1-1-2016

Use of Verb-Noun Collocations by Advanced Learners of Chinese

Xiaolin Peng
University of Pennsylvania, xiaolin@wharton.upenn.edu

Recommended Citation

Use of Verb-Noun Collocations by Advanced Learners of Chinese

Abstract

The important role of collocations has been widely accepted in the current literature, but to date there are still relatively few studies on language learners' collocation knowledge and development within different local contexts. The current study intends to contribute to the literature by investigating the oral production of Chinese verb-noun (V-N) collocations by a group of highly proficient learners comprised of both Chinese as a foreign language learners (CFL learners) and Chinese heritage language learners (CHL learners), as compared to Chinese native speakers (CNSs). The study brings together current literature on collocation and heritage language learners both from a Western perspective and from the Chinese linguistic and sociolinguistic perspective.

Samples of spoken language data discussing both academic and non-academic topics were collected through one-on-one interviews with 10 CFL learners, 10 CHL learners and 10 CNSs. The data are analyzed both quantitatively and qualitatively to yield the following three findings: (1) There is a significant difference in using Chinese verb-noun (V-N) collocations among CFL learners, CHL learners, and CNSs. In general, CNSs produced significantly more V-N collocations in terms of both number (token) and range (type) than CFL learners and CHL learners, (2) The two different oral topics are also found to affect learners' production of collocations. All three groups used more monosyllabic V-N collocations in discussing daily topics and more disyllabic V-N collocations in discussing academic topics. Moreover, CFL learners and CFL learners exhibited both similarities and differences in applying collocations under the two oral contexts, (3) There are different categories and characteristics of collocation usage in terms of the acceptability and communicativeness of non-conventional collocations produced by learners. The discussion further analyzes several factors that tend to influence CFL learners' and CHL learners' production of collocations.

The findings of this study expand our understanding about advanced learners' knowledge and production of Chinese V-N collocations. Moreover, they also provide invaluable information for educators and practitioners who are involved in FL and HL instruction of Chinese.

Subject Categories
Education | Linguistics | Other Languages, Societies, and Cultures

USE OF VERB-NOUN COLLOCATIONS BY ADVANCED

LEARNERS OF CHINESE

Xiaolin Peng

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Yuko G. Butler, Associate Professor of Education

Graduate Group Chairperson

_____

J. Matthew Hartley, Professor of Education

Dissertation Committee

Yuko G. Butler, Associate Professor of Education

Nancy H. Hornberger, Professor of Education

Christina Frei, Adjunct Associate Professor of Education

DEDICATION

I dedicate this dissertation to my parents, Shikui Peng and Runfang Li, and to my

husband Xiaofang Yang.

ACKNOWLEDGMENT

First of all, I would like to express my deepest gratitude to my advisor and committee chair, Dr. Yuko Butler, for her constructive comments, practical assistance, and endless patience throughout the past two years

My heartfelt thanks also go to Dr. Hornberger and Dr. Christina Frei for serving as my committee members and for providing their guidance and comments on my thesis proposal, literature review and research design.

I am very grateful to my supervisor and mentor, Dr. Theresa Jen at the Lauder Institute, for her continuous encouragement and support.

I also would like to extend my thanks and appreciation to all my students, my research participants, the two native speaker raters, as well as my colleagues and friends in both Graduate School of Education and Lauder Institute, who have inspired me and encouraged me along the way.

I especially want to express my gratitude to members of my family: my parents, my husband, my two boys, for their faith and unfaltering support to my academic pursuit.

ABSTRACT

USE OF VERB-NOUN COLLOCATIONS BY ADVANCED LEARNERS

OF CHINESE

Xiaolin Peng

Yuko G. Butler

The important role of collocations has been widely accepted in the current literature, but to date there are still relatively few studies on language learners' collocation knowledge and development within different local contexts. The current study intends to contribute to the literature by investigating the oral production of Chinese verb-noun (V-N) collocations by a group of highly proficient learners comprised of both Chinese as a foreign language learners (CFL learners) and Chinese heritage language learners (CHL learners), as compared to Chinese native speakers (CNSs). The study brings together current literature on collocation and heritage language learners both from a Western perspective and from the Chinese linguistic and sociolinguistic perspective.

Samples of spoken language data discussing both academic and non-academic topics were collected through one-on-one interviews with 10 CFL learners, 10 CHL learners and 10 CNSs. The data are analyzed both quantitatively and qualitatively to yield the following three findings: (1) There is a significant difference in using Chinese verb-noun (V-N) collocations among CFL learners, CHL learners, and CNSs. In general, CNSs produced significantly more V-N collocations in terms of both number (token) and

range (type) than CFL learners and CHL learners, (2) The two different oral topics are also found to affect learners' production of collocations. All three groups used more monosyllabic V-N collocations in discussing daily topics and more disyllabic V-N collocations in discussing academic topics. Moreover, CFL learners and CFL learners exhibited both similarities and differences in applying collocations under the two oral contexts, (3) There are different categories and characteristics of collocation usage in terms of the acceptability and communicativeness of non-conventional collocations produced by learners. The discussion further analyzes several factors that tend to influence CFL learners' and CHL learners' production of collocations.

The findings of this study expand our understanding about advanced learners' knowledge and production of Chinese V-N collocations. Moreover, they also provide invaluable information for educators and practitioners who are involved in FL and HL instruction of Chinese.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1: INTRODUCTION

1.1. Background of the study

In the past decade, with the rapid development of globalization and modern technology, more and more language learners have opportunities to study abroad or immerse for longer periods in the target language. As a result, language educators see a growing number of "advanced" language learners who can use the target language proficiently, especially in terms of spoken language. Many of these learners have worked or lived in the target language country for many years and have used the target language as a primary means of communication in a wide variety of settings, ranging from daily conversations to highly professional contexts.

Despite their already highly proficient language competence, many of these learners are still quite motivated to maintain or enhance their language capacities for academic or professional level performance. Thus, a new challenge faced by language educators is how to facilitate the advancement of these high-achieving language learners by providing advanced-level language courses or content courses taught in the target language (Wang, 2010).

In the field of Second Language Acquisition (SLA), the issue of using "native-like" language proficiency as the standard for L2 language learners has generated much discussion among researchers. While some traditional SLA methods have implicitly treated L2 learners as defective by using monolingual native speakers as the sole standard for comparison (Cook, 1997), a growing number of researchers now view L2 use from a

lingua franca perspective and treat bilingual and multilingual competencies as the norm rather than the exception (Cook, 2009). As a result, the goal of L2 development should not be regarded simply as achieving "native-like" language competency, but rather as achieving communicative competency and efficiency in related language communities.

1.2. Significance of collocations in language acquisition

In order to achieve communicative competency in a second or foreign language, one central issue is the acquisition and application of a great number of prefabricated language chunks that are often referred to as formulaic sequences. It has been widely recognized that these formulaic sequences make up a large portion of both oral and written language and play a major role in language processing and use (Nation, 2001; Schmitt, 2004; Wood, 2010; Wray, 2002). Researchers have argued that the command of prefabricated and formulaic language is an essential aspect of communicative competence because it enables language users to process and produce language both fluently and accurately (Schmitt, 2004; Wood, 2010; Wray, 2002).

Among such prefabricated formulaic language, an important and frequently occurring type of expression is called collocations,   i.e., frequently co-occurring lexical combinations such as verb-noun (V-N) combinations (*make a decision*) or adjective-noun combinations (*fully aware*) (Nesselhauf, 2005). Many studies on collocations have shown that even highly proficient learners seem to have problems in using and developing collocation knowledge (e.g., Nesselhauf, 2005; Revier & Henriksen, 2006). One important claim made by researchers is that language learners tend to use "self-created" collocations that are not often used in the target language communities because L2

learners rely heavily on using learned rules to create new expressions rather than resorting to lexicalized routines (Pawley & Syder, 1983; Wray, 2002).

The extent to which such unconventional use of collocations is problematic is another issue of discussion among researchers. Many researchers (Pei, 2008; Koya, 2005; Gyllstad, 2007) have found that L2 learners' collocational knowledge is related in some way to language proficiency. Thus lack of such knowledge or deviant use of collocations may cause problems for L2 learners' overall proficiency and identification with the target language community (Wray, 2002). However, other researchers (Howarth, 1998) have argued that such unconventional use of collocations should be treated more positively as indications of learners' active application of language rules and communicative strategies. Despite these different views of learners' collocation competency, most researchers in this area agree that language users draw on a large inventory of ready-made collocations to support their spontaneous and creative language production (e.g., Ellis, Simpson-Vlach & Maynard, 2008; Erman & Warren, 2000; Hoey, 2005). Thus it is important to explore how learners process and use different types of collocations in a variety of contexts.

Up to now, most studies investigating collocational competence and development have focused on learners of English and have based their analysis on written data drawn from linguistic corpora or elicitation tasks (Granger, 1998; Nesselhauf, 2003). But, as it is widely agreed, many other languages share similar linguistic features with regard to the importance of collocations and other formulaic language.

1.3. Learning collocations in Mandarin Chinese

Mandarin Chinese, as a language spoken by more than 800 million people around the world, has continued to draw the interest of many second and foreign language learners. In the United States, more educational institutions, from elementary schools to universities, are beginning to offer Chinese as a second/foreign language to different levels and types of learners. Although the teaching and learning of Chinese as a second/foreign language is still a relatively new field, many important SLA theories have been applied to explore various aspects of learning this language.

Because the current study focuses exclusively on spoken Chinese, it is necessary to first acknowledge the complexity involved in its teaching and learning of. The main reason for such complexity is that the Chinese language has many varieties, especially in terms of its numerous dialects and the many social, cultural, historical, and political implications attached to them (Wang, 2010). Currently, Mandarin Chinese is an official language of Mainland China, Taiwan, and Singapore and is also spoken as a key minority language in many other countries, including Indonesia, Malaysia, Canada and the U.S. As a result, differences in phonetics/phonology, syntax, and vocabulary have developed in these regions. For language educators working with Chinese language learners, it is particularly challenging to decide what should or should not be included as "standard" or "native" spoken Chinese in the language curriculum. Considering the scope of the current study, Mandarin Chinese spoken in mainland China, Taiwan, and Singapore is considered the standard variety, as almost all the participants in this study either have family backgrounds from these regions or have studied extensively in these regions.

4

As is the case with many other languages, the Chinese language relies heavily on set phrases, fixed collocations, and idiomatic expressions in both oral and written contexts. Among the various prefabricated units in Chinese, collocations have drawn attention from many scholars (Qi, 2005; Li, 2008; Xin, 2014) because of their complex nature in relation to semantics, phonology and syntax.

In terms of the current pedagogical treatment of Chinese collocations in the language classroom, while most instruction in this area deals with more basic uses (such as the collocational use of the verbs "打, *to hit*" and "做, *to do*") and targets intermediate level learners, relatively little is known about difficulties experienced by highly proficient learners in mastering complex and highly idiomatic verb collocations. A better understanding of these difficulties will shape pedagogical approaches and instructional materials for the successful teaching and learning of this topic beyond the basic knowledge that already exists.

Another complex issue regarding Chinese language acquisition is the increasing presence of Chinese learners who have a family or cultural involvement in some varieties of the language. Such learners are generally regarded as Chinese heritage language learners (CHL learners). They include a wide range of learners with various language backgrounds, such as learners who are ethnically Chinese but with no or limited exposure to Chinese language and culture, learners who speak Chinese fluently but have little or no literacy skills, learners who speak a dialect other than Mandarin, etc. (Han, 2011). Due to such complexities, researchers have been unable to provide a full account of the various linguistic, social, and cultural competencies of this group of learners.

5

Motivated by the above mentioned gaps in the literature, the present study sets out to explore the following three issues: (1) differences among CFL learners, CHL learners, and CNSs in producing collocations in the spoken language; (2) influence of spoken contexts (academic topic vs. daily topic) on collocation usage; and (3) factors that could potentially affect the development of CFLs' and CHLs' collocational competency.

1.4. Organization of the current study

Chapter 1 introduces the general background and the organization of the current study.

Chapter 2 outlines the theoretical and methodological foundations regarding the study of collocations. First, drawing from various literatures, collocations are defined and the two major approaches in studying them, the frequency-based approach and the phraseological approach, are discussed. Then the theoretical models explaining the mechanisms of collocation acquisition are presented. Next, previous research on L2 learners' collocational knowledge and development is reviewed and discussed. And, finally, features of Chinese V-N collocations are summarized, and related SLA and cognitive linguistics studies are reviewed. Chapter 2 also provides a review of the common methodologies adopted in studying learners' collocations. Some advantages and limitations of current methodologies are also identified.

Chapter 3 highlights some of the fundamental issues associated with heritage language (HL) learners, with a particular emphasis on CHL learners. The chapter first discusses and defines HL learners in the U.S. context, based on previous studies. The chapter then presents a review of previous research on the comparison between HL

learners and foreign language (FL) learners in terms of their linguistic competencies. The second part of the chapter focuses on CHL learners and provides a review of definitions and issues regarding CHL learners.

Based on the literature review in Chapters 2 and 3, Chapter 4 presents research questions and the methodologies employed to answer them. This chapter describes the participants, the procedures and materials for collecting spoken data, and the analytical tests employed to evaluate the data. The results of the experiment will be presented in Chapter 5, which includes three sections that report the findings for the three research questions. For each question, the results of descriptive statistics and inferential statistics are reported first, followed by the results of the qualitative analysis.

Chapter 6 discusses the findings of the research as they relate to previous literature. The chapter first suggests some explanations for the findings about the differences between CFL learners, CHL learners, and CNSs in using Chinese V-N collocations, and then discusses the influence of the two types of topics on learners' production of V-N collocations. Finally, the chapter presents possible factors that contribute to the non-conventional use of collocations by advanced learners.

Chapter 7 presents the theoretical and pedagogical implications. The chapter discusses limitations of the current study and raises possibilities for future research. The chapter concludes with a summary of how the findings of this study can extend our knowledge of the use of collocations by advanced learners of Chinese with different language backgrounds.

CHAPTER 2: COLLOCATIONS IN LANGUAGE LEARNING

This chapter provides a review of the research regarding collocations in language learning with special attention to Chinese verb-noun (V-N) collocations. It begins with an overview of the key issues regarding the study of collocations in L2 learning, drawing on both the underlying theories and the various empirical studies. The discussion then moves on to examine the case of V-N collocations in Chinese, with reference to current theory and research in Chinese syntax and semantics. Finally, based on the review of literature, the definition and classification of Chinese V-N collocations in the current study are summarized.

2.1. Overview of collocation in language acquisition

Collocation is a general term used to describe the "co-occurrence of two or more words within a short space of each other in a text" (Sinclair, 1990, p.170). It is one type of commonly studied "formulaic sequences" that also include idioms (e.g., "raining cats and dogs"), figurative expressions (e.g., "to freeze to the spot"), pragmatic formulas (e.g., "have a nice day"), discourse markers (e.g., "let me see now"), and so on (Wray, 2002). Such formulaic sequences have been proven to be ubiquitous in a language and thus are considered central to the mastery of a language, whether in recognition or production.

Among the many studies on formulaic sequences, the research on collocation has received special attention, not only because it is a pervasive phenomenon across many languages, but also because it reveals the intricate relationship between lexicon, syntax, and semantics. Its significance in language learning and teaching has long been

recognized by various linguists, language researchers, and language educators. The following review focuses on several of the central issues around collocation that are pertinent to the current study:

1. How can collocations be defined and identified? Which theoretical models have been presented to account for the acquisition and processing of collocations in L1 and L2?

2. What are the empirical findings regarding L2 learners' use and development of collocations?

3. What are some common methodologies in studying L2 learners' collocational competence?

2.1.1. Approaches in defining and identifying collocations

The term "collocation" is used and defined in very different ways in the field of applied linguistics. While some researchers use the frequency of co-occurrence of words in a text as a criterion for defining collocations (Firth, 1957; Sinclair, 1991), others emphasize the syntactic and pragmatic relations between elements of a collocation (Nesselhauf, 2005). The following two sections review the two major approaches in defining collocations.

2.1.1.1. The frequency-based approach

The first approach adopted by many linguists is the frequency-based approach, which emphasizes the co-occurrence of two or more words in a certain span (Sinclair, 1991). Firth (1957) first used the term collocation as a technical term and defined it as

"actual words in habitual company" (p. 4). Although this definition is rather vague, Firth pointed to its important role in establishing meaning and functional values of words in real use.

Building on Firth's idea, Halliday (1961, 1966) further expanded the notion of collocation to include the syntagmatic association of lexical items and certain restrictions on the co-occurrence of lexical items. Halliday also established some fundamental terms in studying collocation. For example, "node" is introduced to mean the key lexical item under study and "collocate" refers to the co-occurring lexical items. For example, if we are interested in studying different V-N combinations with the verb "study", "study" would be the node and nouns (such as English, math, book, etc.) that occur frequently with "study" would be the collocates.

Following Firth's view of collocation and applying his ideas to many practical projects, Sinclair (1991) made important contributions to the operationalization and interpretation of collocations, defining collocation as "the occurrence of two or more words within a short space of each other in a text" (p. 170). He further expanded on the notion of co-occurrence within a short distance by specifying a span of 4:4, which means that most collocates can be found in the span of four words before and after the node (Jones & Sinclair, 1974). He also proposed two principles for interpreting collocations: the open-choice principle and the idiom principle. The former refers to the regular but complex choices made in creating utterances from single words. And the latter claims that these choices are further restricted by a large number of prefabricated phrases already available to the speakers. Sinclair stressed the importance of the idiom principle

10

in creating the meaning of words and thus identified collocation as an essential

component of the production of language.

The frequency-based approach is accepted by many corpus linguists (Biber and

Barbieri, 2007; Hunston, 2002; Partington, 1998; Shin & Nation, 2008; Stubbs, 2003),

whose analysis often employs object criteria such as frequency, range and collocational

span (the distance between constituents of the collocation). For example, Shin and Nation

(2008) carried out an analysis of spoken English using six strict criteria that concern

frequency, range, and collocational span and presented a list of the highest frequency

collocations of spoken English. Such studies generally do not consider the syntactic

relation between elements as an important factor in determining the formation of a

collocation.

2.1.1.2. The phraseological approach

The phraseological approach, unlike the frequency-based approach, which is

concerned mostly with frequencies and statistical significance, emphasizes grammatical

structure and the degree of semantic transparency as underlying principles for the

identification and analysis of collocations. Researchers following the phraseological

tradition (Cowie, 1981, 1994; Howarth, 1996, 1998; Nesselhauf, 2003, 2005) have

attempted to develop typologies and frameworks for classifying various collocations and

other combinations. In the seminal works of Cowie (1981, 1988, 1994, 1998), Howarth

(1996, 1998), and Nesselhauf (2005), distinctions of different types of word combinations

were largely based on two criteria: transparency and combinability (also called

substitutability). Transparency refers to whether the elements of the combination and the

11

combination itself have a literal or a non-literal meaning (Nesselhauf, 2005). And combinability refers to whether and to what degree the paradigmatic substitution of the elements of the combination are restricted (Cowie, 1981). Based on both criteria, Howarth (1996, 1998) proposed four categories of word combinations under the Continuum Model:

1. *Free combinations* (e.g., pay a bill)
   - each element may be substituted without affecting the meaning of the other
   - all lexical elements are used in a literal sense
2. *Restricted collocations* (e.g., pay a visit)
   - some substitution is possible, but there are arbitrary limitations
   - at least one element has a non-literal (figurative, technical, or delexical) meaning, and the other element is used in its literal sense
3. *Figurative idioms* (e.g., pay the price)
   - substitution of the elements is seldom possible
   - the combination has a figurative meaning, but preserves a current literal interpretation
4. *Pure idioms* (e.g., pay the piper)
   - combination is fixed, substitution of the elements is impossible
   - the combination has only a figurative meaning

In an attempt to further simplify the classification process by employing only one criterion instead of two, Nesselhauf (2003, p. 226-227) classified V-N combinations into the following three categories, using only one criterion "degree of restriction":

1. *Free combinations*: both the noun and the verb are used in an unrestricted sense, so they can be freely substituted by other verbs and nouns (e.g., want a car)
2. *Restricted collocations*: the verb is used in a restricted sense and can only be used with certain nouns (e.g., take a picture/photograph; but e.g., * take a film/movie)

12

3. *Idioms*: both the verb and the noun are used in a restricted sense, so substitution is mostly impossible (e.g., sweeten the pill)

From the above illustration of Howarth and Nesselhauf's classification of collocations, we can see that the phraseological approach identifies word combinations with clear semantic and syntactic relations between the constituents. Also, this approach prefers to view word combinations on a continuum, ranging from the most transparent and free combinations to the completely fixed and invariable idioms.

2.1.1.3. Summary of the two approaches

After presenting the two major conceptual views of collocation, this study summarizes some of the strengths and weaknesses of the two approaches as discussed in the literature.

The frequency-based approach employs objective criteria and can generate many statistically important findings, but it lacks a syntactic and semantic analysis of the relationship between the collocation and the language context in which the collocation is used and may lead to the identification of lexical chunks that have little psycholinguistic validity for language users (Henriksen, 2013).

The phraseological approach, on the other hand, provides a descriptive and typological framework for identifying word combinations and reveals the complex syntactic and semantic relationship between words in collocations. However, this approach remains largely speculative and lacks empirical validation.

2.1.2. Theoretical models on the acquisition and processing of collocations

By analyzing large amounts of natural language data, researchers have argued that instead of being peripheral to language meaning and grammatical rules, phrases or fixed strings of words are actually central in both first and second language acquisition processes (Meunier & Granger, 2008). A number of cognitive principles and models have also been proposed to explain the crucial role of collocations in both written and spoken language.

