and its website is given below.


By adding the Word files created by Excel, Textstat shows the type and token number of morphemes or words at the status line at the bottom.
CHAPTER 3

DATA ANALYSIS

This chapter will tackle the two research questions: (1) how do the textbooks vary from each other; and (2) is there systematicity in the selection of vocabulary items found in the CFL textbooks.

VOCABULARY PRESENTATION

In most CFL textbooks, there are two places to list vocabulary items: one such place is in each lesson where the vocabulary is first introduced; the other is in the vocabulary index. The vocabulary index appears at the end of the textbook to help locate a word. A summary of the format of the vocabulary lists in each lesson, in each of the three sets of CFL textbooks is shown in Table 3.

As shown in Table 3, the three series of textbooks have different ways to present vocabulary. Both CL and IC provide simplified and traditional scripts for the vocabulary lists; however, no traditional scripts are provided for the vocabulary lists in NPCR. This is due to the fact that NPCR is mainly developed and used within Mainland China, where simplified characters are used as the standard writing system.

There are also differences in the organization of example sentences and phrases given for the vocabulary items. CL does not have example sentences or phrases listed for the vocabulary items. On the contrary, NPCR provides example phrases or sentences for every entry. IC, however, prefers to pick and choose amongst its vocabulary items thereby singling out the difficult words for which they provide example. Compared to NPCR, IC explicitly makes the key words stand out. Assistance is provided for particular words when the contextual information is considered crucial.

The vocabulary items chosen for exemplification in IC are mostly verbs, adjectives and adverbs, but rarely nouns. Nevertheless, there are exceptions where the nouns, as new vocabulary items, are provided with example sentences or phrases in IC. The purpose of doing so may have been to recycle the previously learned vocabulary items in order to
Table 3. Comparison of Vocabulary List in Each Lesson for the Three Textbooks Series

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>NPCR</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplified Script</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Traditional Script</td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Pinyin</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>English Equivalent</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Part of Speech</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Example Phrase/Sentence of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Given Word</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(key words)</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>(every word)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proper Noun</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supplementary Words</td>
<td>√</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Difficult Character list</td>
<td>√</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

explain the new vocabulary items, thereby ensuring that students have more of a chance to encounter older words. For instance, the new vocabulary 叔叔 (uncle) (Lesson 14 Book 2, IC) is reinforced by the other kinship words learned before: 爸爸 (father) from Lesson 2, Book 1 and 弟弟 (younger brother) from Lesson 2, Book 1.

叔叔：爸爸的弟弟

Pinyin: shushu: baba de didi

Translation: Uncle: Father’s younger brother

NPCR provides a larger amount of examples for vocabulary than the other two textbooks. It commonly provides more than one example for nearly each vocabulary item on its lists. In other words, when compared with the other two series of textbooks, NPCR use the presentation of vocabulary as both an introduction of new words and a learning resource for CFL students to refer to when they study by themselves.

The last thing worth noticing in Table 3 is the section of Proper Nouns (i.e. people’s names, places’ names). In IC and CL, there are Proper Nouns provided at the end of the
vocabulary list in each lesson; while there are none in NPCR. NPCR does not separate them from the common vocabulary items. Because there is no form difference in Chinese proper nouns as in English (i.e. White House and white house), the merit of marking proper nouns is to help students locate them in the texts and distinguish them from other words.

Table 4 demonstrates both the similarities and differences of vocabulary index in the three sets of CFL textbooks. Among the three series of textbooks, the simplest format of a vocabulary index is the arrangement IC chooses, that is, a list of vocabulary items categorized according to Pinyin’s alphabetic order. Compared to IC, both NPCR and CL have their indexes categorized by lesson, Pinyin, and character as well as Pinyin’s alphabetic order. CL even provides a list in which CFL students can look up Chinese words’ via English translation.

Table 4. Comparison of Vocabulary Index in Each Lesson for the Three Textbooks Series

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>NPCR</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>By Lesson</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Pinyin Index</td>
<td>✓</td>
<td>✓(new words)</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>✓(supplementary words)</td>
<td></td>
</tr>
<tr>
<td>English Index (by English Equivalent)</td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Character index in the Character Book (by lesson or Pinyin)</td>
<td>✓(by Pinyin)</td>
<td>✓(by lesson)</td>
<td></td>
</tr>
</tbody>
</table>

**VOCABULARY EXERCISES**

In addition to examining the varying vocabulary presentations and vocabulary indexes, it is also important to learn how the textbook developers design vocabulary exercises. Table 5 demonstrates the vocabulary exercises from the three series of textbooks. As can be seen, CL has the most comprehensive vocabulary exercise; while, surprisingly,
Table 5. Comparison of Vocabulary Exercise in Each Lesson for the Three Textbooks Series

<table>
<thead>
<tr>
<th>Exercise</th>
<th>IC</th>
<th>NPCR</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Make up words</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Semantic grouping</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Translation</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Transformation (traditional and simplified)</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Radical knowledge</td>
<td></td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>Collocation</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character recognition</td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

there is no practice for vocabulary in NPCR.

Vocabulary exercises aim to highlight both internal structures and external context. Therefore, CFL students can be exposed to the language features of vocabulary multiple times thereby consolidating students’ vocabulary learning. Also, vocabulary exercises function to emphasize the internal structures as in the following: try to use the following characters to make words, phrase and then sentences (Wu et al., 2006, Chinese Link Elementary, p.118). In this practice, the given words are new in the lesson. The students need to bring forth their previously learned vocabulary to make words, phrase and sentences. An example is shown below:

New word: 課 (lesson)

Make word: 英語課 (English lesson)

Make phrase: 上課 (taking lessons)

Make sentence: 我上中文課 (I take Chinese lesson.)

In addition to the meaning-based vocabulary exercises shown above, there also are exercises designed to focus more on external context. One of the vocabulary exercises in CL, to create/match phrases by (utilizing the) given words, is widely used among children who learn Chinese as their first language in elementary school. An excerpt of a lesson in Volume
1 from Yu Wen, which is chosen by the Education Bureau for usage in most elementary schools in China, is shown below.

This kind of vocabulary exercise highlights the context information (collocation) of the words, in this case especially focusing on the adjectives and their nouns. This exercise demonstrates the kinds of nouns that can be described by particular adjectives.

IC also has vocabulary exercises in the form of collocation pattern drills. However, students only mimic the sample sentence patterns. Compared to CL, it is less creative because students are not actively producing the language. An example from IC is shown blow.

給 (gei_) as a position

我    給你    介紹一下
李老師    ____    看她的書
小高    ____    看他爸爸媽媽的照片
小王    ____    聽中國音樂
她    ____    介紹一個朋友
高小音    ____    喝中國啤酒

Translation

I    to you    introduce.
Teacher Li    ____    show her book.
Gao    ____    show his parents’ photo
Wang    ____    listen Chinese music.
She    ____    introduce a friend.
Gao Xiaoyin    ____    drink Chinese beer.

Note: Adapted from p.120, Lesson 6 making appointments, Level 1 Part 1, IC

NPCR does not include any vocabulary exercises in its textbooks except the instruction for characters. On the other hand, the other two textbooks provide a large amount of practice of both vocabulary and characters as, for example, in: Try to use the following
characters to make words, phrases, and then sentences and Write out the radical of each of the following characters, then give some examples from the text of characters having the same radical.

All in all, the CFL textbooks chosen in this thesis are very different from each other in the presentation of vocabulary and vocabulary exercises.

In order to investigate the systematicality of the presentation of vocabulary in the CFL textbooks, two factors are discussed central to this issue, including the quantity of vocabulary types and the comparison between the textbook vocabulary and the HSK vocabulary list.

**THE QUANTITATIVE ANALYSIS OF VOCABULARY IN TEXTBOOKS**

Besides the vocabulary lists and the vocabulary exercises, the vocabulary’s quantitative distribution is worth paying attention to when analyzing the differences among CFL textbooks and the systematicality of vocabulary in each of the textbook series.

The quantity of vocabulary types in each textbook is shown in Figure 2. There are evident differences in the total number of words (word types; cf. word tokens) for each textbook series. IC has the largest amount of words among the three sets of CFL textbooks. There are approximately 2,000 new vocabulary items selected in the first two-year period of Chinese learning. NPCR has the fewest, which is as few as half of the number of vocabulary items in IC.