Usage-based models are a number of cognitive linguistic models emphasizing that actual language use has a significant influence on linguistic structures (Tyler, 2010). According to Ellis, "Usage-based theories hold that the acquisition of language is exemplar based. It is the piecemeal learning of many thousands of constructions and the frequency-biased abstraction of regularities within them" (Ellis, 2002, p. 143). In other words, when human beings use words to communicate, patterns of use emerge and language users are constantly categorizing, consolidating, and creating language structures to shape the future of the language.

In relation to collocations, Ellis (2002, 2003) argued that collocation is developed from a psychological mechanism known as "chunking", which was used to explain how short-term memory is tied to "chunks" of information for fast processing. Ellis believed that the same process can be applied to collocation acquisition. Two or more words that frequently co-occur are recorded as a chunk and are treated as a single entity. By constantly repeating the "chunking" process, language users can store greater amounts of

information in short-term memory, thus increasing the efficiency and fluency of communication.

Ellis (2003) also claimed that the chunking process is driven by "the Law of Contiguity." This rule states that, "Objects once experienced together tend to become associated in the imagination, so that when any one of them is thought of, the others are likely to be thought of also" (James, 1890, quoted in Ellis, 2003, p. 77). Based on this rule, the frequent co-occurrence of two words in linguistic input will lead to their becoming associated in long-term memory and perceived as "chunks." In first language acquisition, the knowledge of collocation is often acquired implicitly from extensive usage. In L2 learning, many formulaic languages are learned explicitly. Thus, Ellis suggested that the complex interface between implicit knowledge and explicit knowledge should be studied.

Wray (2002, 2008) also presented a comprehensive model to explain the use and process of collocations. She defined a "dual-processing system" in language speakers that includes two types of processing mechanisms: analytic and holistic. While holistic processing enables speakers to produce formulaic language patterns, analytic processing helps speakers formulate novel utterances. Wray also pointed out that there is relatively less effort involved in holistic processing. In terms of adult second language acquisition, her model suggests that, unlike children who learn a second language by focusing more on phrases, adult second language learners tend to separate fixed phrases and expressions into single lexical items and thus make more mistakes while trying to reconstruct the lexical items into formulaic language.

Wray suggested that this fundamental difference between child L1 learners and adult L2 learners is the result of a combination of social and cognitive factors. On the social side, adult learners (especially those in a classroom environment) rarely have the pressing need to memorize helpful communicative sequences for immediate communication. In addition, traditional classroom teaching methods often focus more on vocabulary learning and grammar drills that do not facilitate the learning of native-like formulas. On the cognitive side, mature adult learners who possess a full set of L1 vocabulary and rules will tend to break down formulaic L2 languages into individual words rather than treating them as holistic units.

Both the usage-based model and the dual-processing model originated from L1 acquisition and consider the significance of storing and processing frequently occurring chunks of language. In terms of the acquisition of collocations, the theories seem to suggest that collocations are formulaic for native speakers but not for language learners, and that word combinations that are familiar and frequent should be processed faster, at least in an L1. However, whether such an acquisition model can be fully applied to L2 learning of collocations remains questionable. A limited number of studies regarding the processing of collocation in L2 have presented mixed findings. For example, regarding the processing of idioms, some studies have found that idioms were processed faster by both NSs and NNSs than novel sequences were (Conklin & Schmitt, 2008), while others observed no processing advantages for NNSs reading texts with embedded idioms compared to matched control phrases (Siyanova-Chanturia, Conklin, & Schmitt, 2011). Therefore, to better understand the underlying mechanism of L2 acquisition of

16

collocations, researchers have called for more comprehensive and empirically based

models of word combination processing and representation in an L2 (Gyllstad & Wolter,

2015).

2.1.3. Main findings on L2 learners' collocational competence

Many studies have compared the collocational behavior and competence between

native speakers and non-native speakers using comparable datasets or corpus, and the

focus of such studies is usually to identify patterns of learners' use of collocations (Bahns

& Eldaw, 1993; Fan, 2009; Granger, 1998; Howard, 1996, 1998; Nesselhauf, 2003).

The overall picture that has emerged from such comparative studies is that

learners tend to either underuse, overuse, or misuse word combinations (Henriksen,

2013). A recent study by Tsai (2015) compares Taiwanese EFL learners' written

production of collocations with that of native speakers using part of the British National

Corpus (BNC). The results show that, compared with native speakers, EFL learners' use

of collocations is characterized by an inordinate number of collocations with little

variation in terms of types. Another study by Chen and Baker (2010) found that learners'

academic writing showed the smallest range of lexical phrases, as opposed to the widest

range exhibited by published journal papers. Such results are in line with earlier findings

(Granger, 1998; Kaszubski, 2000) that learners tend to frequently use a limited number of

"safe" collocations while underusing less common and more idiomatic collocations. In

terms of the implications for collocation learning, Durrant & Schmitt (2009) pointed out

that "learners are quick to pick up highly frequent collocations, but less common,

17

strongly associated items (e.g., densely populated, bated breath, preconceived notions) take longer to acquire" (p. 175).

In addition to such overuse and underuse of collocations, learners also tend to make self-created and sometimes deviant uses of collocations. A general approach to measure the collocational strength in learner language is to use native speaker corpora or dictionary as the basis, rank the frequency of combinations in the learner language, and then judge the acceptability of collocations produced by the learners. Nesselhauf (2005) adopted three measures (dictionaries, corpora, and native speakers) to evaluate the degree of acceptability of approximately 2,000 V-N collocations produced by German EFL learners and concluded that approximately one-third could be considered unacceptable or questionable. More specifically, she found that learners made the greatest number of deviant uses with restricted collocations, followed by free combinations and idioms. Similar conclusions have been drawn by researchers employing different approaches and studying different types of collocations.

Based on the findings of many studies (e.g., Nesselhauf, 2005; Serrano, Stengers, & Housen, 2015), collocations present special challenges to language learners, even at an advanced proficiency level in both written and oral tasks. For example, Serrano et al. (2015) investigate the difference between three levels of EFL learners and native speakers in producing formulaic sequences while performing an oral narrative and found that there is still a marked difference between the most advanced EFL learners and native speakers, in terms of both the number and range of formulaic sequences produced. While there is no clear link between language proficiency and collocation competence, it has

been observed that learners with high proficiency levels make greater use of collocations. However, more does not necessarily mean better, as advanced learners continue to produce many non-conventional uses of collocations.

The development patterns of ESL/EFL learners' collocational knowledge have also attracted the attention of researchers. Gitsaki's study (1996) contributed to the field by presenting the developmental process that L2 learners follow in the acquisition of collocations. She examined 275 Greek ESL learners (junior high school students) at three different proficiency levels (post-beginner, intermediate, and post-intermediate), using three measurements: essay writing, a translation test, and a cloze test. Thirty-seven collocation types, operationalized in the BBI Combinatory Dictionary of English, were adopted. The data yielded a number of interesting results with respect to the free production and cued production of collocations. In the essay writing test, Gitsaki reported that there was a significant difference in the production of collocations between and within the different proficiency groups, in relation to accuracy and the range of collocations used. The post-intermediate level was reported to be more accurate in the production of both grammatical and lexical collocations, as well as the use of various collocation types, than the other groups. Similarly, considerable differences were found across and within the three groups in the results of the translation and cloze tests, with the post-intermediate students being more accurate in their production of collocations. Also, among the 37 types, V-N (creation) lexical collocations (e.g., draw conclusions, face problems) were the most difficult for all subjects in all three tasks. Gitsaki explained that this is due to the arbitrariness and unpredictability of such collocations, which makes it

difficult for L2 learners to cope with them. Finally, Gitsaki concluded that as language

proficiency develops, collocational knowledge expands steadily. However, she argued

that the acquisition of collocations is affected by factors such as familiarity, frequency of

the input, and "salience" of the collocation types.

While the above findings have confirmed the difficulty associated with

collocation production, it should be noted that the findings from these studies were based

largely on "production," and few empirical studies directly account for the variety of

factors that may influence the acquisition process.

To further understand the nature of collocational competence, researchers have

explored major factors that could influence learners' productive and receptive knowledge

of collocations. Many experimental studies and corpus-based studies on leanrers'

collocation uses (Fan, 2009; Granger, 1998; Nesselhauf, 2005; Yamashita & Jiang, 2010)

tend to support the claim that learners' first language has the most direct and significant

influence on their processing and production of word combinations. The role of L1 in L2

collocation learning can be both inhibitive and facilitative, depending on the degree of

overlap between L1 and L2. For example, Yamashita and Jiang (2010) examined three

groups' (English NSs, lower proficiency Japanese EFL learners, and higher proficiency

Japanese EFL learners) acceptability judgment of two types of collocations: congruent

(collocations with an exact translation in the L1) and incongruent (collocations without

direct L1 translations). The results of their study led to two interesting claims. First,

learners' L1 does show a strong influence on the acquisition of collocations, as

incongruent collocations are more difficult to acquire than congruent collocations. And

second, with increased exposure to the L2, higher proficiency learners may be able to process L2 collocations independently from L1.

Besides the influence of L1 transfer, frequency effect is another widely studied factor. Many researchers argue that L2 learners do not have enough exposure to the varied use of collocations in different contexts and cannot build sufficiently strong associations between the different items within a collocation (Durrant & Schmitt, 2010). A recent study by Peters (2014) investigated the effects of repetition and other factors on EFL learners' form recall of single words and collocations. This study confirmed the strong and durable effect of repetition on learners' recall of collocations.

In addition to the two main factors above, a number of other factors--including the semantic properties of different types of collocations (Webb & Kagimoto, 2010), different levels of attention (Fan, 2009), different input conditions (Sonbul & Schmitt, 2013) and different program settings (Serrano et al., 2015)--have all been discussed as related to acquiring collocations.

2.1.4. Research methodologies in studying L2 collocational competence

Studies of L2 learners' collocational competence have adopted a wide range of research methodologies, from the analysis of learner corpora to online and offline experimental tasks. More recently, there has been a proliferation of collocation studies based on learners' productive knowledge of various written and oral tasks. And a number of trends have been identified for research along this theme.

First, regarding the data used for analyzing learners' collocation usage, many studies make use of large scale learner corpora such as the International Corpus of

21

Learner English (ICLE), which contains about 3.7 million words of EFL writing from learners representing 16 mother tongue backgrounds (Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Turkish and Tswana).

While most of the studies have focused on collocation usage in learner writing (Nesselhauf, 2005; Waibel, 2008), an increasing number also aim to investigate learners' collocation performance in oral contexts. As a result, research institutions have also developed several learner spoken language databases, including the Louvan International Database of Spoken English Interlanguage (LINDSEI), which contains about 800,000 words of oral data produced by learners from 11 mother-tongue backgrounds (Gilquin, De Cock & Granger, 2010).

While large scale corpora such as the ICLE could provide the benefits of enhancing the representativeness and generalizability of the data, other scholars have chosen to focus more on smaller datasets with better control of topic and lexical choices (Wang & Shaw, 2008; Javis, Grant, Bikowski, & Ferris, 2003).

Second, three main types of data-elicitation tools have been adopted for studying collocations: (1) written online tasks, often in the form of essays (especially argumentative essays) produced by both NSs and NNSs and often used by large corpora; (2) offline elicitation tools that include productive translation tasks, cloze format tasks, learner interviews, and receptive multiple-choice and judgment tasks; and (3) online reaction tasks mainly used in studying the processing of collocations. (Henriksen, 2013).

Third, in terms of the scope of this study, a large majority of the studies have

examined the use of collocation by English as a foreign language (EFL) learners, who are

often upper-intermediate to advanced learners. In addition, most studies contain data

collected at a single point in time (e.g., Chen & Baker, 2010; Durrant & Schmitt, 2009;

Gilquin, 2007), but a few longitudinal studies have been carried out recently (e.g.,

Crossley & Salsbury, 2011; Li & Schmitt, 2010)

Fourth, researchers have chosen to focus on different types of collocations. The

most widely studied collocations are V-N collocations, especially in English. Other

commonly studied types include adjective + noun and verb + preposition collocations.

Finally, another methodological issue examined in many studies is determining

the degree of acceptability of the collocations produced by language learners. As

discussed earlier, in evaluating learners' language production, the traditional approach of

using the native speaker's language as the only standard no longer holds its ground in the

context of language globalization. But for researchers working with collocations, it is still

important to capture the different features and problems of learners' production in a

systematic way. Ideally, every collocation made by the learners should have been

checked against huge corpora and/or judged by a large number of native speakers from a

variety of target language communities. However, it is often beyond the scope of a single

study to have every collocational use checked by both. Thus, researchers have adopted

more practical approaches.

Nesselhauf (2005) used three types of sources to determine the degree of

acceptability of the combinations extracted from a learner corpus: dictionaries, corpora,

and native speakers. By checking learners' usage against three different sources, she claimed that the study can make relatively reliable evaluations of the collocations produced by language learners.

This approach to judging the acceptability of combinations produced by learners is limited in that "there is not necessarily a one-to-one relation between what native speakers find acceptable or unacceptable when explicitly asked about a certain language phenomenon and what they themselves produced frequently" (Nesselhauf, 2005, p.53). Other researchers have also claimed that "there might be different standards of acceptability/appropriateness for foreigners and native speakers" (Johansson, 1979, p. 196). On the other hand, there are also indications that the correlation between corpus data and native speaker judgments on lexical combinations is fairly good (Hoffmann & Lehmann, 2000; Shei, 1999).

In summary, researchers have taken a great many approaches to fulfill different goals in studying collocations. Their different focuses and the heterogeneity in research instruments have made comparisons across the research area quite difficult and thus complicating attempts to make valid generalizations about L2 learners' collocational use and development.

2.1.5. Summary of research challenges

After presenting the many complex conceptual, theoretical, experimental and methodological issues surrounding the study of collocations, it is also necessary to reflect on some research challenges in the field.

First, a major challenge concerns the different definitions of the concept of "collocation" (Gyllstad, 2007; Granger, 2009). Various approaches will result in different focus areas and varying research findings. Therefore, it seems necessary to reach a more generally acceptable view of collocation.

Second, while many studies choose to focus on the frequency, type and accuracy of collocations, there is still a dearth of studies that probe the semantic association between the collocation constituents and the degree of restrictedness and opacity of the collocations.

Third, because of the variety of tools employed for eliciting collocations, it seems necessary to develop standardized receptive and productive instruments for measuring collocational knowledge (Gyllstad, 2007; Revier, 2009)

Fourth, in terms of the scope of collocation research, more studies should be carried out in languages other than English and in a variety of contexts including written and oral contexts, academic and nonacademic contexts, second language and foreign language contexts, and so on.

Finally, current studies on collocations are largely descriptive and short-term. Thus, researchers are calling for more longitudinal studies that follow the development of learners' collocational competence over time to gain insights into the developmental sequences of different types of collocations as well as the many variables that influence individual learner's collocation use (Henriksen, 2013).

Having considered the various issues related to the study of collocations, I will now turn to the focus of the current study, learning V-N collocations in Mandarin Chinese.

2.2. Chinese V-N collocations in learning Chinese as a foreign language

2.2.1. Characteristics and classification of Chinese V-N collocations

The Chinese language is rich in many types of prefabricated and idiomatic language chunks, including collocations, set phrases, and idioms. Due to the unique system of the Chinese written and spoken language, applied linguists have researched different aspects of it (vocabulary, syntax, semantics and pragmatics) for both L1 and L2 learning. Among the various aspects of Chinese linguistics, the relationship between verbs and their objects has received special attention. The numerous choices for V-N collocations and the complex semantic relationship in V-N collocations have posed learning challenges for learners of Chinese as a second or foreign language. This section reviews some basic classifications and characteristics of Chinese V-N collocations.

Chinese is often considered a largely parataxis, language meaning that its sentences are often connected by invisible logical relations instead of by regular grammatical rules (Tse, 2010). Chinese sentences are often sequenced through the collocation of many verbs or "run-on sentences" in which things are stated one by one in time sequence. As a result, many Chinese verbs can take a wider range of different objects. In English, most verbs can take only target objects as regular objects, as in "write a letter" or "write a sentence." However, Chinese verbs can take a number of different

types of objects. Take "写, *to write*" for example. In addition to regular objects such as letter or report, it can also take the following types of nouns as objects:

1. instrument: 写毛笔，写钢笔 (*to write with a brush/pen*)
2. manner: 写楷书，写草书 (*to write characters of regular script/cursive script*)
3. location: 写黑板，写纸条 (*to write on a blackboard/to write on a note*)

There are many different types of objects in the Chinese V-N constructions such as noun object, double object, verb object, adjective object, and small sentence object. Within the category of noun object, Meng (1989) further classified 14 types of noun objects on the basis of V-N semantic relation. The list below covers these objects and provides an example for each type.

Object type	Examples
1. Patient objects	安排活动 (*arrange activities*)
2. Result objects	包饺子(*make dumplings*)
3. Object objects	帮助同学 (*help classmates*)
4. Causation objects	关灯 (*turn off a light*)
5. Equate objects	他是老师 (*He is a teacher*)
6. Instrument objects	写毛笔 (*write with a brush*)
7. Manner objects	唱高音 (*sing at a high pitch*)
8. Location objects	参观博物馆 (*visit a museum*)
9. Time objects	过春节 (*spend the Spring Festival*)
10. Goal objects	考研究生 (*test for graduate study*)
11. Reason objects	避雨 (*take shelter from rain*)
12. Agent objects	来了一个人 (*come a man*)
13. Cognate objects	走路（*walk the road*）
14. Others objects	哭鼻子 (*cry the nose*)

27

Based on the unique features and complex classifications of Chinese verb-object constructions, researchers have identified a number of characteristics regarding Chinese V-N collocations.

First, while verbs can collocate with 14 different types of noun objects, the frequencies of these different types of objects in natural language are not the same. A number of studies have examined the overall trend of common Chinese verbs to take different types of noun objects (Kang & Dong, 1998; Li, 2003; Wei 2008). For example, Wei (2008) studied the 1,223 most commonly used verbs collected in the Dictionary of Common Chinese Verbs Usage (Meng, 1989) and found that there is great variability regarding the frequencies of different types of objects. Based on her analysis, patient objects co-occur most frequently with verbs (66% common verbs can collocate with patient objects). Other commonly used noun objects are location objects (44% common verbs), object objects (28% common verbs), and agent objects (20% common verbs).

Second, the collocation range of different Chinese verbs also varies greatly. While some verbs can take only one type of object (e.g., the verb "安排, *to arrange*" can take only patient objects), other verbs can take more types of objects (e.g., the verb "参观, *to visit*" can take two types of objects; the verb "打 *to hit, to play*" can take 11 types of objects). Factors influencing a verb's ability to collocate with objects include the semantic meanings of verbs, the transitivity of verbs, the number of syllabi of verbs, etc.

Third, the number of syllabi within a verb has a strong influence on the verb's ability to take noun objects. The Chinese language places a strong emphasis on syllabi and its relationship with phonology, semantics and grammar. Thus therefore the

following section will review the characteristics of monosyllabic verbs and disyllabic verbs in V-N collocations.

2.2.2. Monosyllabic verbs and disyllabic verbs in Chinese V-N collocations

A monosyllabic verb (Vms) refers to a verb with only one character (e.g., "打, *to hit*" and "做, *to do*"), and a disyllabic verb (Vds) refers to a verb with two characters (e.g., "告诉, *to tell*"and "知道, *to know*"). Chinese language historically has been considered to be a monosyllabic language, because the vast majority of Chinese morphemes are monosyllabic and most morphemes can be used as free words (Kalgren, 1949). However, with the development of the modern Chinese language, more disyllabic words have been adopted into the system and some scholars claim that Chinese has lost more than 50% of its monosyllabic words (Duanmu, 1999). Thus, modern Chinese is considered to be comprised of a large number of compound words that are mostly disyllabic. And through analysis of modern Chinese oral and written discourses, it has been observed that Vms are used more extensively in oral language and in daily communications. In addition, Vms are associated more often with idiomatic expressions. And this is likely due to the fact that modern Chinese is derived from ancient Chinese, which is predominantly a monosyllabic tongue (Duanmu, 1999).

Here are some examples of Vms in V-N collocations:

1. 关 guān (*to close; to shut off*)
    关门 (*to close doors*); 关空调 (*to turn off air conditioner*)
2. 过 guò (*to spend time, to cross, to live*)
    过生日 (*to celebrate birthday*); 过大桥 (*to cross a bridge*)

3. 开 kāi (*to open, to start*)

开玩笑 (*to play a joke, to make fun of*)

4. 谈 tán (*to talk about*)

谈恋爱 (*to be in love*)

From the above examples, it can be seen that Vms can be used to express many basic daily topics (examples 1 and 2) and can also be used in various idiomatic expressions (examples 3 and 4).

Vds, on the other hand, are rarely found in fixed idiomatic expressions but are observed frequently in both written texts and formal spoken contexts. Some examples of V-N combinations using Vds are as follows:

1. 发生 fā shēng (*to happen, to occur, to take place*)

发生问题 (*occur a problem*); 发生情况 (*occur a situation*)

2. 参加 cān jiā (*to participate, to take part in*)

参加会议(*attend a meeting*); 参加活动 (*take part in activities*)

In addition to the different uses of Vms and Vds in various language contexts (oral versus written; formal versus informal), Vms and Vds also differ in some other semantic connotations. Zhang (2004), by comparing Vms-N and Vds-N combinations, concluded that the Vms category tends to contain more strong action verbs than the Vds category. Wei (2008) found that the verbs that can collocate with the most types of noun objects are all Vms. For example, "跑, *to run*" can take 7 types of objects, "吹, *to blow*" can take 9 types of objects and "打, *to hit/to play*" can collocate with 11 types of objects.