![Figure 2. Total number of the vocabulary items (type) from the investigated textbooks.](image)

Besides the total number of vocabulary items in each textbook series, the distribution of vocabulary in the first and the second year study of Chinese is also analyzed. Figure 3
Figure 3. The number of the vocabulary items (type) distribution in 1st and 2nd year study of Chinese.

demonstrates that, for each level of CFL study, the number of vocabulary items increases as the students progress. In other words, the number of vocabulary types CFL students learn in the second-year study is more than that in the first year. Even though the total number of vocabulary items is different in the investigated textbooks, the rate of increase of the number of vocabulary types in the 1st and 2nd year study of Chinese ends up being the same, as shown in Figure 4.

COMPARISON OF WORD TYPES BETWEEN THE TEXTBOOK VOCABULARY AND THE HSK VOCABULARY LIST

In order to investigate the systematicity of the vocabulary in the CFL textbooks, the HSK vocabulary list is used in this study. The distribution of the four grades (Grade A, B, C and D) of HSK vocabulary in each series of the investigated textbooks, is also discussed in this section. The HSK includes words from not only the elementary level but also the advanced ones. In this section, the distribution of the four grades of HSK vocabulary and the textbook vocabulary are compared. Second, since the textbooks chosen for this study are taken from elementary and intermediate levels, Grade A words are emphasized and investigated closely.

The Distribution of the Four Grades in the Textbooks

An overview of the vocabulary distribution of HSK vocabulary and textbook vocabulary is shown in Figure 5. According to Figure 5, the HSK vocabulary list demonstrates that the number of less frequently used words is more than the number of more frequently used words; however, the least amount of words in the HSK vocabulary list (Grade A) can cover 86% of Chinese words and phrases. On the other hand, the distribution of vocabulary within the investigated textbooks turns out to go in the opposite direction: the largest numbers of vocabulary items early on are taken from Grade A.

![Figure 5. The distribution of the four levels of the vocabulary items in the three textbook series.](image-url)

The inverse trend shown in Figure 5 can be explained by the HSK vocabulary list’s use as a standardized list for learners across all levels, while the textbooks investigated in this
study are only designed for elementary and intermediate level college students. Therefore, for the elementary and intermediate level textbooks, it is logical that larger amounts of frequently used words are focused upon rather than the lesser used ones. On the other hand, the HSK vocabulary list reflects the vocabulary as used in natural speech that CFL students might not have access to until later on. Therefore, even after studying at the intermediate level of CFL, the vocabulary which students must master is not equivalent to vocabulary distribution found in the HSK vocabulary list.

The Distribution of the Four Grades Words at Each Level

The four grades from the HSK vocabulary list are distributed in a similar pattern in each textbook. Figures 6, 7 and 8 illustrate that the total number of Grade A words are the majority of all vocabulary items in each of the textbook series. When students move to the next level, it is noticeable that number of Grade A words decline while the vocabulary items found in Grade B increases. In CL and IC, the number of Grade B vocabulary items eventually exceeds Grade A and becomes dominant (Figure 6 and Figure 7) while Grade A words make up the majority of vocabulary items across both the elementary and intermediate levels in NPCR (Figure 8). As for Grade C and D, they show surprisingly similar patterns.

Figure 6. The distribution of four grades’ vocabulary in CL.
along with the textbooks’ advancement. The number of Grade C and D words is both less than the number of Grade A and B; and they increase very slightly when the textbooks become more advanced.

In spite of the similar patterns of distribution, there are differences among the textbooks. According to the data from CL shown in Figure 6, the number of Grade A and Grade B words in the elementary levels are distributed relatively closely to each other; meanwhile, the other two series of textbooks (Figure 7 and Figure 8) demonstrate that there are quite evident differences in the number of vocabulary items between Grade A and Grade B. In other words, CL selects more Grade B words at the elementary level than those in IC and NPCR. Meanwhile, there are fewer Grade A words adopted for elementary level students.
in CL. The distribution suggests that CL vocabulary is more difficult than the other two textbooks because a larger number of the vocabulary items are from the less frequently used Grade B list than from Grade A.

In sum, among the three series of textbooks investigated in this study, Grade A words are predominant at both elementary level and intermediate level. Along with the progress of the textbooks, the number of Grade B, C and D increases. Furthermore, all of the textbook series involved in this study demonstrate changes in the number of vocabulary items at different levels.

**Grade A Words in the Textbooks**

Grade A words in the HSK list represent the most frequently used words in natural speech. Figure 9 illustrates the difference between the number of Grade A words in the HSK vocabulary list and the number of Grade A words found in the three sets of textbooks. The result shows that none of the textbooks, after elementary and intermediate level, completely covers the HSK words of Grade A.