In addition to such semantic differences, collocations with Vms and Vds are also affected by other linguistic factors. One important and unique factor is the rhythmic

constraint (prosody) in the formation and processing of V-N combinations (Zhou, Ye, Cheung, & Chen, 2009). In Chinese, the rhythmic pattern refers to the combination of words with different numbers of syllables. Yip and Rimmington (2004) used the term "rhythmic principle," which refers to the phenomenon that in V-N collocations, if the verb is monosyllabic, the noun object can be either monosyllabic or disyllabic; but if the verb is disyllabic, the noun object mostly can be only disyllabic. For example,

1.	to read newspaper	看报	看报纸
		阅读报纸	*阅读报
2.	to drive	开车	开汽车
		驾驶汽车	*驾驶车
3.	to perform a play	演戏	演话剧
		表演戏剧	*表演戏

Although all of the above V-N collocations seem to conform to grammatical and semantic rules, the collocations in which a Vds takes a monosyllabic noun object violate the rhythmic principle and are thus considered inappropriate uses.

One final feature worth noting about Chinese V-N collocations is that many monosyllabic verbs have Vds counterparts; and these monosyllabic and disyllabic synonyms often have similarities in both their forms and their meanings (Zhou et al., 2009). For example "学" and "学习" both mean *to study, to learn*; "找" and "寻找" both mean *to find, to look for*. When forming V-N collocations, these monosyllabic and disyllabic synonyms are interchangeable in some cases, but not in others. While native speakers can often switch back and forth between these forms effortlessly as the situation

requires, learners of Chinese often find it hard to distinguish between these two types of verbs.

2.2.3. Research on L1 and L2 use of Chinese V-N collocations

Following the cognitive linguistic analysis of collocations in languages such as English, a number of Chinese researchers have also tried to examine the nature and classification of collocations in the Chinese language, with particular attention to verb-object collocations.

Li (1983) raised the issue of object restriction in V-N collocations. He distinguished between restricted objects and free objectives and pointed out that there are different degrees of restrictions between a verb and its objects. Wei (2008) classified the objects of commonly used verbs based on two criteria: substitutability and expandability. First, substitutability refers to whether an object word can be replaced by other words from the same category. For example, the verb "来, *to come*", when followed by locality objects, can be substituted with many locality nouns such as "来学校, t*o come to school*"，and "来北京火车站, *to come to Beijing train station*". However, the same verb "来" when followed by patient objects, can be substituted only with a limited number of nouns. While "来包裹了, *a package arrives*" and "来信了, *a letter arrives*" are acceptable, "来书了, *a book arrives*"，and "来西瓜了, *a watermelon arrives*" are not acceptable. Second, expandability refers to whether the verb-object structure can be expanded to include more elements. Many idiomatic verb-object expressions cannot be expanded structurally to include more elements.

32

Other researchers have examined the syntactic and semantic features of Chinese verb-object collocations. Since the 1980s, Chinese linguists have introduced ideas from the valence theory developed by the French linguist Lucien Tesniere (1959) to study the properties of Chinese grammar. The word "valence," borrowed from chemistry, refers to the number of objects that can be controlled by a single verb predicate. With regard to Chinese V-N collocations, Shao (1995) applied the valence theory to study the Vds-N collocations and drew the conclusion that such collocations have a high degree of freedom and thus are often open to substitutions. For example, in the Vds-N collocation "提高收入, *to raise income*", both the verb "提高, *to raise*" and the noun "收入, *income*" can be substituted freely with many other verbs and nouns.

With regard to the L2 learners' usage of Chinese V-N collocations, due to the complex nature of V-N collocations in that language, it has been generally observed that L2 Chinese learners experience difficulties in applying V-N collocations in a systematic and fluent way (Qi, 2005; Xin, 2014). Despite such observations, relatively few studies have investigated the issue of V-N collocations in Chinese learners' written and oral production. One important attempt to address this gap was made by Xin (2014). Adopting a corpus linguistic approach, Xin compared the collocation uses between CNSs and CFL learners. By analyzing a learner corpus comprised of non-heritage CFL learners' written essays at different proficiency levels, Xin identified, classified, and evaluated different types of V-N combinations made by CFL learners, as compared to CNSs. The results showed that while more advanced learners tend to use more V-N collocations in their written essays, they also tend to produce more deviant usages in terms of the choice of

33

appropriate constituents of the collocations. Following the same line, further studies should be conducted to investigate the nature and cause of advanced learners' difficulty in applying V-N collocations.

2.3. Definition, classification and research methods of collocation in the current study

2.3.1. Definition and classification of collocation in the current study

In defining collocation, the current study takes into consideration both the frequency-based approach and the phraseological-based approach. More specifically, this study follows Henrisken's definition, which described collocations as "frequently recurring two-to-three word syntagmatic units which can include both lexical and grammatical words" (2013, p. 30). By adopting this combined approach, the current study hopes to emphasize both the importance of the statistical analysis of the frequency and token of collocations produced by learners and the significance of detailed semantic and lexical analysis of the collocations across learners' L1 and L2.

The classification of collocations in the current study is based on a broad view of V-N collocations in Chinese which include free combinations, restricted collocations and idioms. As noted above, because many Chinese verbs possess a number of different meanings and can be followed by a variety of nouns, the semantic meanings of V-N collocations may also have subtle variations in relationship to topic and context. Therefore, it would not be feasible to separate restricted collocations clearly from free combinations and idioms, especially in spoken language.

As noted above, a wide range of different types of noun objects follow Chinese verbs, some of which involve complex semantic and syntax interpretations. Considering that the primary goal of studying language learners' use of Chinese V-N collocations, the current study will follow the criterion set by Wei (2009) and Xin (2014) to focus primarily on collocations that consist of simple verb plus simple noun, and will exclude the examination of the following types of V-N collocations:

1. verbs with double objects
   e.g., 他告诉我一件事。(*He told me one thing.*)
2. Verb object combination in "被" (passive) construction
   e.g., 这本书被打开了。(*This book was opened.*)
3. verbs in existential sentences
   e.g., 墙上挂着一幅画。(*On the wall hangs a picture.*)

2.3.2. Research methods of Chinese V-N collocations in the current study

In order to examine collocation use across different speaker groups and language contexts, the current study is designed to compare three groups of Chinese speakers (CFL learners, CHL learners and CNSs) across two spoken topics (daily conversation topic versus academic topics).

Because the focus of the current study is collocation use in learners' spoken language, the main elicitation tools for data collocation include learner interview and picture narration tasks. Both of these elicitations tools have been commonly adopted by researchers in acquiring oral language data. For example, two of the largest spoken corpora of learner English--LINDSEI and LOCNEC--employ both learner interview and picture-based story-telling tasks as their main instruments for eliciting oral data from

35

learners. More specifically, De Cock (2007, p. 219) described the design of these learner interviews as follows:

> The informal interviews in LINDSEI and LOCNEC are of similar length (approximately 2,000 words of interviewee speech each) and follow the same set pattern: the main body of the interviews takes the form of an informal and open discussion mainly centered around topics such as university life, hobbies, foreign travel or plans for the future, although many different subjects were touched upon when the interviewees introduced them into the conversation. …..Each interview concludes with a short picture-based story-telling activity.

Based on the goal and research questions of the current study, a modified version of the learner interview was adopted as the main research instrument. In the current study, it comprises two parts: (1) an informal 30-minute interview on daily conversation topics (language learning background, hobbies, travel, etc.); and (2) a 30-minutes interview on academic topics using both open-ended questions (narration of academic research project) and cartoon picture narration tasks (two political cartoons on environmental protection and higher education reform).

By incorporating these conceptual and methodological approaches, the current study will add to the current literature on collocation use from the perspective of Chinese language learning.

CHAPTER 3: CHINESE HERITAGE LANGUAGE LEARNERS

This chapter provides a review of the literature regarding the definition and characteristics of heritage language learners (HLLs) compared to second or foreign language learners (L2 learners) in higher education, with particular attention to Chinese heritage learners (CHL learners).

3.1 Definition of heritage language learners

In the U.S., the term "heritage language learners" (HLLs) has been used in a variety of contexts such as immigrant languages, indigenous languages, and endangered languages. In a broader sense, the term is used to reflect the ethnic, historical, ideological, or sociopolitical connection and investment in a language (Fishman, 2001; Wiley, 2001; Hornberger & Wang, 2008). For instance, Hornberger and Wang's (2008) definition views heritage language learners as "individuals with familial or ancestral ties to a language other than English who exert their agency in determining if they are heritage language learners of that language" (p. 6).

In a narrower sense, HLLs can be defined based on linguistic competence and language affiliation. For example, Valdés (2001) defined a HLL as a "student who is raised in a home where a non-English language is spoken, who speaks or merely understands the HL, and who is to some degree bilingual in English and the HL" (p. 38).

Both of these perspectives in defining HLLs are valuable for the current study. On one hand, it has been widely acknowledged that the study of heritage language learning and teaching is closely related to a number of issues, including language proficiency,

cultural identity, community involvement, curriculum and assessment, and policy making. In a comprehensive model to capture the many aspects involved in the research, teaching, and language planning for HLLs, Hornberger and Wang (2008) proposed a framework to situate the biliterate development of HLLs on the continua of biliteracy (Hornberger, 2003; Hornberger & Wang, 2008). Using this model, HLLs could position themselves or could be positioned by the society in terms of the following four continua:

1. context (micro ↔ macro; oral ↔ literate; bi/multilingual ↔ monolingual)
2. content (minority ↔ majority; vernacular ↔ literary; contextualized ↔ decontextualized)
3. media (simultaneous exposure ↔ successive exposure; dissimilar structures ↔ similar structures; divergent scripts ↔ convergent scripts)
4. development (reception ↔ production; oral ↔ written; L1↔ L2)

These four dimensions of biliteracy, each encompassing three continua, constantly intersect with each other to create different learning and development possibilities in linguistically diverse settings. In terms of linguistic development, due to the complex development of biliteracy along the receptive-productive, oral-written, and L1-L2 language skills continua, HLLs may demonstrate different levels of language competencies related to phonology, grammatical rules, vocabulary and literacy skills. In addition, due to the tradition of academic schooling that emphasizes formal, literary and decontextualized contents, even highly fluent heritage speakers can be considered less proficient in functioning in an academic or professional setting due to a lack of mainstream discourses acquired through formal schooling.

Considering the vastly different backgrounds and beliefs of HLLs, a number of researchers have argued for an ecological and resource-oriented approach to examine the

identity and development of HLLs. Instead of viewing HLLs as non-native and problematic, they are focusing more on the expertise and resources brought by HLLs regarding language, culture, and ideology into the classroom and the society. It has also been widely acknowledged that developing HLLs' linguistic and cultural knowledge to advanced levels is "valuable not only for the learners themselves and their families and communities, but also for individuals' sense of personal identity and connectedness to their past and to their extended families, and for society more broadly" (Li & Duff, 2008, p. 14).

In addition to a sociocultural understanding of HLLs, for the purpose of designing effective and appropriate curricula, materials, and programs to assist HLLs in advancing their language proficiency in both L1 and L2, proficiency-based definitions of HLL are also important. These definitions set several criteria for identifying and categorizing HLLs. For example, Valdés (2001) sets three criteria for identifying HLLs: (1) the home language, (2) minimal proficiency in the HL, and (3) the dominant or societal language.

However, even such a proficiency-based definition of HLL covers a highly heterogeneous group of learners who vary greatly in terms of their linguistic repertoires, literacy skills, sociolinguistic and pragmatic competence, age of immigration, extent of formal education in HL, etc. Another question that arises with regard to proficiency is how to decide the desired attainment and necessary instruction for different HLLs.

In order to account for the diverse issues in HLLs, researchers have attempted to define or classify specific HLL groups. For example, Kagan and Dillon (2004) divided Russian HL students at the University of California, Los Angeles into three categories:

(1) students who graduated from high school in Russia, (2) students who attended junior high schools in Russia, and (3) students who were born in the U.S. to Russian-speaking families. These classifications would provide language programs with a basic idea for assigning classes and designing curricula for different types of HLLs.

In addition to efforts for defining and categorizing HLLs on the basis of linguistic background and proficiency, researchers and teachers should be cognizant of the diverse identities, motivations, expertise, and resources that both HLLs and L2 learners bring into the classroom.

Researchers in the field of language acquisition have also proposed different theories for understanding the fundamental processes involved in learning and speaking a HL.

3.2. Theoretical approaches to understand HL acquisition

Much of the earlier work on HLLs focused primarily on describing the characteristics of different groups of HLLs and understanding the pedagogical challenges of language-specific HLLs. However, scholars have long recognized the need for coherent theories regarding the acquisition and development of HL among different groups of HLLs. Researchers have sought to situate HLLs in relation to first and second language learners and to establish a theory of HL learning on the basis of L1 and L2 acquisition (Carreira, 2004; Valdés, 2005).

A number of different approaches have been adopted to study the acquisition process of heritage learners versus non-heritage learners. Among them, two major approaches have emerged regarding the differential language use by and knowledge of HL speakers.

Researchers following the first approach view HL acquisition as an incomplete or partial process of L1 acquisition. Montrul (2010) described HL acquisition as "incomplete L1 acquisition that takes place in a bilingual environment rather than a monolingual one" (p. 11).

Theories of L1 acquisition generally characterize the process of L1 acquisition with early exposure to the native language, abundant input and interaction in a naturalistic context, early development of key linguistic features (basic grammatical structures, phonology, some vocabulary), as well as an often successful and complete outcome attained after various types of formal schooling. In contrast, theories of L2 acquisition have emphasized later exposure to the language, a varying amount of input in

an instructed and/or naturalistic setting, incomplete structures and vocabulary development, and highly variable proficiency outcome.

Montrul (2010) argued that while HL acquisition has some features of early L1 acquisition (e.g., development of phonology and early structures), due to the often restricted input and output learning environment for HL, HLLs can experience many difficulties similar to L2 learners, such as developmental errors, negative transfer of L1 (majority language), and fossilization.

The incomplete acquisition approach emphasizes the variability of HL competence and attempts to evaluate the language proficiency of HLLs on a scale of different language attainment as compared to native speaker norm. Lynch (2008), by highlighting grammatical and lexical similarities between HLLs and L2 learners of Spanish, argued for the theoretical stance of conceptualizing students' language abilities on a continuum rather than positioning them along the traditional dichotomies (heritage vs. foreign, native vs. non-native, bilingual vs. monolingual).

A number of researchers have cautioned against the notion of "incomplete acquisition" and the use of native speakers or monolinguals as the sole criterion for evaluating HLLs' language competency. There are two main reasons for arguing against the "incomplete acquisition" interpretation. First, some researchers claim it is not justifiable to compare bilingual/multilingual speakers' language proficiency with monolinguals, as the former generally have multiple native languages and the term "native" actually covers a wide range of variation. As pointed out by Rothman and Treffers-Daller (2014, p. 97), "nativeness can and should apply to states of linguistic

42

knowledge that can be described as varying, even significantly, from monolingual baselines."

Second, the so-called "incomplete acquisition" of HL may well be viewed as the result of linguistic change rather than the partial attainment of HL acquisition. Data from recent studies on English-Spanish and German-Spanish bilinguals as well as on adolescent and adult monolinguals (Guijarro-Fuentes, Pires, & Nediger, 2015; Schmitz, 2015) have revealed that the variation in HS language is likely due to an intrinsic change in the HL rather than deficits of the HS language. For example, in a recent research study, Guijarro-Fuentes et al. (2015) examine the acquisition of the different semantic conditions related to Differential Object Marking (DOM) by English-Spanish heritage adolescents and monolingual adolescents. DOM refers to a difference in the form of overt case marking that depends on the properties of the direct object. The findings of their study not only revealed that heritage bilingual Spanish is similar to L1 monolingual Spanish but, more importantly, showed that adolescent monolingual speakers and adult monolingual speakers differed significantly in perceiving and using DOM. Based on the findings, the authors argue that heritage language grammar should not be considered as "incomplete", rather it is a delay and/or change in process and such change may "trigger intrinsic language change" (p. 246).

In summary, there are different views regarding the underlying mechanism of heritage language acquisition and whether "nativeness" should be used as the baseline for evaluating heritage language attainment. While it is undeniable that there is some difference between heritage speakers' competencies and the monolingual variety of the

43

homeland language, further research is needed regarding the differences between heritage speakers and early/late bilingual speakers. In the current study, although native speakers' spoken data are collected, they are not used as the sole criterion for evaluate both CFL learners' and CHL learners' competences. Rather, they are included as a reference to show the different patterns of language use among monolingual speakers and the different types of bilingual speakers.

3.3. Factors that characterize and influence heritage language acquisition

A number of factors have been identified that influence the acquisition process of heritage languages.

Traditionally the quantity and quality of input has been considered a major factor in bilingual language acquisition. In L1 acquisition, certain structures that are considered macroparameters of language (parameters that distinguish one language from another) tend to be acquired early in childhood and thus require less input than other structures. Therefore, for bilingual language learners who may not have a sufficient amount of input for their less dominant language, even though they could acquire some language structures in early childhood, it may take them longer to acquire certain structures which need extended periods of input (Tsimpli, 2004). Moreover, monolingual speakers of a language usually develop knowledge of formal registers through schooling and written language. Heritage learners often do not have the opportunity to go through extensive formal schooling in the HL and thus tend to have difficulty accessing a full range of acquire a particular vocabulary and structure (Pires and Rothman, 2009).

44

More recently, the role of input has been discussed more thoroughly as providing essential building blocks for both monolingual acquisition and bilingual acquisition. A number of studies have provided evidence that heritage acquisition is indeed quite similar to native acquisition, and different types of competences result from various input situations. For example, Rinke and Flores (2014) examine the use of clitic forms (Clitic is a morpheme that has syntactic characteristics of a word but is phonologically bound to another word) and strong pronouns by adult German-Portuguese heritage speakers and monolingual speakers of European Portuguese. The authors attribute different preferences in using clitic forms and strong pronouns to the fact that HS come into contact more often with oral forms of European Portuguese, which allows for some variation in the use of pronouns. Instead of treating the performance of HSs as "incomplete" or "deficit", the authors claim that the heritage grammar promotes linguistic changes which are inherent to the speech of monolingual speakers.

In addition to the important role of input, age and timing of acquisition comprise another siginificant factor. A number of recent studies have identified age as the most significant factor in affecting the linguistic features of HLLs (Montrul, 2008; Livert & Otheguy, 2010; Flores & Barbosa, 2014). For example, Livert and Otheguy (2010) investigated first generation speakers of Spanish in New York city and found that HSs who arrived in the U.S. between the ages of 4 and 14 showed a significantly higher rate of pronoun use than those who arrived as young adults or later, which showed that earlier arrival age would enable HSs to acquire more linguistic features from the dominant societal language.

Another key factor which is relevant in bilingual language acquisition is the influence of the dominant language (Flores, 2015). While transfer from a dominant language can cause HLLs to produce a divergent use of certain structures in HL, such a process is viewed by some scholars as the processing and representation of two grammars in the learner language rather than the deficient acquisition of HL (Flores, 2015; Sorace, 2011).

Other factors, including generation, and language learning context, have also been identified as affecting HL acquisition.

3.4. Linguistic similarities and differences between HL and Non-HL Learners

As discussed above, the increasing number of HLLs has challenged language researchers and educators to pay special attention to designing appropriate curricula for these learners. Concurrently, researchers have been interested in understanding the similarities and differences between HL learning and L2 learning for both pedagogical and theoretical purposes.

As Montrul pointed out, **"**without proper understanding of how similar or different these two types of learners are, it is difficult to tell at this point whether the exact same methods applied to L2 learners in the classroom should also be applied to heritage language learners" (2008, p. 500). Many studies have been carried out to compare HLLs and L2 learners in terms of specific areas of linguistic knowledge (phonology, morphology, semantics, vocabulary, etc.).

In terms of phonology, some studies have provided evidence that HLLs benefit from early exposure to the HL and tend to have an advantage over L2 learners in terms of

phonology (Chang et al., 2008; Saddah, 2011). For example, Au, Knightly, Jun, and Oh (2002) and Knightly, Jun, Oh, and Au (2003) compared the receptive and productive knowledge of different aspects of Spanish phonology and morphosyntax features. Their results showed that HLLs achieved a higher level of phonology and pronunciation than L2 learners did. However, for morphosynatatic features, HLLs and L2 learners performed similarly in terms of accuracy. But other researchers have also pointed out HLLs may also experience difficulty in pronouncing certain sounds in the HL (Godson, 2004).

Morphology is an area of particular difficulty in many languages for both HLLs and L2 learners. Many studies have yielded mixed results in comparing the two groups. While some researchers have found no significant advantages for HLLs in aspects of morphology (Au et al., 2002; Montrul & Ionin, 2013), others have identified certain aspects of morphology in which HLLs seem to have an advantage over L2 learners (Håkansson's, 1995).

For example, Montrul, Foote, and Perpiñán (2008) examined the knowledge of Spanish gender agreement among 69 HLLs, 72 L2 learners, and 22 native speakers of Spanish. HLLs were chosen based on the following three criteria: (1) born and schooled in the U.S., (2) had no schooling experience in their home country, and (3) became bilingual in Spanish and English before age five. All the participants completed one oral picture description task and two written tests. The results showed that HL learners performed more like native speakers than the L2 learners on the oral test, which required more implicit knowledge. However, they were less accurate on the two written tests, which required more explicit knowledge. Moreover, HLLs and L2 learners systematically

47

made different errors in their production. Such findings seem to suggest that HLLs' knowledge of gender in Spanish "might be stored, represented, and deployed differently" from L2 learners due to differences in early learning and formal education experiences (p. 541).

The acquisition of vocabulary is also key to language learning and is closely related to knowledge of semantics, syntax, and morphology. Studies comparing the lexical knowledge of L2 learners and heritage learners have taken a psycholinguistic and experimental approach (Hulsen, 2000; Montrul & Foote, 2014; Polinsky, 2008). Findings from these studies have generally indicated that "HLLs and L2 learners differ in their knowledge of vocabulary, which is highly dependent on frequency, the context of acquisition and language use" (Montrul, 2012, p. 18). For example, Montrul and Foote (2014) examine whether the selective retention of vocabulary is affected by age of acquisition and whether such retention is the same for HLLs and L2 learners. Study participants were 28 HLLs of Spanish and 28 L2 learners of Spanish at the intermediate to advanced proficiency levels. The study found that both HLLs and L2 learners had better command of nouns than verbs, and HLLs did not hold an advantage over L2 learners in terms of speed and accuracy of lexical access.

While most of the research comparing HLLs and L2 learners' linguistic knowledge employed experimental methods, Lynch (2008) conducted a qualitative study on five HLLs and four L2 learners of Spanish. The former group were either born in the U.S. or migrated before age two and had two to five years of formal study of Spanish. In other words, they were all typical lower-proficiency HL learners. On the other hand, the

latter group had more than five years of formal study of Spanish, meaning they were generally at a more advanced level of Spanish L2 learning. Data were collected from individual interviews in Spanish by the researcher. Selected grammatical features were analyzed quantitatively, such as noun-adjective gender agreement, aspectual and mood distinction, subject-verb word order, etc. The overall results showed more similarities than differences between the two groups. The two most advanced L2 learners consistently outperformed or performed as well as the most advanced HL learner.

In summary, both similarities and differences have been found between HLLs and L2 learners in different areas of linguistic knowledge. Other factors, such as timing and type of input, also seem to have an effect on learners' performance. For language educators and researchers, it is also important to bear in mind that there is no dichotomous relationship between HLLs and L2 learners (Lynch, 2008). On the contrary, with their diverse but also intercepted background and experiences, HLLs and L2 learners should all be placed along a continuum of different linguistic repertoires and skills. By doing so, language researchers will be inspired to explore connections between theories of SLA and the learning of HLs, and language educators will be encouraged to draw on different pedagogical elements from both the HL classroom and the L2 classroom to serve the various needs of learners.

3.5. HLLs' acquisition of collocations

Up to now, only a few studies have examined the use of collocations by HLLs. While no significant trend or patterns can be drawn from these studies, they certainly

contribute to our understanding of the characteristics and process of acquiring collocations by HLLs.

Kanno, Hasegawa, Ikeda, and Long (2008) investigated the relationship between type of language learning experience and different linguistic repertoires of advanced English-speaking learners of Japanese. The targeted linguistic profiles include knowledge and production of structural patterns and lexical items/collocations. Results of the study showed that, regardless of the type of language learning experience, HLLs' weakest area is vocabulary and collocations. In particular, HLLs use sophisticated lexical items and collocations (such as idioms) much less frequently than native speakers of Japanese do. The findings are actually in line with a number of previous studies claiming that even advanced language learners use less collocations, in terms of both frequency and range, than native speakers.

Kim (2009) carried out a descriptive study of Korean HLLs' productive and receptive knowledge of Korean "noun-verb" constructions. The participants comprised 15 heritage Korean HLLs in a high-intermediate undergraduate Korean class and the instruments were Korean to English and English to Korean translation tasks. By analyzing participants' use of Korean "noun-verb" collocations, the study revealed that HLLs had more difficulty producing collocations than comprehending collocations. In particular, "N-Vt" collocations tended to present the greatest challenges for Korean HLLs. In addition, the author also interviewed all the participants and finding they had little awareness of the importance of learning collocations thus called for a curriculum which raises HLLs' awareness of collocational knowledge.

50

Finally, Treffers-Daller, Daller, Furman, and Rothman (2015) compare the use of lexical collocations among three groups: heritage speakers of Turkish in Germany, Turkish returnees, and adult Turkish monolinguals. Previous studies have found that employing conventional combinations of words such as collocations in a language can be very challenging for language learners, even at the very advanced level. Thus in this study, the authors are particularly interested in studying whether heritage speakers who return to their native country after the critical period could still acquire the conventional use of collocations in their heritage language. Findings of the study do reveal that although Turkish heritage speakers in Germany use collocations differently from Turkish monolingual speakers, Turkish returnees started to pick up the conventional use of collocations with the verb "to do" in Turkish only after one year of returning to Thurkey. The authors argued that their findings can lend support to the view that HLLs have the ability to adapt to their native language variety, even after a short period of immersion. Thus their language ability should not be viewed as "incomplete", rather it should be viewed only as a unique language variety resulting from different kinds of input conditions. And unlike L2 learners, who usually have a hard time picking up conventional collocations, HLLs can adapt quickly to the standard variety of their native language by receiving enough input in the native country.

The aforementioned studies examined HLLs' use of collocations based on different languages and different research goals. Nevertheless, they contribute to the current literature by providing evidence that HLLs have both similarities and differences in using collocations with L2 learners.

Having discussed the theories regarding HLLs in general, the current study turns to a group of HLLs of particular interest: Chinese heritage language (CHL) learners.

3.6. Characteristics of Chinese heritage language learners

To understand the characteristics of CHL learners, it is necessary to first clarify the meaning of Chinese for HL learning. Essentially, there are seven major dialects in the Chinese language family: Beifang Hua (the northern dialect, or Mandarin, the native language of more than 70% of the Chinese population), Yue (Cantonese, 5% of the Chinese population); Kejia (Hakka, 3.7%), Min (including Taiwanese, 4.1%), Wu (including Shanghai dialect, 8.5%), Xiang (4.8%), and Gan (2.4%) (Wiley, 2008). These dialects are largely unintelligible to each other and some also have different written scripts.

In addition to the many regional dialects of Chinese, there is also a spoken norm which can be called Mandarin, Putonghua, or Guoyu (mostly in Taiwan) and has served since the 1920s as the *lingua franca* in China. The current study uses Mandarin Chinese to refer to this Chinese spoken norm.

Mandarin Chinese has been widely used and developed in mainland China, Taiwan, Singapore and elsewhere (e.g., in Chinese communities overseas). Although it is considered intelligible to people from all the above regions, there are still regional differences in pronunciation, vocabulary, syntax, and semantics.

In discussing the Chinese language, it is also important to distinguish between its spoken and written forms. Although most Chinese dialects share the same written script, there are varities in different regions. For example, mainland China has been promoting

the use of simplified characters since 1949, while Taiwan and Hong Kong maintain the use of traditional characters.

Thus, it is obvious that the term Chinese language is far from monolithic and can be used only as a very general term. In addition, due to the many historical, political, cultural, and social connotations associated with Chinese language, both native speakers and heritage speakers hold different views of the language.

According to the U.S. Census Bureau, in 2000, among the over 2.2 million people of Chinese ethnicity in the U.S., 83% spoke Chinese, while the rest were either English monolinguals or speakers of other languages. Moreover, the majority of the Chinese speakers were foreign-born, and ranked as the fourth-largest immigrant group in the U.S. Chinese (including both Mandarin and Cantonese) also became the second most common FL spoken by people in the U.S., following Spanish (U.S. Census Bureau, 2000). These people, or their ancestors, moved to the U.S. primarily through three waves of large-scale immigration from different parts of China to the U.S. during the past 160 years. As a result of the diverse backgrounds of these Chinese immigrants, they are considered to be a highly heterogeneous group of language speakers because of the different regions, dialects, and identities associated with the Chinese language.

In the U.S., CHL instruction and studies comprise a newly emerging field. Since the mid-1990s, it has attracted a rapidly growing body of scholars from various disciplines, such as SLA, bilingualism, reading research, discourse analysis, orthography analysis, and/or language pedagogy (He, 2008). CHL learners have presented unique characteristics, which differ from their non-CHL counterparts in terms of linguistic

knowledge system (Dai & Zhang, 2008; Hendryx, 2008), literacy skills (Jia, 2009; Xiao, 2008), morphological awareness (Koda et al., 2008), Chinese character learning (Ke, 1998; Xiao, 2006), and their motivation (Lu & Li, 2008; Weger-Guntharp, 2006).

An inherent problem in talking about CHL education in the U.S. is that the term Chinese is generally assumed to mean Mandarin Chinese and does not include the other dialects of the Chinese language family. As discussed below, the practice of using Mandarin Chinese as the "standard" heritage language can become problematic for many CHL learners.

The following sections will discuss questions related to defining and categorizing CHL learners, understanding their identities and motivations for HL learning, and identifying the linguistic skills and needs of these learners.

3.6.1. Defining and categorizing CHL learners

As noted above, the term CHL learners actually covers a very heterogeneous population. There are Mandarin speaking CHL learners and non-Mandarin speaking CHL learners who were exposed at home to different dialects of Chinese that may be mutually unintelligible to each other (e.g., Mandarin Chinese and Cantonese). Moreover, CHL learners possess very different language experiences and thus vary greatly in terms of their birthplace, family background, age of immigration, linguistic skills, literacy skills, etc.

Based on actual linguistic competence and familial affiliation, following Valdés (2001), a CHL learner is defined as one who "is raised in a home where Chinese is spoken and who speaks or at least understands the language and is to some degree

bilingual in Chinese and English" (He, 2006, p.1). However, CHL learners identified

under this definition still have a very "uneven grasp of the HL, falling along a continuum

of having very little HL knowledge to being highly proficient" (Li & Duff, 2008, p. 17).

As for HL learners of other languages, CHL learners' uneven proficiency results

from their different language experiences. The influence of their home dialect and the

orthography system play key roles in the development of their Chinese language and

literacy skills. A CHL learner whose home dialect is unintelligible to Mandarin may have

difficulty with aspects of the spoken language in Mandarin Chinese classes, but may not

have any trouble in writing if the classroom script is the same as his/ her home script (He,

2008). Other CHL learners may speak standard Mandarin Chinese with a high level of

oral proficiency but may have limited reading and writing skills or a limited range of

sociolinguistic and pragmatic competence to function in a formal or professional context.

Furthermore, immigrant students from China, Taiwan, and Hong Kong who have

received some formal schooling in Chinese in their native Chinese-speaking countries are

also joining the population of CHL learners in U.S. postsecondary schools (Li & Duff,

2008). Since these immigrant students are already partly literate in their native language,

questions have been raised regarding what could be considered "fully literate" in the HL

for such learners.

In order to cope with the increasingly complex profiles of CHL learners,

universities and colleges in the U.S. have sought ways to categorize CHL learners. While

most institutions still put all CHL learners into the same heritage language classes, others

have begun to explore the possibilities of enrolling students based on their dialectal

background and prior formal schooling. But just as Hornberger and Wang (2008) noted, there is no single profile of HLLs. Thus, regardless of institutional classifications, researchers and teachers should always be cognizant of the diversity of linguistic profiles and culture identities CHL learners bring to the classroom.

Due to the nature of a small-scale linguistic study, the current study's analysis of CHL learners will include only 10 learners who were born in the U.S. or moved from mainland China to the U.S. before age five. Although the focus group was to be CHL learners whose home language was Mandarin Chinese only, due to the diverse background of most heritage learners, three of the CHL learners in the study also had exposure to other Chinese dialects during their childhood. This is yet another piece of evidence that language educators should be made aware of the often highly complex language learning experience of HLLs.

3.6.2. CHL learners' motivation and identity in learning the HL

Since the 1970s, sociolinguistic research on the role of motivation in language learning has attempted to explore the various motives students have for studying a different language. Gardner's (1985) well-known sociolinguistic model has categorized the various motives of language learning into two major categories: the instrumental orientation and the integrative orientation. The former refers to an external and practical incentive for learning a language—e.g., to obtain employment or to move to another country. The latter is more of an internal drive—e.g., to understand a different culture or to become member of a community.

The existing research argues convincingly that learners' attitudes and motivations remain the strongest and most consistent predictors of second language success (Dornyei & Skehan, 2003). However, rather than being a static and direct criteria, motivation is examined more as a dynamic and multifaceted concept that is closely related to the constantly changing learning context, socioeconomic status, and identities of learners.

In the case of the Chinese HL learners, several studies on CHL learners' motivation have yielded mixed results about learners' motives for learning Chinese. While some researchers reported that CHL learners show strong integrative orientation (Wen, 1997), others found that CHL learners are motivated more strongly by instrumental orientation (Lu & Li, 2008). For example, Lu and Li (2008) compared the different motivational factors among 59 CHL learners and 61 non-CHL learners from 9 Chinese college classes through questionnaires and interviews. Their results showed that even though both types of motivation played important roles for both learner groups, CHL learners were influenced significantly more by instrumental motivation than their non-CHL counterparts were. In addition, their integrative motivation was more highly correlated to their perceptions of their listening and speaking abilities, but not their reading or writing skills. For the non-CHL learners, the correlation was found between their integrative motivation and listening, speaking, and writing skills, but not their reading skills.

Other researchers have also argued for a more integrative view of motivation in understanding CHL learners' performance. Li (2005), in surveying a large sample of university CHL learners, found that there learners were strongly motivated by both the

57

instrumental orientation and integrative orientation. With regard to instrumental

orientation, CHL learners were more interested in a long-term investment in their future

careera than in simply obtaining "easy credits" for academic success. In other words,

Chinese learners are still very interested in knowing more about their culture roots and

identities, but they also want to be part of China's rapid economic growth.

In terms of identity, it has been widely acknowledged that the learners' identity,

which involves their perceptions and evaluations of themselves in relation to the outer

world, is crucial to their development and learning (Hornberger & Wang, 2008).

Similar to other HLLs, CHL learners carry with them multiple identities, which

are constantly negotiated and shaped by themselves and others. Their self-identification is

closely related to their language environment and social context, such as their place of

birth, length of residence in the U.S., age of immigration, and family socio-economic,

educational, and political backgrounds (He, 2004; Tse, 1998). Thus, CHL learners'

identities vary from individual to individual.

Dai and Zhang (2008) surveyed 80 college students with CHL backgrounds about

their culture identity. The majority of the participants viewed themselves as a

combination of both Chinese and American cultures, and they drew on the diversity of

their cultural knowledge when interacting with different people at different times and in

different social contexts. The remaining participants considered themselves either "less

Chinese than the Chinese, and less American than the Americans" (p.44), and thus

indicated that they belonged to neither the Chinese or American culture, or accepted one

culture but rejected the other.

Jia (2008) found that HL speakers who identified a stronger connection to and preference for their HL culture tended to not only use more HL, but also self-rated their reading and writing skills higher. In a related study, He (2008) considers learners" identity as a prime dynamic force rather than just the background in CHL learning. She argued that through CHL learning, which takes place in a three-dimensional framework with intersecting planes of time, space, and identity, a CHL learner will form a new cultural identity which inherits "some of the Chineseness" from his family and his neighborhood but will enable him to become a very different kind of Chinese-American from his family and his neighbors" (p. 110). In addition, she noted that, "the degree to which a learner's CHL develops is dependent upon the degree to which s/he is able to find continuity and coherence in multiple communicative and social worlds in time and space and to develop hybrid, situated identities, and stances" (p. 116). In her identity-based model for CHL development, she hypothesized that the degree of success in CHL development correlates positively with learners' desires to be connected with their heritage culture and CHL community members in the long term.

3.6.3. CHL learners' linguistic repertories and needs

CHL learners bring a variety of linguistic repertoires and skills to their language classroom. Due to the wealth of their heterogeneous past exposure and learning experiences in Chinese, many of them are able to speak a number of different Chinese dialects from different regions; some are able to write different scripts (e.g., the traditional script and the simplified script); most have a good command of a wide range of vocabulary and syntactic structures; and many are familiar with implicit and explicit

cultural norms associated with their home language. As a result, these learners also tend

to have diverse characteristics and needs for learning Chinese as compared to other non-

heritage learners.

First, while some CHL learners have certain advantage in terms of pronunciation

and phonology, others may not necessarily possess "standard pronunciation". According

to the existing literature on HL learning, many HLLs tend to have "native pronunciation

and fluency" (Campbell, 2000). Some studies on CHL learners also show that they tend

to outperform non-heritage Chinese learners in terms of their speaking and listening skills

(Jia, 2008). Despite these results, the spoken skills of CHL learners actually present a

very complex issue. As noted above, although Mandarin Chinese is often taught as the

"standard" heritage language of students with Chinese backgrounds, many of these

students actually speak a different dialect (e.g., Taiwanese, Cantonese) that is (almost)

incomprehensible to speakers of Mandarin Chinese. Thus, students who speak non-

Chinese dialects do not demonstrate phonological advantage in learning the language,

and many actually find themselves marginalized in the CHL classroom (Wu, 2013).

Second, CHL learners have different skill levels of reading and writing based on

their backgrounds and learning experiences. A number of studies have found that CHL

learners' Chinese acquisition places a strong emphasis on speaking and listening rather

than reading and writing skills (Dai & Zhang, 2008). There are two reasons for this. First,

CHL learners generally began learning Chinese at home, relying mainly on spoken

interactions with their parents or grandparents. Thus, many CHL learners do not formally

learn to read and write in Chinese until they take Chinese courses in secondary schools or

in college. And for some CHL learners, their home written script might be different than the standard script taught in the HL classes. Second, unlike some other HLs where students' speaking and listening competencies could facilitate their reading and writing skills, Chinese has a very different and sophisticated logographic writing system that differs considertably from its phonological system. As a result, it has been found that CHL learners' home background in Chinese has little or no effect on their learning to write Chinese characters (Ke, 1998; Xiao, 2006). On the contrary, in some studies, L2 learners of Chinese tend to perform better in reading comprehension and character writing than CHL learners do (Xiao, 2006).

Finally, although some CHL learners tend to have a good command of basic vocabulary and grammar, they still need to acquire higher-level registers and more sophisticated repertoires. As the existing literature has pointed out, developing HLLs' linguistic and cultural competencies to advanced levels is valuable not only for the learners themselves, but also for more closely connected families, communities and society in general (Fishman, 1991; Hornberger, 2003). A number of linguistic studies have identified CHL learners' strengths and weaknesses in terms of vocabulary and grammar (Wu, 2008; Hendryx, 2008; Koda et al., 2008); and most of these studies have focused on basic Chinese lexical and grammatical features, such as the use of aspect market "了"，the use of "把" sentence construction, etc. Although the results of these studies showed that CHL learners in general have a good command of such basic lexical and grammatical features, since many CHL learners are aiming to learn the language for future careers and professions, it also seems necessary that they acquire more

sophisticated, academic, and professional varieties of the language (Li & Duff, 2008). Currently, very few studies have focused on the advanced level registers, genres, and language varieties that learners need to acquire to reach academic and professional Chinese competence.

3.6.4. Summary of the discussion regarding CHL learners

In summary, CHL learners exhibit wide variability in terms of their motivations, identities, linguistic skills, and repertoires due to their diverse learning contexts and backgrounds. Language educators must acknowledge the individual differences and needs of this group of learners and to motivate them to fulfill their goals for learning the language. On one hand, it is essential to maintain and preserve these learners' often highly developed competencies in phonology, vocabulary, and grammar, especially in the domain of the spoken language. On the other hand, it is also an appropriate goal for CHL learners to acquire sophisticated, academic, and other registers that will enable them to function successfully in a wide variety of contexts.

CHAPTER 4: RESEARCH QUESTIONS AND METHODOLOGIES

Chapter 2 and chapter 3 provide an overview of the major issues surrounding the important role of collocation in language learning as well as recent research on heritage language learning. Drawing on both theoretical and empirical evidence from SLA, corpus linguistics and HL education, the current study aims to explore the influence of learner background and oral topic on the use of collocations. Chapter 4 first introduces the three research questions and then discusses in details the methodologies adopted in this study.

4.1. Research Questions

The current study seeks to explore collocation use among learners with different language profiles. As one of the first attempts to examine the use of Chinese collocations in oral context, I am particularly interested in comparing Chinese heritage language learners (CHL learners) and Chinese foreign language learners' (CFL learners) production of collocation and studying the effect of daily conversation topic and academic abstract topic on collocation use. More specifically, this study aims to examine the following three research questions:

1. What is the difference (type and token) in using Chinese V-N collocations among CFL learners, CHL learners and CNSs in an oral interview context?

2. How do two types of topics (daily conversation topics vs. abstract academic topics) affect CFL learners, CHL learners and NSs' use of Chinese V-N collocations?

3. What are the types and characteristics of non-conventional collocations produced by CFL learners and CHL learners?

4.2. Research Methodology

In order to study collocation use in an oral context, the present study employs learner interview as the main elicitation instrument and carries out both statistical analysis (e.g., types and tokens of collocations) and qualitative analysis (e.g., lexical and semantic characteristics of collocations) to analyze the spoken data.

4.2.1. Setting of the study

The present study is conducted at a private four-year research university in northeast United States. I have been an instructor in an advanced Chinese language and culture program at this university. The program is part of a two-year Master Program in International Studies. All students enrolled in the program need to complete two years (four semesters) of advanced language course and need to reach "superior" level language proficiency as measured by the Oral Proficiency Interview (OPI) test offered by the American Council on the Teaching of Foreign Languages (ACTFL) at time of graduation. Prior to entering the program, all students were tested by OPI once and they need to achieve at least "advanced low" to be able to enroll in the program. Once enrolled, students in the Chinese program usually start by spending 8 weeks during the summer in China taking one intensive Chinese course in an immersion environment. For the following 4 semesters in the United States, they are required to take one advanced Chinese language course every semester. The course meets twice a week for 80 minutes and covers a wide range of topics related to the current social, economic and political issues in the greater China region. The language of instruction is Mandarin Chinese and the main goal of the course is to enhance students' spoken proficiency in Mandarin

Chinese as well as to deepen students' understanding of issues related to the greater China region.

4.2.2. Participants

The participants of the study are composed of two experimental groups and one control group.

Experimental group one consists of 10 Chinese as a foreign language (CFL) learners who are all students of the above mentioned 2 year master program in International Studies. Table 1 summarizes the basic demographic and language learning background information of these 10 CFL learners as obtained from a brief language background questionnaire (Appendix A) filled out by learners at the beginning of their individual interview.