![Figure 9. Comparison of the three textbook series with the HSK vocabulary list.](image)

Ideally, after elementary and intermediate level study, the words found in HSK Grade A should all have been introduced to CFL students. Nevertheless, the results show that the vocabulary selection from neither level does well at including the words from Grade A, which is supposed to be the most frequently used words. Among the three series of textbooks, IC includes more Grade A words than the other two sets and it only covers approximately 2/3 of the Grade A words from the HSK vocabulary list. Therefore, the
vocabulary items in the textbooks do not provide CFL students with a sufficient number of high frequency words.

Ideally, after elementary and intermediate level study, the words found in HSK Grade A should all have been introduced to CFL students. Nevertheless, the results show that the vocabulary selection from neither level does well at including the words from Grade A, which is supposed to be the most frequently used words. Among the three series of textbooks, IC includes more Grade A words than the other two sets and it only covers approximately 2/3 of the Grade A words from the HSK vocabulary list. Therefore, the vocabulary items in the textbooks do not provide CFL students with a sufficient number of high frequency words.

To generalize, the data reflect the problems of the CFL vocabulary system in the textbooks. At the elementary and intermediate level CFL study, the vocabulary items have not been shown a similarity in distribution to the HSK list. On the other hand, the three textbook series do not include the frequently used words successfully either, for instance, the Grade A words. As a result, the textbook vocabulary is not systematic due to the fact that the current vocabulary distribution does not exhibit the word distribution in the HSK vocabulary list nor manage to involve the most frequently used words.

**Vocabulary Repetition/Recycling**

In order to investigate the issue of vocabulary recycling, it is useful to look at the ways in which the different textbooks recycle the meaning units in vocabulary, words and morphemes. As was talked about earlier, the TTR indicates the repetition of words/morphemes. The higher the TTR, the less the word/morpheme is repeated in the texts; while the lower the TTR, the more the meaning unit is recycled and thus has a greater chance of helping students’ retention. Both word TTR and morpheme TTR are discussed in this thesis. The data of TTR from in CL and IC are shown in Table 6.

<table>
<thead>
<tr>
<th></th>
<th>CL</th>
<th>IC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word TTR</td>
<td>26%</td>
<td>27%</td>
</tr>
<tr>
<td>Morpheme TTR</td>
<td>24%</td>
<td>24%</td>
</tr>
</tbody>
</table>
As can be seen in the table above, the results of word TTR from these two series of textbooks are very close. The word TTR in CL is 26%; while the word TTR of IC is 27%. The word TTR of the two series of textbooks shows that each vocabulary item repeats less than four times in the two-year study of Chinese. The data clearly indicates that the repetition of vocabulary items in the investigated textbooks is rather insufficient.

It also must be noted that the morpheme TTR turned out to be very close to the word TTR. Both CL and IC have a morpheme TTR of 24%. In other words, each morpheme appears approximately around four times across the elementary and intermediate textbook levels of IC and CL.

In addition to the morpheme TTR and the word TTR across the whole of the texts, the morpheme TTR in the textbook vocabulary lists reflects a low repetition frequency of morpheme as well. That is, each morpheme appears only slightly more than each word does (100%) in the vocabulary lists. According to Table 7, the morpheme TTR of the vocabulary lists found in the three sets of textbooks are shown to be 71% and 80% respectively.

**Table 7. Morpheme TTR in the Vocabulary Lists in the Textbooks**

<table>
<thead>
<tr>
<th></th>
<th>IC</th>
<th>NPCR</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morpheme TTR</td>
<td>71%</td>
<td>80%</td>
<td>80%</td>
</tr>
</tbody>
</table>

Both morpheme TTR and word TTR across the texts and morpheme TTR of the vocabulary lists point out two issues of CFL vocabulary teaching: (1) the vocabulary teaching approach used in the investigated textbooks is word-based; (2) the shared morphemes are not emphasized when choosing vocabulary for the CFL textbooks. The similarity in percentage between word TTR and morpheme TTR in the texts across the different textbooks as well as the morpheme TTR in the vocabulary lists indicates that the textbook writers consider the new morphemes as new vocabulary items, because the occurrence of a recycled morpheme is approximately equal to the frequency of a word’s appearance in the texts or vocabulary lists. Moreover, as there is little difference between morpheme TTR and word TTR, the morphemes are not emphasized when choosing vocabulary for the CFL textbooks. As already stated, the textbooks do not emphasize the significance of the morphemes in the compound words, and thus do not recycle morphemes adequately.
All in all, morpheme TTR and word TTR from the investigated textbooks in the texts and in the vocabulary lists show that morpheme/word recycling is insufficient. The occurrence of recycled morphemes/words is far less than is sufficient to help CFL students acquire vocabulary. Also, the data show that the connections among compounds, which are the shared morphemes, are ignored by the textbook’s designers.