In terms of CFL learners' oral language proficiency in Chinese, 8 have achieved "advanced high" and 2 have achieved "superior" in their most recent OPI test. Within this group, 7 are males and 3 are females and their age range is between 24 and 30. All of them have English as their native tongue. The CFL learners' interest in learning Chinese generally started in their secondary school or college years. Besides taking formal Chinese courses at the secondary and university level, most of them have attended short term or long term exchange programs or intensive language programs in China. Moreover, 8 of them also gained work experience in China ranging from 1 to 4 years. They worked in a variety of different fields in China including education, consulting, finance, business and government.

65

Table 1. CFL Learners' information and language learning background

Name	Gender	Age	Current program of study	Native language	OPI rating	Years of formal Chinese learning	Years working in China
Bill	**M**	**24**	**1st year Master student**	**English**	**Advanced High**	**6.5 (1.5 years in China)**	**1.5 years**
David	**M**	**28**	**2nd year Master student**	**English**	**Advanced High**	**4 years (1 year in China)**	**N/A**
Julia	**F**	**26**	**2nd year Master student**	**English**	**Superior**	**5 (1 summer in China)**	**1 year**
Kathy	**F**	**26**	**2nd years Master student**	**English**	**Advanced High**	**4 (1 semester in China)**	**2 years**
Lucas	**M**	**25**	**1st year Master student**	**English**	**Advanced High**	**3 years (1 summer in China)**	**4 years**
Mary	**F**	**27**	**2nd A Master student**	**English**	**Advanced High**	**6 (2 summers in China)**	**N/A**
Ryan	**M**	**25**	**2nd year Master student**	**English**	**Advanced High**	**6 (2 years in China)**	**2 years**
Tom	**M**	**24**	**1st year Master student**	**English**	**Advanced High**	**5 (1 year in China)**	**2.5 years**
Tyler	**M**	**30**	**2nd year Master student**	**English**	**Advanced High**	**4.5 years (1 month in China)**	**1 year**
William	**M**	**26**	**1st year Master student**	**English**	**Superior**	**5 (2 months in China)**	**3 years**

The second experimental group contains 10 heritage learners of Chinese (CHL). Table 2 below reports CHL learners' basic information as obtained from the language background questionnaire. Among these CHL learners, 7 are enrolled in the above mentioned 2 year master program, and 3 are university seniors who attend an advanced business Chinese course. For their most recent OPI taken within the past year, 3 received

"superior" rating and 7 received "advanced high" rating. In terms of age and gender, 6 are females and 4 are males and their age range is between 21 and 29. Regarding their heritage background, 6 were born in the US to parents with Chinese background and 4 moved from mainland China to US before the age of 5. Therefore, all these learners grew up in households where Mandarin Chinese was spoken by at least one of their parents.

In terms of language learning background, although most CHL learners have taken formal Chinese classes either in high school or in college, 2 of them did not have any formal Chinese learning before they entered the two-year Master Program in International Studies. Also, on the language background questionnaire, 4 CHL learners indicated their experience of attending Chinese Sunday school weekly when they were in elementary school. Finally, unlike CFL learners who mostly had rich working experience in China, only 2 CHL learners had short-term (4 months to 6 months) experience working in mainland China. However, almost all have had chance to visit family or take short trips to China.

Table 2 : CHL learners' information and language learning background

Name	Gender	Current program of study	Age (Age moving to US	OPI rating	First Language at home	Years of formal Chinese learning	Years working in China
Amber	F	college senior	21	Advanced High	Mandarin Chinese	3 years (college Chinese class)	No
Claire	F	2nd year Master student	29 (2.5(	Advanced High	Mandarin Chinese/Cantonese/English	2 years (no formal training before graduate school, 1 summer in China)	6 months
Ethan	M	1st year Master student	27 (3)	Superior	Mandarin Chinese	1 year (college, 1 summer in China)	No
Grace	F	2nd year Master student	27	Superior	Mandarin Chinese/English	2 years (college Chinese class, 1 summer in China)	No
Kevin	M	2nd year Master student	26 (4)	Advanced High	Mandarin Chinese/English	2 years (no formal training before graduate school, 1 summer in China)	4 months (2 summers)
Lily	F	2nd year Master student	25	Superior	Mandarin Chinese	4 (1 year in college, 1 year in Taiwan, , 1 summer in China)	No
Mark	M	2nd year Master student	25 (3.5)	Advanced High	Mandarin Chinese/English	3 years (1 year high school, 2 years in college, 1 summer in China)	No
Mike	M	college senior	21	Advanced High	Mandarin Chinese/English	4 years (2 years elementary level Sunday school, 2 years in college)	No
Sarah	F	college senior	22	Advanced High	Mandarin Chinese	2 years (2 years in college)	No
Sue	F	1st year Master student	26	Advanced High	Mandarin Chinese	3 years (2 years college Chinese class, 1 summer in China)	No

The control group consists of 10 native speakers of Chinese who are also graduate students major in education and East Asian Studies at the same private university. Among the 10 native speakers, 6 are females and 4 are males and their age range is between 23 to 28 years. All of the native speakers grew up in major cities of mainland China and they all speak standard Mandarin as their native language. Table 3 summarizes the gender, age and birth place of participants in the control group.

Table 3. CNSs' background and interview recording data

Name	Gender	Age	Birth place	Native language
Bai	**F**	**23**	**Kunming, China**	**Mandarin Chinese**
Gang	**M**	**27**	**Shanghai, China**	**Mandarin Chinese**
Hong	**F**	**23**	**Beijing, China**	**Mandarin Chinese**
Hua	**F**	**23**	**Beijing, China**	**Mandarin Chinese**
Huang	**F**	**24**	**Hangzhou, China**	**Mandarin Chinese**
Jing	**F**	**26**	**Shijiazhuang, China**	**Mandarin Chinese**
Lan	**F**	**23**	**Nanjing, China**	**Mandarin Chinese**
Li	**M**	**25**	**Hangzhou, China**	**Mandarin Chinese**
Yong	**M**	**25**	**Guangzhou, China**	**Mandarin Chinese**
Zhu	**M**	**28**	**Nanning, China**	**Mandarin Chinese**

4.2.3. Targeted Chinese V-N collocations

As discussed in section 2.3.1, the current study adopts a broad view of V-N collocations to include free V-N combinations, restricted V-N collocations and idioms. Also, as reviewed in Chapter 2, there are many different forms of Chinese verb-object collocations as related to the complex word order, semantics and pragmatics of the Chinese language. Considering the aim and scope of the current study, the collocations to be focused on will exclude the following types of V-N collocations:

1. verbs with double objects

   e.g. 他 告诉 我一件事。(*He told me one thing*)

2. Verb object combination in "被" (passive) construction

   e.g. 这本书被打开了。(*This book was opened.*)

3. verbs in existential sentences

   e.g. 墙上挂着一幅画。(*On the wall hangs a picture.*)

4.2.4. Instrument and data collocation procedure

The main instrument of the current study is one-on-one interview with all participants. The following section provides a description of the format and procedure of the interviews and other data collection methods.

4.2.4.1. Instruments

In the present study, spoken data were elicited through one-on-one interviews with all 30 participants. Each 1-hour long interview consists of two parts: 30 minutes of conversation on daily topics and 30 minutes of discussion on academic topics. The description of these two parts of the interview is presented below.

One part of the interview is a casual conversation of common non-academic daily topics between me and the participant. Each participant was asked a number of questions regarding their personal background, language learning experience, daily activities and hobbies in a casual and spontaneous way. The list of questions used in this part of the interview can be found in Appendix B. Although the same list of questions was used as guiding topics for each interview, this part of the conversation is still very open and includes a variety of different topics which were brought up by the participants during the interview.

70

The other part of the interview is structured around the discussion and narration of abstract academic topics. Two types of prompts (open-ended question prompt and cartoon interpretation prompt) are employed in this part to elicit participants' spoken response to abstract topics. First, because all participants in this study are enrolled in graduate-level coursework and are trained to perform academic research in their own fields, each participant was asked to introduce in detail an academic research project they were involved in or were planning to carry out.

After being asked about their own research projects, participants were further asked to comment on two political cartoons that the author chose for this study from major news agencies. The two political cartoons concern two widely recognized social and cultural issues: China's environmental pollution and China's educational system reform (Appendix C). Each participant was given 2 minutes to read each cartoon carefully and then was simply prompted to describe the cartoon and offer their analysis and opinions on the issue reflected by the cartoon. The participant then gave a narration largely in the form of a monologue. In some cases, I asked simple follow-up questions to elicit more responses.

4.2.4.2. Interview Procedure

Each participant had one-on-one interview with me in a common office setting. Because each interview consists of two parts, the study employs a counterbalanced design to minimize the order effect of sequencing the two types of topic. For each group of 10 participants, I would begin 5 interviews with the topic of daily conversation and begin the other 5 interviews with the topic of academic discussion.

The interview process with the three groups is basically the same, except that both CFL learners and CHL learners were asked to fill in a simple language background questionnaire at the beginning of interview whereas native speaker did not fill in the questionnaire. The questionnaire was used to collect general information about learners' language learning including native language, years of formal Chinese learning, study-abroad experience and other related information. The language questionnaire used in this study could be found in Appendix A.

For each interview, I would start by introducing the topic and purpose of the current study and asking the participant to sign the consent form. I would then move on to one of the two types of topics.

For the daily conversation topics, the interview is carried out in casual question-and-answer format and I tried the best to create relaxed environment for the participants. For example, the following questions were asked during the interview:

- "我很想听听你是怎么开始对中文/英文有兴趣的。（*I really want to hear about how you first became interested in learning Chinese/English.*）"

- "你好像很喜欢旅游，可以聊聊你最近一次旅行的经历吗？（*It looks like you like travelling a lot, could we talk about one of your recent travel experiences*?"

While interviewing on more abstract academic topics, I tried to elicit more formal and extended spoken language. For one elicitation prompt, I raised the following open-ended question:

- "请向我具体地描述一下你读研究生时做过的一个研究项目，你认为这个研究项目有什么重要性呢？(*Can you describe to me in details about a research project you have done for your graduate work and discuss why you think it is important.)*"

And for the cartoon picture prompt, the spoken instruction was given as follows:

- "请你仔细看一下这幅漫画，然后谈谈漫画的内容和意义。(*Please read this cartoon carefully and comment on the content and meaning of it.*"

In instances when a participant only had very little to say about the political cartoon, simple follow-up questions would be asked to extend the discussion:

- "很好的想法，还有别的方面吗？(*Good thoughts, any other aspects*?)"
- "你觉得中国的教育体系需要怎样的变化呢？(*What kind of change do you think is needed for China's education system*?"

Since the current study is about collocational use, I also tried to minimize the use of V-N collocations in my own questions and prompts. This was done through careful planning of questions and prompts in advance.

Both parts of the interview were recorded using two audio-recording devices. For each part of the interview, I tried to control the length to be as close to 30 minutes as possible, so the total length of each interview is about 1 hour.

After all interviews were carried out, I transcribed all interviews and organized all transcriptions into three datasets: CFL learners, CHL learners and CNSs. All instances of V-N collocations were then manually identified and classified in the three datasets.

It is also important to note the limitations regarding the multiple identities of me, the researcher, as both the language instructor of the participants and the investigator of the research study. On the one hand, the familiarity between myself and most of the learners in the study could to a certain extent facilitate the interview process. Learners may feel more relaxed in answering questions and thus can produce longer discourses. On the other hand, such familiarity may also present a problem for examining the influence of different topics. This is because most one-on-one interactions between learners and the language instructor tend to be casual and conversational, thus learners might not feel the need to use formal and academic language even in discussing abstract topics. Also, instead of being a natural conversation, learners may perceive the interview more as a language practice with the instructor and this can also influence their overall language fluency and accuracy.

4.2.4.3. Examine the degree of acceptability and communicativeness of V-N collocations produced by learners

After all instances of V-N collocations were identified in the data, I then try to decide whether the collocations produced by learners are commonly used in the target language and whether the collocations are communicative in expressing meanings.

The first step is to determine the degree of conventionality. Since the current study involves collocation usage, it would be ideal to check every combination against huge spoken corpora and/or judged by large numbers of native speakers. However, there currently is no extensive Chinese spoken corpus and it is also beyond the scope of this study to have every combination judged by large numbers of native speakers. Instead, a more practical approach is adopted following the methods of Nesselhauf (2004) and Xin (2014).

In the present study, three types of sources are adopted to judge the overall degree of conventionality of each collocation produced by learners in the data: dictionaries, written corpora and experts' judgements.

First, collocations were judged conventional if they occurred in identical form and with the same (or similar) meaning in one of the following three widely used Chinese dictionaries: 《现代汉语词典》(*Contemporary Chinese Dictionary* published in mainland China), 《国语辞典》(*Chinese Dictionary* published in Taiwan), and《现代汉语搭配词典》(*Dictionary of Modern Chinese collocations published in mainland China*).

Second, collocations were also judged conventional if they occurred in similar

form and meaning for at least three times in the Peking University Contemporary Chinese

Language Corpus (CCL). Up to now, there is no large scale and easily accessible Chinese

spoken language corpus available to researchers. However, most of the major

contemporary Chinese language corpus includes a wide variety of written language

samples such as literature, newspaper, academic, blogs and micro blogs. And the

inclusion of many informal sources and on-line media would to some degree reflect the

common and everyday use of Chinese language under a spoken context.

All the V-N collocations that could not be judged as conventional on the basis of

dictionaries and the CCL were then presented to expert raters for acceptability

judgements. The two primary expert raters were native speakers of mandarin Chinese

from mainland China and both were Chinese instructors teaching at the college level in

the U.S. The V-N collocations were presented in context (highlighted in the original

sentence produced by the learners) to the native speaker experts, so that the collocations

could be judged in relation to their intended meaning in context. The raters were asked to

judge the collocations on a five-point scale from completely unacceptable (1)

unacceptable (2) not sure (3) acceptable (4) to completely acceptable (5). Whenever a

combination was judged "not sure", "unacceptable" or "completely unacceptable", the

raters were also asked to provide an acceptable or better option to express the intended

meaning.

Initially, two native speaker raters were asked to judge and rate all the non-

conventional V-N collocations used by the learners. If their ratings were the same, their

judgement constituted the final score for the collocation. If their ratings had a difference of 1 point, an average score would be calculated to be the final score. If the two raters' scores had a difference of two or more points, one additional native speaker rater was asked to provide a judgement. After a third score was given for the collocation, an average score would be calculated using all three scores.

This method of judging the acceptability of collocations produced by learners is limited and remains an approximation. As pointed out by Nesselhauf (2004), "there is not necessarily a one-to-one relation between what native speakers find acceptable or unacceptable when explicitly asked about a certain language phenomenon and what they themselves produce frequently" (p. 53) However, a number of studies have argued that there is good correlation between corpus data and native speaker judgement on lexical co-occurrence (e.g., Lapata et al, 1999; Hoffmann & Lehmann 2000; Shei 1999, Neselhauf, 2004). For the present study, the number of native speaker raters for judging the acceptability of collocations is relatively small and therefore the acceptability score for each collocation should be used only as an approximate number to show the overall trend and characteristics.

The second step is to determine the communicativeness of all unconventional collocations. The term "communicativeness" is adopted in the current study to mean if a speaker can convey his/her intended meaning to a listener through the use of certain expressions. In previous literature, most researchers have chosen to study only acceptability of collocations. However, as communicative competence has increasingly been emphasized by language educators at an important goal of learning English and as

more researchers start to view learner language as "asset" rather than "deficit", it seems equally important to find out how listeners perceive the spoken language produced by learners.

As an attempt to determine the communicativeness of learner language, all unconventional collocations were again presented to the two expert raters for judgments and their judgments are based on three scales: (1) not communicative (2) not sure and (3) communicative. The two raters' judgments would then be compared and analyzed.

4.3. Data analysis and hypothesis

4.3.1. Data analysis for Q1

What is the difference (type and token) in using Chinese verb-noun collocations among CFL learners, CHL learners and CNSs in an oral interview context?

To find out answer to research question 1, both quantitative and qualitative analyses of the data were carried out.

The research hypothesis as related to research question 1 is that there is significant difference among the three groups, especially between CFL learners and CNSs. As discussed in Chapter 2, many studies comparing the collocation usage between L2 learners and native speakers have indicated that the two groups are significantly different (Bahns & Eldaw, 1993; Biskup, 1992; Fan, 2009; Farghal & Obiedat, 1995; Granger, 1998; Howarth, 1996, 1998; Lorenz, 1999; Nesselhauf, 2004) both in terms of the number and the range of collocations produced.

To test this hypothesis, two multivariate analysis of variance (ANOVA) tests are carried out to investigate the influence of language background (CFL learners, CHL

78

learners and CNSs) on the types of V-N collocations produced and on the tokens of V-N collocations produced. Multivariate ANOVA tests are chosen for research question 1 because there are two dependent variables involved which are Vms and Vds and the two dependent variables are probably related.

In addition to the statistical analysis, I also performed a qualitative analysis of the characteristics of V-N collocations used by the three groups. By examining a number of examples from the spoken datasets, similarities and differences in producing Vms and Vds collocations across three groups are summarized and discussed.

4.3.2. Data analysis for Q2

How do two types of topics (daily conversation topics vs. abstract academic topics) affect CFL learners, CHL learners and NSs' use of Chinese verb-noun collocations?

The underlying research hypothesis for research question 2 is that types of topics will have a significant influence on the three groups' collocation usage. More specifically, I expect that in discussing daily topics, Vms-N collocations will be used more often than Vds-N collocations whereas in discussing abstract topics, Vds-N collocations will be used significantly more often than Vms-N collocations. This hypothesis is based on the unique linguistic feature of Vms and Vds in Chinese language. As mentioned in chapter 2, in modern Mandarin Chinese, Vms is more closely associated with informal oral language while Vds appears more often in formal and academic written and spoken language (Feng, 2003; Liu, 2007).

In order to assess the hypothesis for research question 2, a series of mixed methods ANOVA tests are performed to find out if the three groups used different Vms-

N collocations and Vds-N collocations in speaking about the two types of topics. Mixed methods ANOVA is a statistical test typically used when different groups of participants are tested multiple times or under multiple conditions and their responses over time or over conditions need to be compared. In the current study, three groups of participants are interviewed on two types of topics which can also be considered as two conditions and the question aims to examine the effect of the two conditions on participants' responses. Therefore, it is reasonable to employ such a method for testing the hypothesis.

For qualitative analysis of research question 2, the collocation use of all three groups are interpreted in relation to the two types of topics. Excerpts and examples from the data would illustrate the features of learners' collocation use in daily topic conversation and in abstract academic topic discussion, as compared to native speakers.

4.3.3. Data analysis for Q3

What are the characteristics of non-conventional V-N collocations produced by CFL learners and CHL learners?

For research question three, the analysis of learners' collocation usage is largely based on the non-conventional collocations used by the two learner groups. The "non-conventional" collocations are defined as learner-produced collocations which are not included in either Chinese dictionaries or large modern Chinese corpora. Answer to the third research question entails a largely qualitative analysis of the different types of non-conventional collocations produced by CFL learners and CHL learners, as well as an analysis on the acceptability and communicativeness of non-conventional collocations.

After descriptive statistics regarding the number and percentage of non-conventional V-N combinations made by learners are presented, the study will present different categories of non-conventional V-N combinations based on lexical and semantic features of these combinations. A more detailed discussion of the factors which lead to the different categories of non-conventional collocational usage will be provided in Chapter 6.

CHAPTER 5: FINDINGS

This chapter summarizes both quantitative data and qualitative data as obtained from the oral interviews with participants in the three research groups: Chinese foreign language learners (CFL learners), Chinese heritage language learners (CHL learners) and Chinese native speakers (CNS). Findings for each research question will be described below in details.

5.1. Difference of CNS, CFL and CHL in producing V-N collocations

The first research question concerned the difference (type and frequency) in using Chinese V-N collocations among CNSs, CFL learners and CHL learners.

A quantitative analysis performed on the types and tokens of verbs (both monosyllabic verbs and disyllabic verbs) in V-N collocations revealed significant difference among the three groups. Qualitative data reporting the most frequently used V-N collocations also displayed different preferences in using collocations across groups.

5.1.1. Overall production of V-N collocations across the three groups

5.1.1.1. Descriptive statistics of spoken datasets used in the study

Data used for this research is composed of spoken language recordings collected from 30 interviews with the participants. All interviews were transcribed and entered into three spoken language datasets: CNSs dataset, CFL dataset and CHL dataset. Instances of all V-N collocations were manually identified in the three datasets. Thus, I would like to first present the descriptive statistics of the three datasets.

Each dataset consists of the transcribed recordings of the 10 participants in one research group. Means and Standard deviations of recording time and of number of Chinese characters for each group were reported in Table 4.

Table 4. Total recording time and total Chinese characters for the three datasets

Spoken language dataset	Number of participants	Interview time (min.)		Number of characters	
		Mean	**Std.**	**Mean**	**Std.**
CNS dataset	**10**	**56.3**	**4.2**	**5024.8**	**391.0**
CFL dataset	**10**	**53.7**	**2.5**	**4694.3**	**293.0**
CHL dataset	**10**	**54.9**	**3.4**	**4856.2**	**355.0**

In order to obtain comparable amount of data for each participant, I controlled the length of each interview to be about the same (around one hour), as shown in the mean and standard deviations of interview time for three groups (Table 4). To verify if the numbers of characters for the three datasets were also similar, a one-way ANOVA test was performed (Table 5). Results of the ANOVA test revealed no statistical difference among the total sizes of the three datasets ($p > .05$).

Table 5. Results of ANOVA test comparing the mean number of characters for each dataset

Tests of Between-Subjects Effects

Dependent variable: number of characters

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Language background	546226.067	2	273113.033	2.242	.126	.142

a. R Squared = .142 (Adjusted R Squared = .079)

The comparable sizes of the three datasets allow direct comparison of collocational usage in the three datasets. Table 6 reports the descriptive statistics for

analyzing research question 1. Mean values and standard deviations of the types and

tokens of verb (monosyllabic verbs and disyllabic verbs)-noun collocations for each

group are provided. Figure 1 and figure 2 further compares means of type and token for

V-N collocations across the three groups.

Table 6. Mean and standard deviation of Vms and Vds (both in type and token) between three groups

	Language background	Mean	Std. Deviation
Vms-Type	CFL	14.1000	4.35762
	CHL	18.8000	3.76534
	CNS	24.4000	4.29987
Vms-Token	CFL	22.3000	6.37791
	CHL	22.1000	5.80134
	CNS	51.5000	6.04152
Vds-Type	CFL	27.3000	8.04225
	CHL	32.0000	6.96020
	CNS	38.3000	5.41705
Vds-Token	CFL	58.3000	15.41320
	CHL	52.4000	11.06747
	CNS	74.0000	9.32142
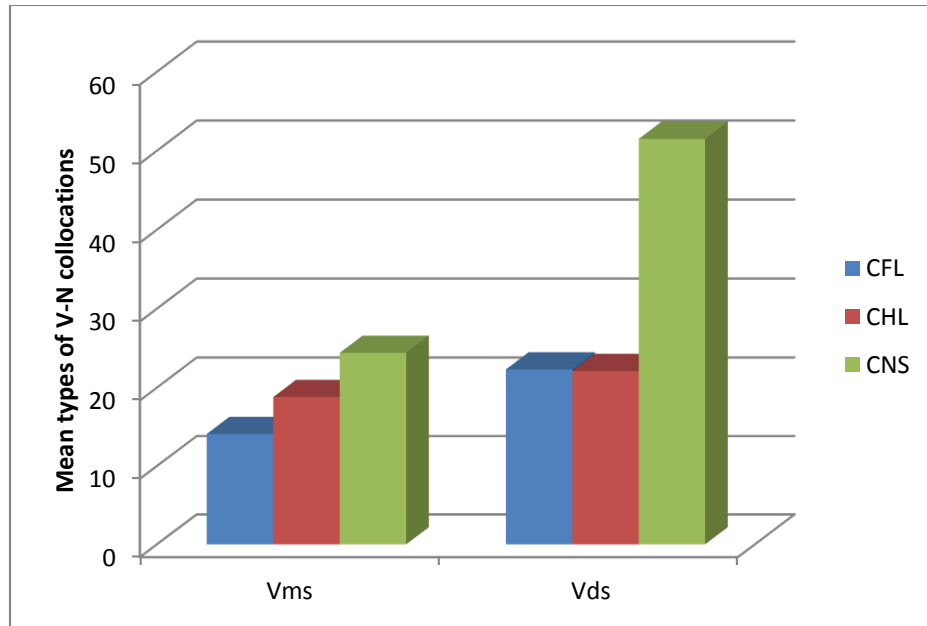
Figure 1. Comparison of means types of Vms and Vds collocations among three groups
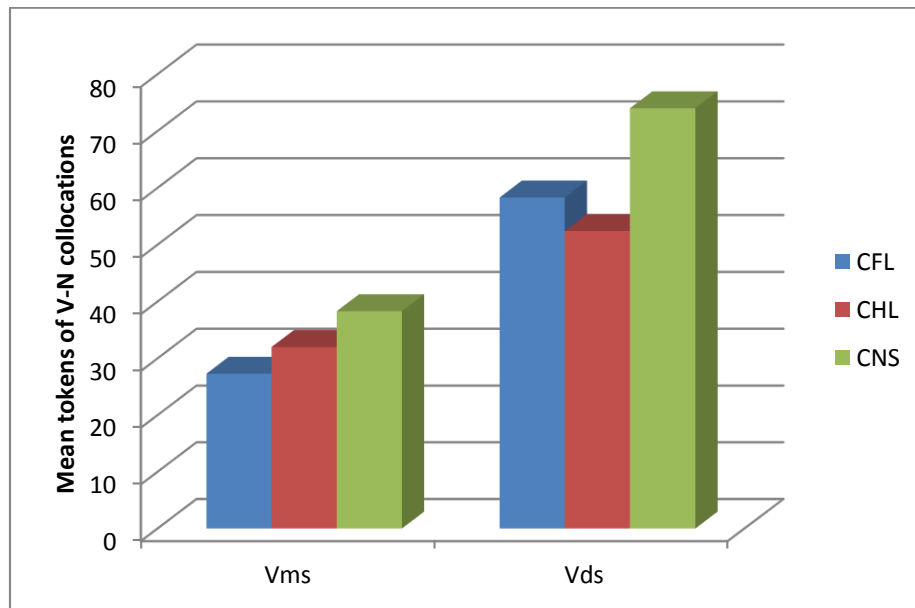
Figure 2. Comparison of mean tokens of Vms and Vds collocations among three groups

5.1.1.2. Inferential statistics of the type and token of verbs in V-N collocations between

three groups

Research question one asks whether language background can affect the way

speakers use Chinese V-N collocations. Since Chinese V-N collocations are comprised of

both Vms-N collocations and Vds-N collocations, it is necessary to examine the effect of

language background on both Vms and Vds collocations. One-way multivariate analysis

of variance (one-way MANOVA) is used to determine whether there are any differences

between independent groups on more than one dependent variable. For research question

one, two one-way MANOVA tests were conducted. The first test was run to understand

whether there were differences in the types of Vms-N and Vds-N collocations used by the

three groups. And the second test was done to investigate the differences in the tokens of

Vms-N and Vds-N collocations used across the three groups.

In order to ensure that the assumption of homogeneity of variance was met,

Levene tests were performed before running the MANOVA. The results indicated that

no significant differences ($p >.05$) were observed in the variances of the three groups on

verb types and verb tokens; thus the assumption was met. Results of the Levene's test of

homogeneity of variance are presented in Table 7.

Table 7. Levene's Test of Equality of Error Variances

	F	df1	df2	Sig.
TypeVms	.351	2	27	.707
TypeVds	.380	2	27	.687
TokenVms	.418	2	27	.662
TokenVds	2.174	2	27	.133

86

The assumption of multicollinearity is further checked by conducting correlations between the dependent variables of the two MANOVA tests. Regression analysis indicated that multicollinearity was not a concern (Type Vms-Vds, Tolerance = .55, VIF = 1.80; Token Vms-Vds, Tolerance = .70, VIF = 1.42). Results of the multicollinearity tests are reported in table 8.

Table 8. Collinearity Statistics

Model	Collinearity Statistics	
	Tolerance	VIF
TypeVds-TypeVms	0.55	1.80
TokenVds-TokenVms	0.70	1.42

Results of the two MANOVA tests were reported in Table 9. Significant difference was found among the three groups in producing types of Vms [F(2, 15.443), p < .001, η_p^2=.534], types of Vds [F(2, 77.460), p < .001, η_p^2=.852], tokens of Vms [F(2, 6.415), p = .005, η_p^2=.322], and tokens of Vds [F(2, 8.366), p = .001, η_p^2=.383]. The results demonstrate that CNSs, CFL learners and CHL learners have large difference in producing types and tokens of V-N collocations.

Table 9. Combined results of the two one-way MANOVA tests on the effect of language background on Vms and Vds

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Language background	Vms Type	531.800	2	265.900	15.443	.000	.534
	Vds Type	5723.467	2	2861.733	77.460	.000	.852
Language background	Vms Token	609.267	2	304.633	6.415	.005	.322
	Vds Token	2492.867	2	1246.433	8.366	.001	.383

As a further step to understand the variation between different groups in using V-N collocations, Post Hoc tests were conducted using SPSS. Turkey HSD Post Hoc tests, as shown in Table 10 and Table 11, further revealed several important findings:

(1) With regard to the types of Vms in collocations, CNSs used significantly more types of Vms than both CFL learners ($p < .001$) and CHL learners ($p < .02$). But the results failed to show a significant difference between CHL and CFL learners in their production of the the types of Vms generated ($p > .05$)

(2) With regard to the types of Vds in collocations, CNSs also used significantly more types of Vds than both CFL learners ($p < .001$) and CHL learners ($p < .001$). But the results again failed to show statistical difference between CFL learners and CHL learners in the types of Vds used ($p > .05$).

(3) In terms of the tokens of Vms in collocations, CFL learners used significantly less Vms than CNSs ($p < .005$). But no statistical difference was found between CHL learners and CFL learners ($p > .05$) or between CHL learners and CNS ($p > .05$) in the tokens of Vms.

(4) In terms of the tokens of Vds in collocations, CNSs again produced significantly more tokens of Vds than both CFL learners ($p < .05$) and CHL learners ($p < .002$). And no significant difference was found between CFL learners and CHL learners in the tokens of Vds used by them ($p > .05$).

Table 10. Multiple comparisons of groups for types of Vms and Vds using Turkey HSD

Dependent Variable	(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Vms-Type	CFL	CHL	-4.70*	1.86	0.05	-9.44	0.04
		CNS	-10.30*	1.86	0.00	-15.04	-5.56
	CHL	CFL	4.70*	1.86	0.05	-0.04	9.44
		CNS	-5.60*	1.86	0.02	-10.34	-0.86
	CNS	CFL	10.30*	1.86	0.00	5.56	15.04
		CHL	5.60*	1.86	0.02	0.86	10.34
Vds-Type	CFL	CHL	0.20	2.72	1.00	-6.74	7.14
		CNS	-29.20*	2.72	0.00	-36.14	-22.26
	CHL	CFL	-0.20	2.72	1.00	-7.14	6.74
		CNS	-29.40*	2.72	0.00	-36.34	-22.46
	CNS	CFL	29.20*	2.72	0.00	22.26	36.14
		CHL	29.40*	2.72	0.00	22.46	36.34

Based on observed means. The error term is Mean Square(Error) = 36.944.*. The mean difference is significant at the .05 level

Table 11. Multiple comparisons of groups for tokens of Vms and Vds using Turkey HSD

Dependent Variable	(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Vms-Token	CFL	CHL	-4.70	3.08	0.42	-12.57	3.17
		CNS	-11.00*	3.08	0.00	-18.87	-3.13
	CHL	CFL	4.70	3.08	0.42	-3.17	12.57
		CNS	-6.30	3.08	0.15	-14.17	1.57
	CNS	CFL	11.00*	3.08	0.00	3.13	18.87
		CHL	6.30	3.08	0.15	-1.57	14.17
Vds-Token	CFL	CHL	5.90	5.46	0.87	-8.03	19.83
		CNS	-15.70*	5.46	0.02	-29.63	-1.77
	CHL	CFL	-5.90	5.46	0.87	-19.83	8.03
		CNS	-21.60*	5.46	0.00	-35.53	-7.67
	CNS	CFL	15.70*	5.46	0.02	1.77	29.63
		CHL	21.60*	5.46	0.00	7.67	35.53

Based on observed means. The error term is Mean Square(Error) = 148.981. *. The mean difference is significant at the .05 level

5.1.2. Qualitative analysis on the difference in using collocations between the three groups

    In the statistical analysis of findings for researcher questions 1, I compared the means and standard deviations of the types and tokens of V-N collocations produced by the 30 participants of the three research groups. In order to examine more closely the actual cases of V-N collocations produced by learners and native speakers in the three datasets, I have also carried out a qualitative analysis of some of the most frequently used V-N collocations in the three datasets.

5.1.2.1. Frequently used Vms and Vds in V-N collocations of the three datasets

    Table 12 summarizes the 5 most frequently used Vms and Vds in V-N collocations as appearing in each dataset. In addition to listing the total number of appearances for each verb, the number of participants who used the same verb in forming V-N collocations has also been listed to show that the frequency of appearance for most of these verbs is distributed quite evenly among participants. In other words, most participants tend to choose and repeat a number of verbs in forming V-N collocations.

Table 12. Five most frequently occurring Vms and Vds in the three datasets

Group	High Frequency Vms	Token	No. of participants using the Vms	High Frequency Vds	Token	No. of participants using the Vds
CFL	看 (to see)	30	10	学习 (to learn)	26	10
	去 (to go)	29	10	了解 (to understand)	25	9
	学 (to learn)	27	10	解决 (to solve)	24	8
	做 (to do, to make)	26	10	支持 (to support)	18	6
	打 (to hit, to play)	19	9	参加 (to participate)	18	10
CHL	学 (to learn)	27	10	喜欢 (to like)	24	10
	做 (to do)	25	10	学习 (to study)	22	10
	去 (to go)	25	9	解决 (to solve)	22	8
	看 (to see)	23	10	帮助 (to help)	18	7
	吃 (to eat)	15	9	了解 (to understand)	16	9
CNS	做 (to make, to do)	21	10	准备(to prepare)	6	4
	学 (to learn)	17	9	参加 (to participate)	6	6
	上 (to take)	15	8	进行 (to carry out)	5	3
	打 (to hit, to play)	13	9	了解 (to understand)	5	4
	选 (to choose)	13	7	产生 (to produce)	4	4

Based on the information in Table 12, we could find some interesting features of the three research groups' preference for Vms and Vds in V-N collocations.

(1) A few verbs appeared frequently in all three datasets. Vms such as "做"(to do, to make) and "学"(to study) and "打" (to hit, to play) together with Vds such as "了解" (to understand) were commonly produced by many CFL learners, CHL learners and CNS. Among these most frequently used verbs for V-N collocations, some are Chinese light verbs such as "做"(to do, to make) and "打" (to hit, to play); and some are verbs that appear often in Chinese spoken language such as "学"(to study) and "了解" (to understand).

91

(2) CFL learners and CHL learners shared more similarities in choosing Vms and Vds whereas the choice of verbs for collocations was more diverse for CNSs. Among the 5 most commonly appearing Vms used by CFL learners and CHL learners in V-N collocations, 4 of them were the same ("看, *to see*", "去, *to go*", "学, *to study*", and "做, *to do*"). Similarly, among the 5 most frequent Vds used by CFL learners and CHL learners, 3 Vds are the same ("学习, *to learn*", "了解, *to understand*", "解决, *to solve*"). The overlap in the most frequent Vms and Vds between CNS and the two learner groups is much smaller.

(2) CFL learners and CHL learners tend to repeat the same verbs more often than CNSs in forming V-N collocations, as shown in Figure 3. For example, the total tokens of the 5 most frequently used Vms by CFL learners are 92 times, which account for 34% of the total numbers of Vms produced by CFL learners in V-N collocations. The total tokens of the 5 most frequent Vds by CFL learners are 112 times, which account for about 19% of the total numbers of Vds used by CFL learners. For CHL learners, the 5 most frequently appearing Vms take up 29% of all Vms and the 5 most frequently appearing Vds take up about 20% of all Vds used by heritage learners in V-N collocations. In comparison, CNSs' usage of the 5 most frequently Vms and Vds only account for 10% and 4% of their total production of Vms and Vds in V-N collocations.
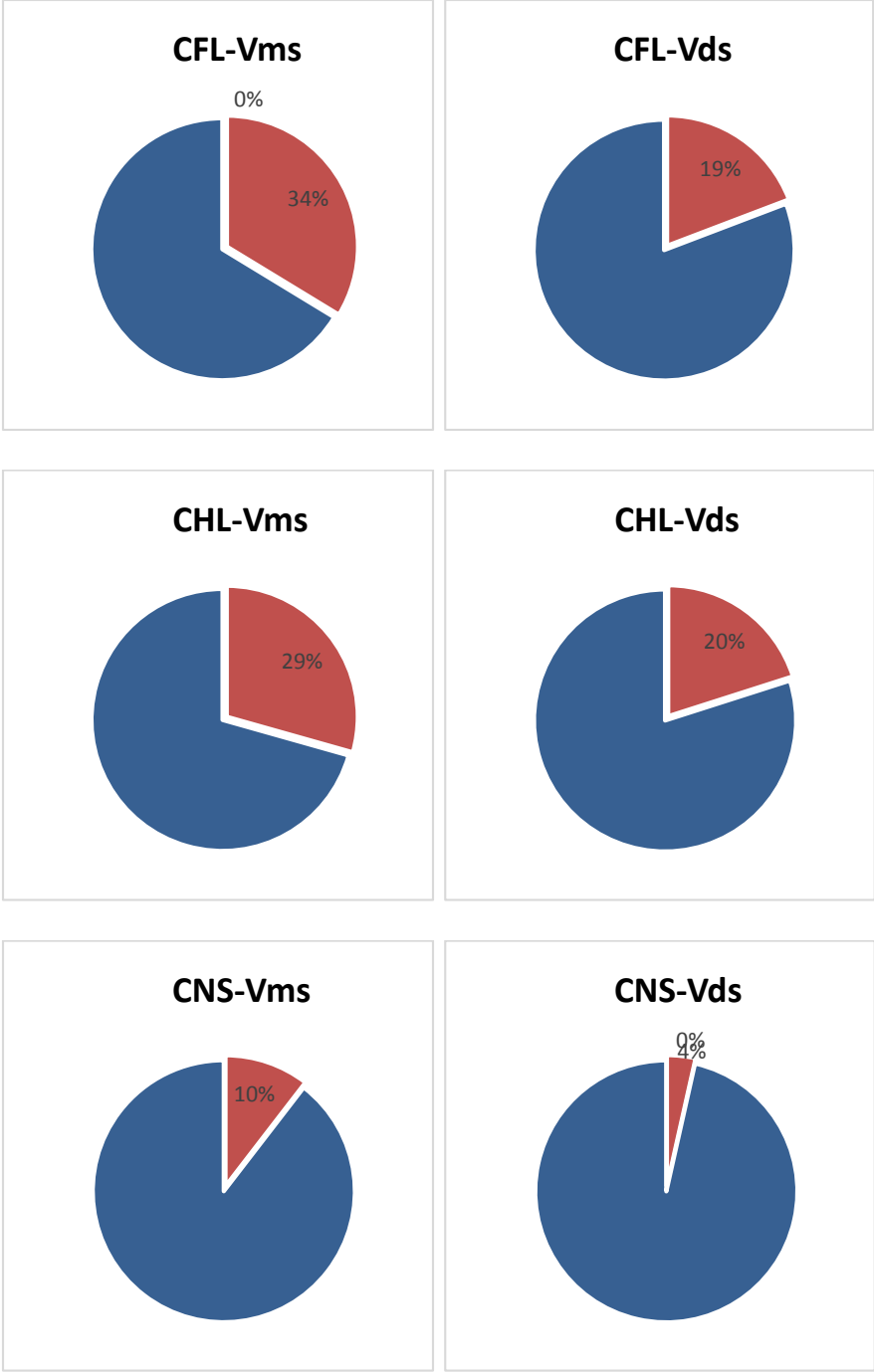
Figure 3. Percentage of 5 most frequently used V-N collocations in total collocations

5.1.2.2. Frequently used types of V-N collocations in the three datasets

      In addition to examining the common verbs in V-N collocations, I also manually identified types of V-N collocations most frequently appearing in the three spoken datasets. Table 13 recorded the five types of V-N collocations most regularly occurring in the datasets. The token and the number of participants using the same type of collocation are both listed to show that most participants have produced these common types of V-N collocation in their language.

      By observing and analyzing Table 13, the following characteristics of the most frequently produced V-N collocations of the three groups could be drawn:

      (1) Overall, CFL learners and CHL learners tend to employ the same types of collocations more often than CNSs. For example, the collocation type "学汉语/学英语" appear 17 times in CFL dataset, 15 times in CHL dataset, but only 7 times in CNS datatset. This is not surprising considering the statistical finding that CNS used a significantly greater range of different types of V-N collocations than both CFL learners and CHL learners.

      (2) Regarding the most frequent collocation types, CFL learners and CHL learners share similar preferences for certain types of collocations. For example, collocations such as "上中学/上大学"(*to attend middle school/college*), "学汉语/学英文"(*to learn Chinese/English*), "去美国/去中国"(*to go to the United States/China*) , "保护环境"(*to protect environment*) , "解决问题 (to solve problems)*" and "了解文化/了解历史"(*to understand culture/history*) occur frequently in both CFL and CHL datasets. In comparison, CNS dataset does not contain many similar common collocation types as the

94

two learner groups which also shows that CNS tend to employ different kinds of

collocations than learners.

Table 13. Most frequent types of V-N collocations identified in the spoken datasets of the three groups

	High Frequency Vms+n	Token	No. of participants	High Frequency Vds+n	Token	No. of participants
CFL	学汉语（中文）(learn Chinese)	17	10	解决问题 (solve problems)	17	9
	去中国（北京，上海等）(go to Beijing, Shanghai, etc.)	11	9	保护环境 (protect environment)	16	9
	说中文 (speak Chinese)	10	10	学习汉语（中文等）(learn Chinese)	14	8
	上中学（大学等）(go to high school, college, etc.)	9	8	了解文化 (历史等) (learn culture, history etc.)	10	8
	买东西(buy something)	6	5	采取措施 (adopt policy)	7	5
CHL	学汉语（中文）(learn Chinese)	15	10	了解文化 (历史等) (learn culture, history etc.)	13	8
	上中学（大学等）(go to high school, college, etc.)	10	8	参加活动(运动等)(participate in activities, sports, etc.)	11	7
	去中国（北京，上海等）(go to Beijing, Shanghai, etc.)	8	9	解决问题 (*solve problems*)	9	9
	坐飞机（高铁等）(take planes, high speed trails, etc.)	7	5	保护环境 (protect environment)	9	7
	吃中国菜（米饭等）(eat Chinese foods)	7	7	学习汉语（中文等）(learn Chinese)	8	7
CNS	上大学/上中学(go to college, go to high school, etc.)	11	8	练习钢琴（书法等）(practice piano, *calligraphy, etc.*)	7	5
	选专业 (choose majors)	7	5	关注问题 (pay attention to problems)	7	6
	学英文/学外语 (learn English, foreign languages, etc.)	7	7	进行改革 (carry out reformation)	6	5
	读本科 (study at college)	6	6	培养人才 (train people)	4	4
	花时间 (spend time)	4	3	找到工作(find jobs)	4	4

5.2. The effect of topic on the production of V-N collocations of CNS, CFL and CHL

Research question two concerns the effect of different topics on the three groups' production of V-N collocations. In order to find out about the effect of topic, a series of mixed method ANOVA tests are performed. Mixed method ANOVA is a statistical test typically used when participants of a research study are tested multiple times or under multiple conditions and their responses over time or over conditions need to be compared.