**Lack of Integration of Semantic Networks**

The connections among vocabulary items are referred to as semantic networks. The use of semantic networks leads to the easier activation of vocabulary items and provides more insight to the semantically shared meanings of words and thus is more beneficial than simply memorizing words individually. The data in this thesis reveal there is a lack of integration of semantic networks in these analyzed textbooks. That is to say, the vocabulary items chosen for CFL textbooks do not form complete semantic networks. Take 下午 (afternoon) for example. 下午 (afternoon) appears in Lesson 10, Book 1, CL. However, 上午 (morning) does not appear anywhere in the selected textbooks. Also, CL does not include 涼快 (pleasantly cool) as one of the vocabulary entries even though it has the vocabulary entry 暖和 (warm) as IC does. The examples show that there are gaps in the semantic networks in the textbooks.

**Sequencing of Multiple Meanings**

There are occasions in which the derived meanings of words are introduced earlier than the basic meanings in the textbooks. Sometimes, the basic meanings of the words are not even introduced to students in the textbooks at all. Carelessly sequencing of words could affect students’ understanding of the basic meaning of words as well as hinder their constructing the semantic connections among related words. Take 碟 as an example. 碟’s most basic meaning is *plate, dish, or saucer*. One of its derived meanings is *DVD* or *disc*. The word is first introduced in (Lesson 11, Level 1 Book 2, of IC) in the form of the compound 看碟 (to watch discs or DVDs). However, the original meaning of 碟 as *plate, dish, saucer* has not been introduced before the meaning of *disc* is learned. Students, therefore, will have no idea of the full derivational process of 看碟. Since CFL students do not know how the meaning of a compound is originally derived from the more basic
meanings, introducing the extended meaning before the basic may lower the possibilities of students’ understanding the association of a word’s extensive meanings and its basic meaning (e.g., *DVD* and *plate*).
CHAPTER 4

CONCLUSION

This thesis set out to explore the vocabulary acquisition issue in teaching Chinese as a foreign language. The primary purpose of this study was to investigate the systematicity of vocabulary selection in the CFL textbooks as well as vocabulary instruction practice. In particular, answers to the following research questions were addressed:

Research Question 1:

How do the textbooks differ from each other from the perspective of vocabulary presentation and vocabulary exercises?

Research Question 2:

Is the presentation of vocabulary in the CFL textbooks systematic?

In addition, the thesis suggests several teaching techniques that aim to assist CFL students with their vocabulary acquisition, such as the recycling of vocabulary and making connections among new words.

RESEARCH QUESTION 1

How do the textbooks differ from each other from the perspective of vocabulary presentation and vocabulary exercises?

One of the first things analyzed in this study were the differences in vocabulary organization utilized by three commonly used textbooks. Particularly close attention was paid to the different methods of presenting vocabulary and the arrangement of vocabulary exercises. From the perspective of vocabulary presentation, NPCR does not supply traditional scripts (traditional characters); nor does it group proper nouns into a distinctive separate section from other vocabulary items. On the other hand, NPCR, along with IC, provides example phrases and/or sentences for the new words; conversely, there are no examples given for the vocabulary items found in CL. It is important to note that NPCR is mainly used in Mainland China where the simplified scripts were adopted as the standard writing system. Also, compared to the other two series of textbooks, NPCR provides plenty
example sentences and phrases for each vocabulary item. IC only provides examples for the key words. CL, on the other hand, does not accommodate any examples for the new words. In some textbooks, however, there is another place for presenting new vocabulary items, and that is a vocabulary index. IC provides the index of the vocabulary items while CL and NPCR include the indexes for vocabulary items as well as for new characters. CL also has an index of the Chinese vocabulary’s English translation.

When it comes to the vocabulary exercises, CL includes the most comprehensive vocabulary practice among the three series of textbooks used in this thesis. Surprisingly, NPCR does not provide any exercises for vocabulary in its textbooks.

**RESEARCH QUESTION 2**

*Is the presentation of vocabulary in the CFL textbooks systematic?*

The data from the currently investigated textbooks shows that the vocabulary selections are not systematic in a few ways. First of all, each series of CFL textbooks consists of its own different size of vocabulary items. When looked at in this manner, the differences between the studied textbooks and their respective vocabulary lists are quite evident. However, all the CFL textbooks in this study show an identical tendency for the total number of vocabulary types to increase from the first year books to the second, i.e. there are more words introduced in the second year of study than the first. As there is a tendency for increased numbers of vocabulary items to be introduced in higher levels, a standardized vocabulary list, the HSK vocabulary list, was used to examine the grading of the vocabulary items found within the textbooks. The HSK is divided into four vocabulary categories, otherwise known as Grades, Grade A being the most commonly used words in Chinese, and Grade D being relatively less frequent. The findings of this comparison between the textbooks and the HSK vocabulary list show that (1) within the elementary and intermediate level textbooks there is tendency for a an inverse rate of Grade A through D distribution; i.e., there are a significant number of Grade A words but decreasingly fewer C through D used in the textbooks ; (2) the number of Grade A words found in the textbooks (elementary and intermediate) do not exhaustively cover the whole HSK Grade A list; and (3) there is not much of an increase in the number of Grade C and D words across the elementary and intermediate levels of the textbooks; i.e., the rate in which new C and D words are introduced
is relatively low and increases only slightly.

To restate more simply what was explained above, the number of words in the HSK vocabulary list increases from Grade A to Grade D while the distribution of the vocabulary of the same four Grade levels is declining in the textbooks. In other words, the list and number of words the HSK proficiency test expects CFL learners to master increases over time, while new words taught by the texts declines over time. The opposite directions of vocabulary distribution between the textbook vocabulary and the words found in natural speech can be explained by the difficulty levels of the investigated textbooks. The elementary and intermediate level textbooks tend to focus more on the basic and frequently used vocabulary items (Grade A and B) rather than harder and less frequently used words (Grade C and D).

Even though the Grade A list consists of the most frequently used words in Chinese, the vocabulary selections found in each of the series of textbooks do not successfully include all of the words from Grade A. Grade C and D words are found the least and there is never much of a change in their overall number. Thus, the textbook vocabulary within the first two years of study does not adhere to the same general trends of word distribution in the HSK vocabulary list.

In sum, vocabulary selection is not systematic in the CFL textbooks because of the evident difference among the number of vocabulary items in the textbooks. Also, the vocabulary selection does not manage to include the most basic and frequently used words; nor is it consistent with the word distribution found in the standardized HSK vocabulary list.

Aside from the issues relating to the lack of systematicity, this thesis also reveals the low repetition rate of vocabulary items and the little emphasis given to shared morphemes. The data show that word repetition within the textbooks is rather low. Also, the word TTR from the whole set of texts (dialogues and narratives) is very close to the morpheme TTR. Thus it seems that the reusing of morphemes is very low, suggesting there is little difference between word and morpheme TTR in the textbooks. This assessment is verified by the high morpheme TTR in the vocabulary list, which restates again that the textbooks treat morphemes and words without much difference.

Last, an investigation of the textbooks’ morpheme and word frequencies shows that semantic networks are not all well developed. This was determined by the lack of word
association integration (such as “warm” and “cool” in Chapter 3) within the semantic networks and the sequencing of multiple meanings of words, such as DVD and plate. These factors all show the shortcomings of the semantic networks that exist in the CFL textbooks.
CHAPTER 5

PEDAGOGICAL IMPLICATIONS

Based on the findings of the present study, the textbook writers and CFL teachers could improve vocabulary teaching by paying closer attention to the systematicity of vocabulary selection. Also, several teaching techniques are advocated in this thesis, including paying greater attention to the individual morphemes when teaching, making use of shared morphemes among different compounds and increasing vocabulary recycling.