In the current study, all participants are interviewed on two types of topics which can also be considered as two conditions and the question aims to examine the effect of the two conditions on participants' responses. More specifically, four different mixed method ANOVA tests were carried out using SPSS to examine the influence of topic on the types of Vms-N collocations, types of Vds-N collocations, tokens of Vms-N collocations, and tokens of Vds-N collocations respectively. Based on the overall results of the tests, the two oral topics did have a significant effect on the three groups' collocation usage.

5.2.1. Statistical findings on the effect of topic on V-N collocations across groups

5.2.1.1. Descriptive statistics and comparison of means

To answer the second research question, the dataset for each group was further divided into two parts: one consisted of speakers' discussion of academic topics and another consisted of speakers' discussion of daily topics.

Table 14 and Table 15 summarized the means and standard deviations of verb types and verb tokens under the two types of topics as obtained from the three datasets.

96

Figure 4 and Figure 5 displayed the overall comparison of means (in type and in token) of

the three groups in using Vms and Vds across two different topics. The bar graphs clearly

showed that the three groups do not produce even amounts of V-N collocations in

academic topics and in daily topics. Also, CNSs tended to use more V-N collocations

than both CFL learners and CHL learners in discussing both academic topic and daily

topic.

Table 14. Mean and standard deviation per person in producing types of Vms and Vds under two topics

	Academic topic				Daily topic			
	Vms-Type		Vds-Type		Vms-Type		Vds-Type	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CFL	5.9	3.87	16.1	5.28	12.3	3.12	11.3	2.98
CHL	8.5	2.91	15.8	4.49	15.7	3.68	12.8	2.44
CNS	9.2	2.20	38.4	7.41	19.5	3.53	16.5	2.63

Table 15. Mean and standard deviation per person in producing tokens of Vms and Vds under two topics

	Academic topic				Daily topic			
	Vms-Token		Vds-Token		Vms-Token		Vds-Token	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
CFL	8.4	4.94	18.9	5.06	40.5	11.4	17.8	5.82
CHL	12.3	3.74	19.7	4.19	31.8	7.77	20.6	4.81
CNS	12.7	2.11	25.6	4.29	50.5	7.10	23.5	3.02
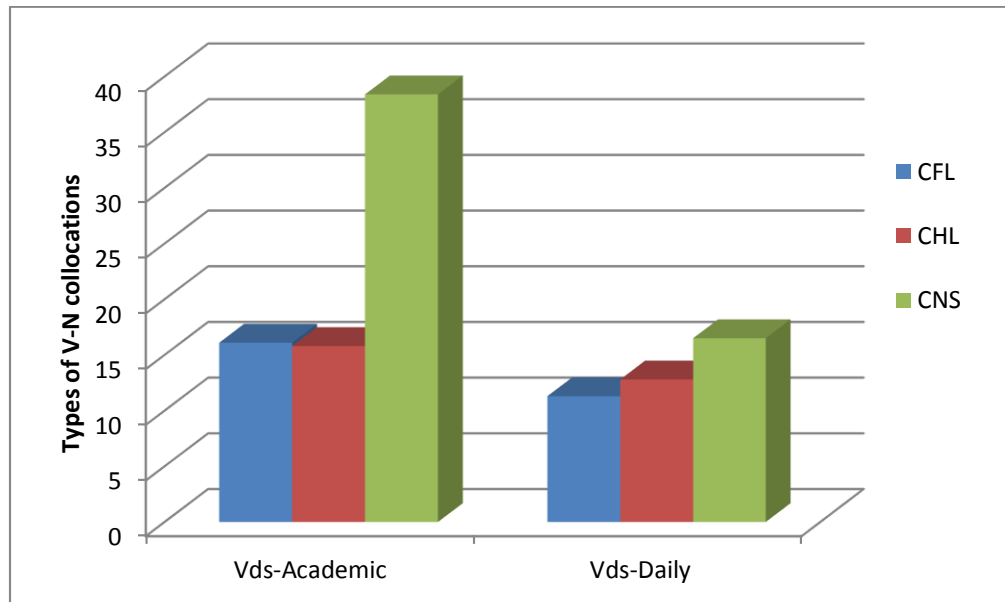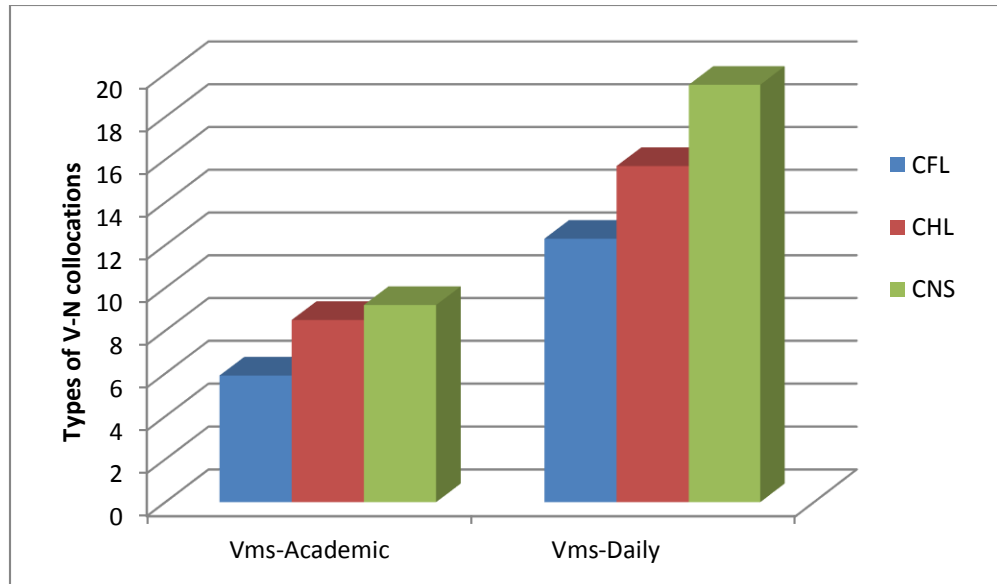
Figure 4. Comparison of mean values of types of Vms and Vds under two different oral topics
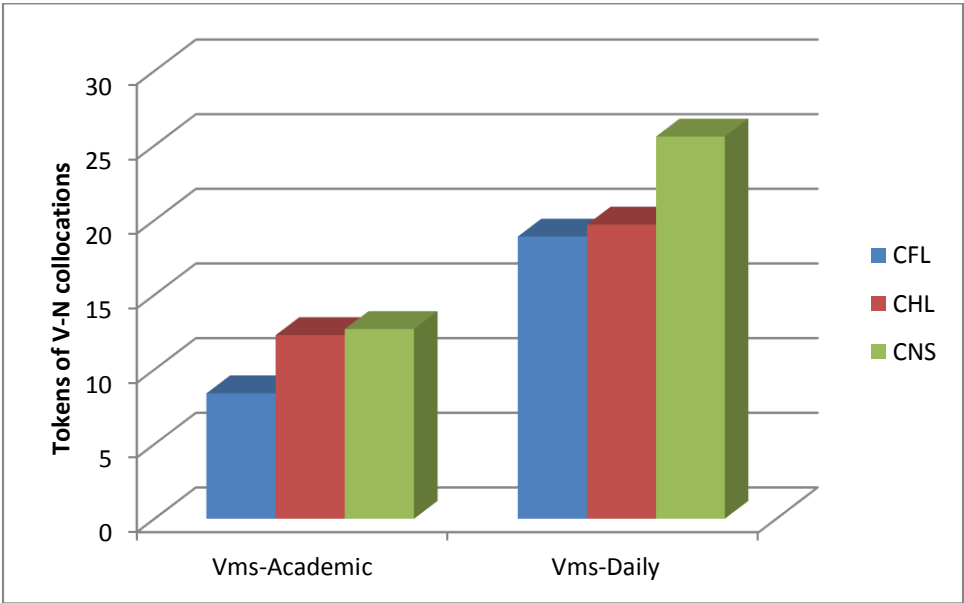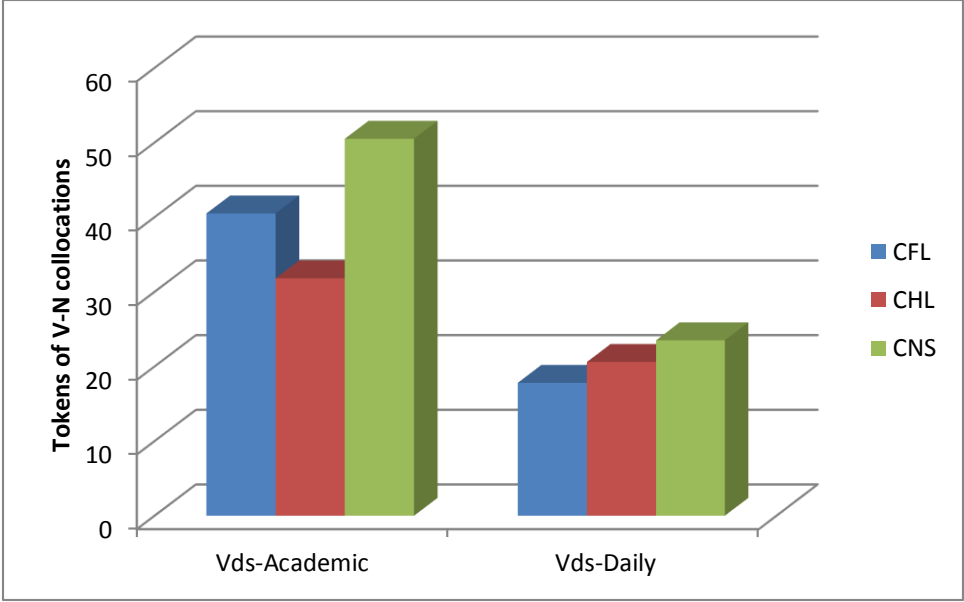
Figure 5. Comparison of mean values of tokens of Vms and Vds under two different oral topics

5.2.1.2. Results of mixed method ANOVA tests

The results of the four mixed method ANOVA tests were shown in Table 16 through Table 23.

Table 16 displays the test results of the mixed method ANOVA on the influence of topic (academic topic vs. daily topic) on the types of Vms used by the three groups. For this test, independent variable was language background and dependent variable was type of Vms used in academic topic discussion and Vms used in daily topic discussion.

The results show that there was a significant effect of Topic (academic topic vs. daily topic) on types of Vms, $F (1, 27)=103.216$, $p < .001$, $\eta_p^2 =.793$, and a significant effect for Language background, $F (2, 27)=11.399$, $p < .001$, $\eta_p^2 =.458$. And there was no significant interaction effect between Topic and Language background, $F (2, 27)=2.3$, $p > .05$, $\eta_p^2 =.146$.

All three groups showed significantly differences in using Vms-N collocations under the two topics. And the insignificant interaction for topic and language background suggested that the different usage under the two topics was similar across groups. Turkey HSD Post Hoc analyses was also conducted (see Table 17). Significant difference was found between CNSs and CFL learners, and between CFL learners and CHL learners. But the results failed to find any significant difference no significant difference between CNSs and CHL learners.

100

Table 16. Results of mixed method ANOVA for the mean types of Vms of three groups in the academic topic discussion and daily topic discussion

		Sum of squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subjects Effects	**topic**	952.017	1.000	952.017	103.210	.000	.793
	topic * Language background	42.433	2.000	21.217	2.300	.120	.146
	Error	249.050	27.000	9.224			
Between-Subjects Effects	**topic * Language background**	277.500	2	138.750	11.399	.000	.458
	Error	328.650	27	12.172			

Table 17. Multiple comparisons of groups for types of Vms produced under academic and daily topics using Turkey HSD (α = .05)

(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CFL	**CHL**	-3.0000*	1.10328	.029	-5.7355	-.2645
	CNS	-5.2500*	1.10328	.000	-7.9855	-2.5145
CHL	**CFL**	3.0000*	1.10328	.029	.2645	5.7355
	CNS	-2.2500	1.10328	.122	-4.9855	.4855
CNS	**CFL**	5.2500*	1.10328	.000	2.5145	7.9855
	CHL	2.2500	1.10328	.122	-.4855	4.9855

Based on observed means. The error term is Mean Square(Error) = 6.086. *. The mean difference is significant at the .05 level

Table 18 summarizes the test results of the mixed method ANOVA of the influence of topic on the types of Vds used by the three groups. The results revealed a significant effect of Topic (academic topic vs. daily topic) on types of Vds, $F(1, 27)=99.422$, $p < .001$, $\eta_p^2 = .786$, a significant interaction effect between Topic and Language background, $F(2, 27)=36.792$, $p < .001$, $\eta_p^2 = .732$, and a significant effect for Language background, $F(2, 27)=45.052$, $p < .001$, $\eta_p^2 = .769$.

All three groups were significantly different in using Vds-N collocations in discussing academic and daily topics. And the significant interaction for topic and language background suggested that the different usage under the two topics differed across groups. Turkey HSD Post Hoc analyses was also conducted (see Table 19). Significant difference was found between CNSs and both learner groups. But the analysis failed to find any significant difference between CFL learners and CHL learners.

Table 18. Results of mixed method ANOVA for the mean types of Vds of three groups in the academic topic discussion and daily topic discussion

		Sum of squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subjects Effects	topic	1470.150	1.000	1470.150	99.422	.000	.786
	topic * Language background	1088.100	2.000	544.050	36.792	.000	.732
	Error	399.250	27.000	14.787			
Between-Subjects Effects	Language background	2415.633	2	1207.817	45.052	.000	.769
	Error	723.850	27	26.809			

Table 19. Multiple comparisons of groups for types of Vds produced under academic and daily topics using Turkey HSD (α = .05)

(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CFL	CHL	-.60	1.63	.929	-4.65	3.45
	CNS	-13.75*	1.63	.000	-17.80	-9.69
CHL	CFL	.60	1.63	.929	-3.45	4.65
	CNS	-13.15*	1.63	.000	-17.20	-9.09
CNS	CFL	13.75*	1.63	.000	9.69	17.80
	CHL	13.15*	1.63	.000	9.09	17.20

Based on observed means. The error term is Mean Square(Error) = 6.086. *. The mean difference is significant at the .05 level

Table 20 summarizes the test results of the mixed method ANOVA of the influence of topic on the tokens of Vms used by the three groups. The results revealed a

102

significant effect of Topic (academic topic vs. daily topic) on types of Vds, F (1,

27)=141.87, $p < .001$, $\eta_p^2 = .840$, a significant interaction effect between Topic and

Language background, F (2, 27)=3.441, $p < .05$, $\eta_p^2 = .732$, and a significant effect for

Language background, F (2, 27)=6.415, $p < .05$, $\eta_p^2 = .322$.

Tokens of Vms-N collocations produced by all three groups were significantly

different under the two topics. And the significant interaction for topic and language

background suggested that the different usage under the two topics differed across

groups. Turkey HSD Post Hoc analysis was also conducted (see Table 21). The analysis

indicates significant difference between CNSs and CFL learners, but fails to show any

significant difference between CNSs and CHL learners or between CFL learners and

CHL learners.

Table 20. Results of mixed method ANOVA for the mean tokens of Vms of three groups in the academic topic discussion and daily topic discussion

		Sum of squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subjects Effects	**topic**	1581.067	1.000	1581.067	141.870	.000	.840
	topic * Language background	76.033	2.000	38.017	3.411	.048	.202
	Error	300.900	27.000	11.144			
Between-Subjects Effects	**Language background**	304.633	2	152.317	6.415	.005	.322
	Error	641.100	27	23.744			

Table 21. Multiple comparisons of groups for tokens of Vms produced under academic and daily topics using Turkey HSD (α = .05)

(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
CFL	CHL	-2.35	1.54	.295	-6.17	1.47
	CNS	-5.50*	1.54	.004	-9.32	-1.67
CHL	CFL	2.35	1.54	.295	-1.47	6.17
	CNS	-3.15	1.54	.121	-6.97	.67
CNS	CFL	5.50*	1.54	.004	1.67	9.32
	CHL	3.15	1.54	.121	-.67	6.97

Based on observed means. The error term is Mean Square(Error) = 6.086. *. The mean difference is significant at the .05 level

Table 22 summarizes the results of the mixed method ANOVA of the influence of topic on the tokens of Vds used by the three groups. The results revealed a significant effect of Topic (academic topic vs. daily topic) on tokens of Vds, $F(1, 27)=221.981$, $p < .001$, $\eta_p^2 =.892$, a significant interaction effect between Topic and Language background, $F(2, 27)=11.982$, $p < .001$, $\eta_p^2 = .470$, and a significant effect for Language background, $F(2, 27)=8.366$, $p < .05$, $\eta_p^2 =.383$.

Tokens of Vds-N collocations produced by all three groups were significantly different under the two topics. And the significant interaction for topic and language background suggested that the different usage under the two topics differed across groups. Turkey HSD Post Hoc analysis was also conducted (see Table 23). Results revealed significant difference between CNSs and both learners groups. But the test failed to find any significant difference between CFL learners and CHL learners.

Table 22. Results of mixed method ANOVA for the mean tokens of Vds of three groups in the academic topic discussion and daily topic discussion

		Sum of squares	df	Mean Square	F	Sig.	Partial Eta Squared
Within-Subjects Effects	topic	6181.350	1.000	6181.350	221.981	.000	.892
	topic * Language background	667.300	2.000	333.650	11.982	.000	.470
	Error	751.850	27.000	27.846			
Between-Subjects Effects	Language background	1246.433	2	623.217	8.366	.001	.383
	Error	2011.250	27	74.491			

Table 23. Multiple comparisons of groups for tokens of Vds produced under academic and daily topics using Turkey HSD (α = .05)

(I) Language background	(J) Language background	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
CFL	CHL	2.95	2.72	.534	-3.81	9.71
	CNS	-7.85*	2.72	.021	-14.61	-1.08
CHL	CFL	-2.95	2.72	.534	-9.71	3.81
	CNS	-10.80*	2.72	.001	-17.56	-4.03
CNS	CFL	7.85*	2.72	.021	1.08	14.61
	CHL	10.8\0*	2.72	.001	4.03	17.56

Based on observed means. The error term is Mean Square(Error) = 6.086. *. The mean difference is significant at the .05 level

In summary, all four repeated ANOVA tests reveal that discussion topic does have a significant influence on language speakers' oral production of V-N collocations, whether in terms of type or in terms of token. In order to further analyze language speakers' usage and preference under the two types of topics, the following section will focus on qualitative findings from the three datasets.

5.2.2. Qualitative analysis on production of collocations in discussing two types of topics

    In this section, I will examine more closely a number of cases from the three datasets which will generate interesting observations regarding some common characteristics of participants' collocational usage across the three groups.

5.2.2.1. Observations of V-N collocation usage in daily topic conversation

    In this part, I will provide excerpts and examples from the three datasets to show a few interesting observations regarding participants' choice of V-N collocations in discussing common everyday topics.

    Observation 1: Overall speaking, CNSs produced more idiomatic Vms-N collocations in talking about common everyday topics and a total of 16 idiomatic Vms-N collocations appear in the CNS datasets. In comparison, learners' use of idiomatic Vms-N is quite restricted and there are only 3 cases of Vms-N collocations in the CFL and CHL datasets. And 2 of the 3 cases concern the use of idiomatic expression "开玩笑 (to tell a joke)" by one CFL learner.

- Idiomatic collocations used by CNSs
  1) 我在中国找工作时，**碰 (*v. to bump*)**了不少**钉子 (*n. nails*)**。（碰钉子: bump one's head against a nail; meet setbacks）
     I met with many setbacks when I was look for a job in China.
  2) 我学习钢琴和画画其实都是在**打(*v. to buy*)酱油(*n. soy sauce*)**，没有学到很多东西。(打酱油: *to buy some soy sauce; to be a bystander*)
     Regarding piano and painting, I was only a passer-by and did not learn many things.

106

3) 我那个时候有点**走**(***v. to walk***)**极端**(***n. extreme***)，觉得自己没有别的选择了。(走极端: go to extremes)

   I was going to some extremes at that time and felt that I did not have any other choices.

4) 我大学毕业的时候，我爸其实**托**(***v. to support from under***)**关系**(***n. relation***)帮我找了一份政府的工作。(托关系*: to use one's personal network*)

   When I graduated from college, my dad used his personal network and found a government job for me.

- Idiomatic collocations used by CFL learners

  1) 我的很多中国朋友喜欢**开**(***v. to open***)**玩笑**(**n. joke**)。(开玩笑: *to tell a joke*)

     Many of my Chinese friends liked to tell jokes.

  2) 我那时候也开始**谈**(***v. to chat***)**恋爱**(***n. love***)，所以就想留下来找工作。(谈恋爱: *to fall in love*)

     I was falling in love at that time, so I wanted to stay and look for a job.