SYSTEMATICITY OF VOCABULARY SELECTIONS

To improve the systematicity of vocabulary selection in the CFL textbooks, first of all, a set of well-selected and sufficient vocabulary items should be chosen by the textbook writers. It is reasonable that each textbook have its own preferred size and choice of vocabulary items; however, the vocabulary selection should reflect the word distribution in the natural use of Chinese. Moreover, in the elementary and intermediate levels, the CFL textbooks should include the most frequently used words as much as possible. At subsequent levels of difficulty, however, a more explicitly graded vocabulary should be selected to include words taken from the Grade C and D lists.

THE USE OF SHARED MORPHEMES IN DIFFERENT COMPounds

As mentioned in Chapter 1, Chinese does not have cognates that are shared by Indo-European languages such as English. The lack of shared cognates is one of the most important factors of why learning Chinese vocabulary is hard for speakers of Indo-European languages. Therefore, the language internal connections between shared morphemes that make up different compounds have to be relied upon to connect what the student already knows to what is newly presented. Once students are more aware of the morphemes that form the compounds, they can make use of the knowledge of recurrent morphemes to build associations among words.
A sample vocabulary lesson is provided below to show how many vocabulary items can be interrelated by making use of shared morphemes. The sample vocabulary list is taken from IC lesson 11, level 1 part 2. The list is expanded, however, to reflect the significance of shared morphemes in CFL vocabulary teaching and learning. According to Table 8, when the morpheme-based teaching approach is applied, students have morpheme and semantic connections to 39 related words. This is striking when compared to the 18 vocabulary items shown on the vocabulary list from the textbook.

**Table 8. A Modified Vocabulary List of Lesson 11 Level 1 Part 2, Integrated Chinese**

<table>
<thead>
<tr>
<th>Vocabulary list</th>
<th>Shared Morpheme</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>天氣/weather</td>
<td>天day</td>
<td>今天，明天，後天，昨天，前天，白天 today, tomorrow, the day after tomorrow, yesterday, the day before yesterday, daytime</td>
</tr>
<tr>
<td>下雨，下雪/raining, snowing</td>
<td>下down</td>
<td>下載download</td>
</tr>
<tr>
<td>預報/forecast/forecast</td>
<td>預pre-, beforehand</td>
<td>預習，天氣預報，節目預報 preview, weather forecast, program forecast</td>
</tr>
<tr>
<td>員/person</td>
<td>員person</td>
<td>天氣預報員，售貨員，店員，球員 weather forecaster, salespeople, salespeople, soccer player</td>
</tr>
<tr>
<td>春天，夏天，秋天，冬天/spring, summer, fall, winter</td>
<td>天day</td>
<td></td>
</tr>
<tr>
<td>假/holiday</td>
<td>假holiday</td>
<td>寒假，暑假，春假 winter break, summer break, spring break</td>
</tr>
<tr>
<td>滑冰，滑雪/skiing, snowboarding</td>
<td>滑slide</td>
<td>輪滑，滑蓋手機，速滑，滑草 skateboarding, slide phone, speed skating, grass skating</td>
</tr>
<tr>
<td>約/date</td>
<td>約date</td>
<td>約時間，約會 make an appointment, have a date</td>
</tr>
</tbody>
</table>

*(table continues)*
### Table 8. (continued)

<table>
<thead>
<tr>
<th>Vocabulary list</th>
<th>Shared Morpheme</th>
<th>Expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>(怎麼)辦</td>
<td></td>
<td>怎麼了，怎麼樣，怎麼+Verb</td>
</tr>
<tr>
<td>how…</td>
<td></td>
<td>What happened, what’s up, how+verb</td>
</tr>
<tr>
<td>公園</td>
<td>公</td>
<td>公共汽車</td>
</tr>
<tr>
<td>park</td>
<td>public</td>
<td>bus</td>
</tr>
<tr>
<td>公園</td>
<td>園</td>
<td>動物園，公園門票</td>
</tr>
<tr>
<td>park</td>
<td>park</td>
<td>zoo, park pass</td>
</tr>
<tr>
<td>面試</td>
<td>試</td>
<td>考試，口試，筆試</td>
</tr>
<tr>
<td>interview</td>
<td>test</td>
<td>test, oral test, written test</td>
</tr>
<tr>
<td>看碟</td>
<td>碟</td>
<td>碟子，盤子，光碟，光盤，租碟，飛碟</td>
</tr>
<tr>
<td>watch DVD</td>
<td>DVD/plate</td>
<td>plate, plate, DVD, rent a disc, UFO</td>
</tr>
<tr>
<td>TOTAL: 18</td>
<td>TOTAL: 12</td>
<td>TOTAL: 39</td>
</tr>
</tbody>
</table>

The sample vocabulary lesson shown in Table 8 highlights the morphemes of the new vocabulary entries as well as involving the previously learned morphemes/words. Take 今天 (today), 明天 (tomorrow), and 天氣 (weather) for example. They all include a shared morpheme, 天 (day), in the words. The first two words that are learned, today and tomorrow, are used to connect the new vocabulary item, 天氣 (weather), in this lesson by emphasizing the morpheme of 天 (day). In sum, the shared morphemes between the new words and the learned words emphasize the connections, shared morphemes, among compound words. Also, it relates the learned words with the new vocabulary items and increases the chance to recycle vocabulary items.

Nevertheless, semantic meanings of the constituent morphemes of words are not always so transparent that students can make up words by this technique. Consider 打飯 and *雨日 as examples. 打飯 is an actual word in which the morphemes are not semantically transparent. The word means “to get food (from the cafeteria)” rather than the literal meaning indicated by the two morphemes, to beat the rice. The morphemes of *雨日 are semantically transparent but their combination does not produce an actual word. Similarly, 雨 is rain and...
is day, and these two morphemes, when added together, are by no means the correct way to say rainy day.

Therefore, the advocated teaching method is to introduce the new words to students first and then analyze their constituent morphemes. This teaching technique helps students understand the new words better than simply presenting the words to students and meanwhile avoids the misunderstanding of the words with literal meanings that are not semantically transparent.

**Making Connections Among Words**

In addition to highlighting the significance of morphemes, making connections among words is an important teaching technique. There are some issues that deserve the instructors and textbook developers’ attention. First, when selecting vocabulary, CFL textbook writers need to pay close attention to the completeness of the semantic networks based on the vocabulary choices. The semantic networks should be presented in complete sets so that students can make connections among the words by their meanings.

Also, when arranging the vocabulary, the basic meanings of morphemes should be introduced before the derived or more complex meanings so that CFL students have a clear understanding of the vocabulary.

Last, since compound words make up the majority of words in Chinese and their meanings exhibit many shared features with their constituent morphemes, shared morphemes become an important element for building connections among words in order to assist students’ vocabulary learning.

**Vocabulary Recycling**

The recycling of each vocabulary item should be emphasized more in CFL vocabulary acquisition. Multiple exposures to a vocabulary item gives CFL students more chances to learn the words. In order to improve vocabulary recycling, the texts can adopt the same vocabulary again and again in different contexts, since contextual information, in addition to repetition, helps students understand and learn new vocabulary (Nation, 2001). Secondly, using the learned vocabulary to explain new vocabulary items in example phrases or sentences arranged alongside the vocabulary lists would increase the frequency in which students encounter the vocabulary as, for example, in the new word presentation from
Chapter 1 叔叔(uncle)：爸爸(father)的弟弟(brother). Also, making use of vocabulary exercises can provide students with the opportunity to recycle the vocabulary items in an active way. Last but not least, when recycling vocabulary, the productive morpheme, should be pointed out and emphasized by instructors so that students are aware of and may make use of them in their vocabulary learning later on.
CHAPTER 6

LIMITATIONS AND FURTHER RESEARCH

It should be pointed out there were limiting factors on the present study. First of all, only elementary and intermediate level textbooks are included in the study. The data did not examine advanced textbooks, which will provide a more complete picture of the current CFL vocabulary issues. If possible, further research ought to investigate the vocabulary issues discussed in this study across all levels of instruction.

In addition to the possibility of looking at more advanced levels of the textbooks, the quantity of the textbooks used in this study also must be addressed. Three series, and from those series seven textbooks in total, were included in this study. It is likely that other textbooks employ different methods for dealing with vocabulary instruction. Moreover, this thesis also did not choose the newest editions of the investigated textbooks for the study. Therefore, it is possible that the textbook writers made improvement in the newer editions.

Last but not least, as the textbooks chosen in this study are focused primarily upon building a full repertoire of general language skills, it might also be worthwhile to study the vocabulary items from textbooks that specialize in specific language skills; i.e. listening and reading, since different language skills might focus on different language styles (i.e. spoken and written). As such, these various and different language styles might affect the vocabulary selection for the textbooks and therefore build specialized repertoires of language competence in the learners.
REFERENCES