Observation 2: CHL learners make use of more colloquial Vms in collocationsVms than CFL learners. As shown below, in the three utterances made by CHL learners, Vms "煮"(*to cook*)，"搞"(*to do*) and "惹"(*to provoke*) are all considered colloquial verbs which appear commonly in casual or dialectal spoken language. As an example, the verb "煮"(*to cook, to boil*) is used often in some southern dialects (Cantonese, Taiwanese, etc.) of China and it is related to the special way of cooking in those regions. To express similar meaning, CFL learners tend to choose light verbs (such as "做, to do") and form collocations such as "做菜" (*to do some cooking*) and "做研究" (*to do some research*).

- 我平常很喜欢自己**煮**(*v. to cook, to boil*)菜，会**煮**(*v. to cook, to boil*)各种各样的食物。

  I usually like to cook myself, and I can cook many different foods.

- 我爸爸以前在中国是**搞**(*v. to do*)科学研究的。

  My dad did some scientific research when he was in China.

- 我小的时候很淘气，常给爸爸妈妈**惹**(*v. to provoke; to stir up*)麻烦。

  When I was a child, I was quite naughty and caused many trouble for my parents.

Observation 3: In talking about personal hobbies and daily activities, learners tend to repeat some light verbs in forming collocations. In the following example, a CFL learner used 3 collocations with the verb "打(v. to play)" and 2 collocations with the verb "做(v. to do)" in describing his hobby of playing tennis.

- 我平常经常**打**(*v. to play*)**网球**(*n. tennis,*)和游泳。**打**(*v. to play*)**网球**(*n. tennis*)是个很好的运动，可以帮助(*v. to help*)我的健康(*n. health*)，也让我有好的精神。我最喜欢**打**(*v. to play*)**双打**(*n. double play*)，因为可以跟朋友一起玩。……我们一般周末早上去(*v. to go to*)学校的网球场(tennis court, n.)，先**做**(*v. to do*)一些**动作**(*n. moves*)，正手球，反手球什么的，再**做**(*v. to do*)几个**比赛**(*n. games*)。

  I often play tennis and go swimming. Tennis is a great sport that is good for my health and spirit. I like to play doubles, because you can play with your friends…… We usually go to tennis courts on weekend mornings. I will practice some moves first, forehand, backhand, etc, and then play a few games.

108

5.2.2.2. Observations of V-N collocation usage in academic topic discussion

In this section, I will provide excerpts from the participants' narration on academic topics to show how individuals choose to use V-N collocations in narrating abstract daily topics. And the following observations were made.

Observation 1: Overall speaking, both learners and native speakers produce a great number of Vds-N collocations but very few Vms-N collocation in expressing their opinions on the abstract topics. In the following two narrations made by one CHL learner and one CNS, all of the V-N collocations are Vds-N collocations except one Vms-N collocation used by the CNS.

- Mark (CHL learner)

  中国教育体制很单一。每个学生有不同的能力和技术，像图画里每一个不同的动物，但中国教育体制只依靠高考来**决定**(*v. to decide*)学生的**能力**(*n. ability*)。在这个情况下，每个学生不可成功。美国的教育体制**认识**(*v. to recognize*)了考试的**弱点**(weakness, n.)。在各种教育和工作申请过程中，美国学校和公司**使用**(*v. to use*)几个**标志**(signs, n.)来**评价**(*v. to evaluate*)**申请者**(*n. applicants*)，包括面试和活动。

  China has a unitary education system. Although every student has different ability and skill, like each animal in the picture, China's educational system relies solely on the entrance exam to judge students' ability. Under this situation, individual students cannot become successful. US educational system recognize the weakness of exams. In various education and job application process, US schools and companies use several sign to evaluate applicants, including interview and activities.

- Gang (CNSs)

109

其实我觉得更重要的问题是教育应该怎样**培养**(*v. to cultivate*)人的各方面**能力**(*n. abilities*)，我自己认为中国的教育太**强调**(*v. to emphasize*) 某些方面的**能力**(abilities, n.)，如计算的能力，准确回答问题的能力等等。但因为学生**花**(to spend, v.)了太多**时间**(time, n.)在这些方面，就会**忽视**(to ignore, v.)其他的**能力**(abilities, n.)，或者没有足够时间和精力**发展**(to develop, v.)一些个人的**兴趣**(interests, n.)。

In fact I think a more important question is how to cultivate different abilities through education. I think China's education places too much emphasis on some abilities such as math and accurately answering questions. But since students spend too much time on these skills, they might overlook other abilities, or they do have enough time and energy to develop some personal interests.

Observation 2: Learners tend to employ a number of high-frequency Vds-N collocations in discussing abstract topics. In the first example below, the Vds-N collocation "面临问题(*face a problem*)" appears three times in a CHL learners' discussion of the environmental issue. In the second example, a CFL learners produced three different collocations using the same Vds "采取". In comparison, the third example below shows how one CNS uses three different collocations, "面临问题(*face a problem*)"，"应对困难(*reply to difficulty*)" and "克服挑战(*overcome challenge*)", to express the idea "China is facing big problems in environmental pollution".

- CHL learner

    中国许多大城市正在**面临**(*v. to face*)大量空气污染的**问题**(*n. problem*)，特别是在中国东岸与东北部**面临**(*v. to face*)更大**问题**(*n. problem*)。就算环

110

保者能够处理好"大气污染"的问题，他还是会需要**面临**(*v. to face*)更大的一个**问题**(*n. problem*)，水污染。

Many Chinese big cities are facing huge air pollution problem, especially the east coast and northeast China are facing bigger problems. Even if the environmentalists can deal with "air pollution", he still needs to face a bigger problem: water pollution.

- CFL learner

中国一直**采取**(*v. to adopt*)高考这个**办法**(*n. measure*)来评价学生的能力。我听说很多家长还有老师对高考都**采取**(*v. to adopt*)了很严肃的**态度**(*n. attitude*)，要准备很长时间，花很多很多时间和力气。对这个问题，我不知道政府**采取**(*v. to adopt*)什么**想法** (*n. attitude*)，但我认为政府已经开始改革高考的办法。

China has been adopting "college entrance exam" as a way to evaluate students' abilities. I heard many parents and teachers adopt very serious attitude towards the entrance exam, preparing for long periods of time and spending a lot of time and energy. I don't know what opinions the government adopts towards this problem, but I think the government is beginning to reform the entrance exam.

- CNS

中国现在**面临**(*v. to face*)了很大的污染**问题**(*n. problem*)，不仅是在空气方面，也在水和土壤方面。为了**应对**(*v. to deal with*)这些**困难**(n. *difficulties*)，政府已经开始采取一些具体的办法，比如制定更严格的法律和加强执法。我认为大概 10 年左右，中国很有希望能**战胜**(*v. to overcome*)这些**挑战**(*n. challenges*)。

China is facing big pollution problem, not only in terms of air, but also in terms of water and soil. To cope with these difficulties, government has started to

adopt some concrete measures, such as making more strict laws and
strengthening the enforcement of laws. I think in about 10 years, there is big hope
that China will overcome these challenges.

In summary, both quantitative analysis and qualitative observations reveal that participants use V-N collocations differently in discussing different topics.

5.3. Analysis of the non-conventional use of V-N collocations by CFL learners and CHL learners

This part provides both a quantitative and a qualitative analysis of the non-conventional V-N collocations identified in the spoken language datasets of CFL learners and CHL learners.

5.3.1 Descriptive statistics of the distribution of non-conventional V-N collocations

Altogether, 104 non-conventional uses of V-N collocations were found in the spoken language datasets of learners. Among them, 63 were produced by CFL learners and 41 were produced by CHL learners. Raw data of the number of non-conventional V-N collocations was shown in table 24. Considering the total number of V-N collocations generated by CFL learners (856 times) and CHL learners (844 times), the percentage of non-conventional collocations was quite small for both groups (7% for CFL learners and 5% for CHL learners) suggesting that advanced learners already achieved high level productive knowledge of V-N collocations.

Table 24. Number of non-conventional V-N collocations ut of collocations used by each participant?

CFL	Number of non-conventional V-N collocations	CHL	Number of non-conventional V-N collocations
Bill	5	Amber	2
David	5	Claire	5
Julia	9	Ethan	3
Kathy	6	Grace	1
Lucas	10	Kevin	4
Mary	6	Lily	8
Ryan	6	Mark	3
Tom	8	Mike	2
Tyler	4	Sarah	5
William	4	Sue	5
Total	63 (7% of total V-N collocations)	Total	41(5% of total V-N collocations)

In addition to the percentage of non-conventional collocations among all collocations, it would also be useful to study the percentage of non-conventional Vms and Vds collocations respectively, as shown in Figure 6 and Figure 7. The bar graphs showed that both groups of learners produced a higher percentage of non-conventional Vds-N collocations than Vms-N collocations. Such results indicate that for advanced learners, Vms-N collocations are easier to acquire than Vds-N collocations.
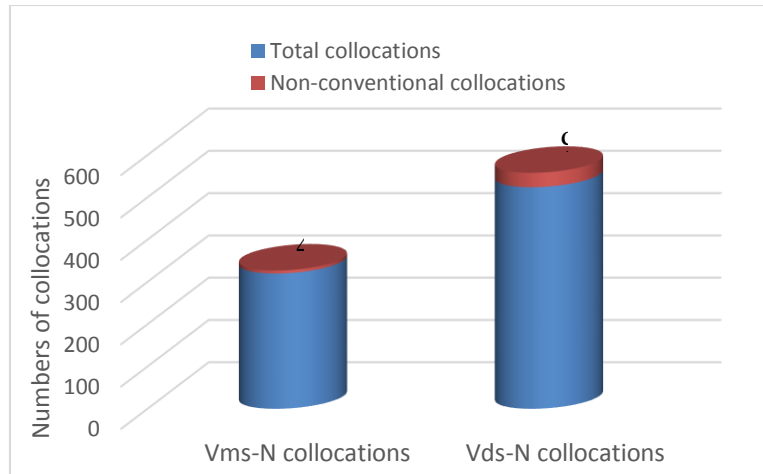
Figure 6. Percentage of non-conventional collocations in CFL total collocations
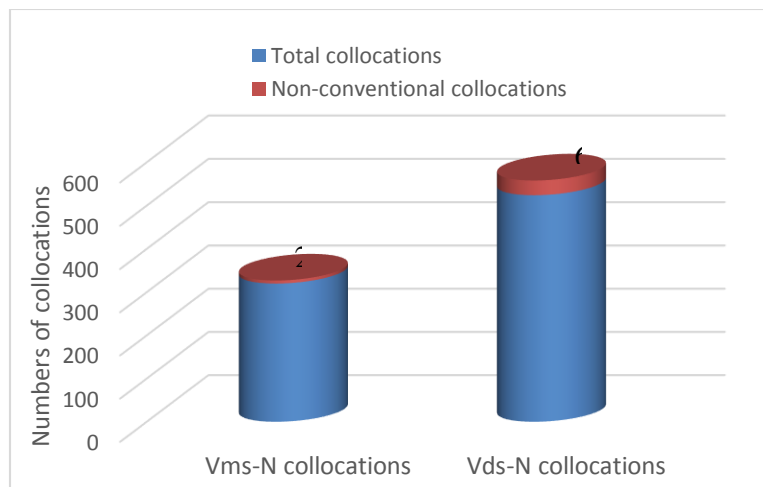
Figure 7. Percentage of non-conventional collocations in CHL total collocations

A close look at learners' unconventional use of collocations across the two different topics also revealed interesting findings. The percentage of non-conventional Vms and Vds collocations uttered in discussing two types of topics across the two learner groups were displayed in figure 8 and figure 9.

114

The pie charts indicated that for both CFL and CHL learners, more non-conventional V-N collocations were created in discussing academic topics. Furthermore, the use of non-conventional Vds-N collocations for academic topics accounted for the largest part of all non-conventional V-N collocation usage. These findings demonstrated that learning of academic Vds-N collocations posed biggest challenges for both CFL learners and CHL learners.
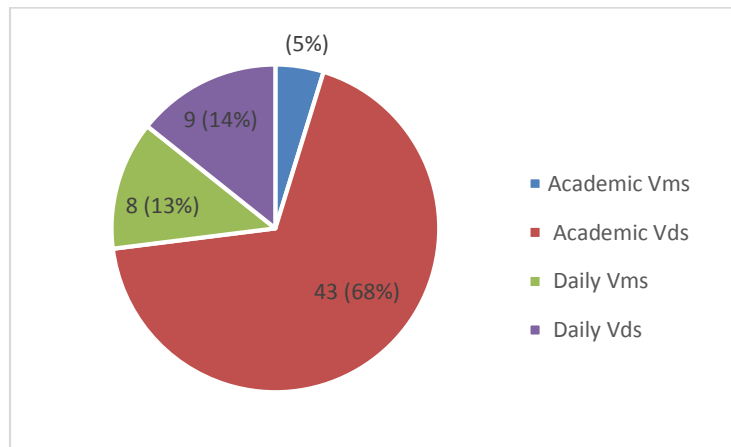
Figure 8. Distribution of non-conventional collocations by CFL learners
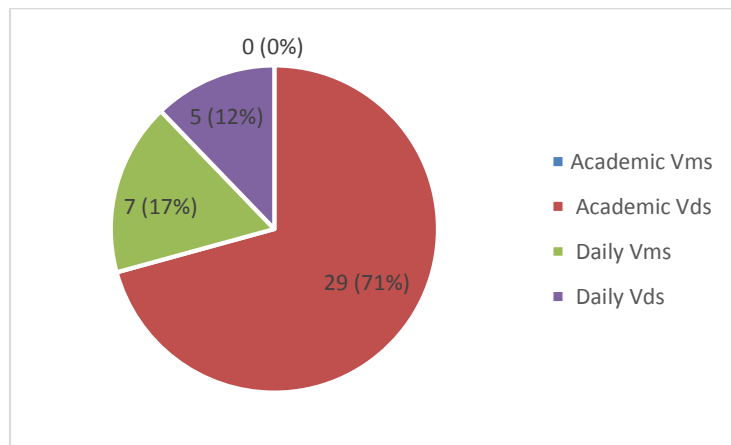
Figure 9 . Distribution of non-conventional collocations by CHL learners

5.3.2. Descriptive statistics of the acceptability scores for non-conventional collocations

After all the non-conventional V-N collocations were identified, these non-conventional usages are further evaluated by native expert raters on a five-point scale from completely unacceptable (1) unacceptable (2) not sure (3) acceptable (4) to completely acceptable (5).

Initially, two native expert raters were asked to judge and rate all the non-conventional V-N collocations used by the learners. If their ratings were the same, their judgement constituted the final score for the collocation. If their ratings had a difference of 1 point, an average score will be calculated to be the final score. If the two raters' scores had a difference of two or more points, one additional native speaker was asked to provide a judgement. After a third score was given for the collocation, an average score would be calculated using all three scores.

Regarding interrater reliability of the two main raters, Cohen's κ test was run to determine if there was agreement between two native speaker raters' judgement on the degree of acceptability of the 104 non-conventional V-N collocations produced by learners. Between the two raters, 73 identical scores, 19 scores with 1 point difference, 11 scores with 2 points difference and 3 scores with 3-4 points difference were identified. Based on the test results in Table 25, there was moderate agreement between the two officers' judgements, κ = .486, p < .0005.

Table 25. Interrater Reliability Test

Rater 1 score * Rater 2 score Crosstabulation

		Rater 2 score					Total
		1	**2**	**3**	**4**	**5**	
Rater 1	**1.0**	15	7	0	1	0	23
score	**2.0**	3	47	2	10	1	63
	3.0	0	4	5	2	0	11
	4.0	0	1	1	5	0	7
Total		18	59	8	18	1	104

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	**Kappa**	.486	.073	7.904	.000
N of Valid Cases		104			

The overall acceptability scores across the two learner groups were shown in Figure 10. Bar graphs indicated similar overall patterns of acceptability ratings between the two learner groups. Most of the non-conventional V-N collocations used by learners were considered "unacceptable" (rating scores 1-2) by native speakers. Furthermore, CFL groups produced more non-conventional collocations which were judged "completely unacceptable" (rating scores 0-1) by native speakers.

117
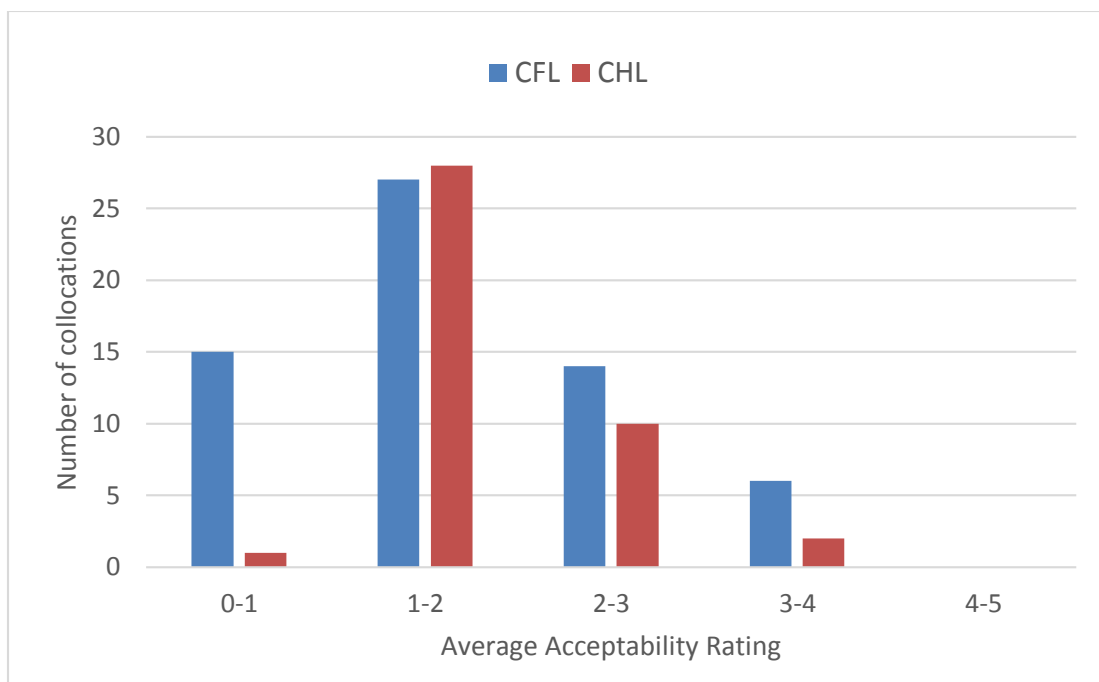
Figure 10. Acceptability rating of non-conventional V-N collocations used by CFL learners and CHL learners

5.3.3. *Qualitative analysis of non-conventional V-N collocations used by learners*

To find out about the linguistic features of non-conventional V-N collocations used by learners, native expert raters were also asked to provide a better option to replace the collocations that were judged as "not sure", "unacceptable" and 'completely unacceptable". Based on their responses and replacements, unconventional collocations could be categorized into three types: unconventional usage concerning the verbs, unconventional usage concerning the nouns, and unconventional usage concerning the entire collocations. Linguistic features of each type of collocations will be discussed below.

5.3.3.1. Unconventional usage concerning the verbs

Altogether 72 cases of unacceptable V-N usage concerning the verbs were identified by native speaker raters.

- Replacing Vds with Vds

The most common case of replacement made by native experts concerning verbs is replacing a Vds with another Vds of similar meaning. This case occurs 54 in all 20 learners across the CFL and CHL groups. Only a few learners used the same Vds inappropriately several times and no learner used the same in appropriate Vds more than three times.

Among the cases of Vds to Vds replacements, there are also different linguistic features concerning the Vds.

First, learners sometimes use a Vds which is similar both in form (Both Vds share one character in common) and in meaning with another more appropriate Vd. Three examples are given below to explain this feature. In example 1, a learner chose to use the verb "测量(*v. to measure*)" in the collocation "测量技能(*v. to measure the skills*)". Native expert replaces the verb "测量(*v. to measure*)" with another verb "测试(v. to test)". Although both verbs share the same character "测(v. to test)" and their meanings are similar too, the two verbs usually collocate with different nouns. "测量(*v. to measure*)"is often used with concrete noun objects such as "长度(*n. length*)" and "重量(*n. weight*), but "测试(*v. to test*)" is more often paired with abstract noun objects such as "能力(*n. ability*)" and "水平(*n.*

119

standard)". Thus native expert believes that it is more appropriate to use "测试技能 (*v. to test skills*)" than to use "测量技能(*v. to measure skills*)". In all of the examples that follow, learners' original usage of V-N will appear in bold characters, and the part which was considered non-conventional will be marked with an underline. The replacing collocation provided by native experts' correction will appear in the parenthesis that follows the original utterance, and the part which is being changed will be marked with an underline.

1) 考试是用来**测量**(*v. to measure*)人的**技能** (*n. skills*)。（测试[*v. to test*]技能/衡量[*v. to evaluate*] 技能, as replaced by native speakers）
   Exams are used to measure people's skills.

2) 高考**控制** (*v. to control*)了学生创新的**机会** (*n. opportunity*)。（限制[*v. to limit*]机会, as replaced by native speakers）
   College entrance exam limits students' opportunity for innovation.

3) 我们应该利用中国政府的控制能力来**转移**(*v. to transfer*)中国的经济**体制** (*n. system*)。（转变[*v. to transform*]体制, as replaced by native speakers）
   We should use Chinese government's control power to transform China's economic system.

Second, learners could choose a Vd is similar only in meaning with another more appropriate Vd. Some examples are:

1) 我希望通过上课来**提高**(*v. to raise*)我的**性格**(*n. personality*)。（改善[*v. improve*] 性格, as replaced by native speakers）
   I hope to improve my personality through taking classes.

120

2) 我 2012 年在上海**经历**(*v. to experience*)一个**暑假** (*n. summer break*)。

（度过[*v. to spend time*]暑假，as replaced by native speakers）

I spent a summer in Shanghai in 2012.

Third, other non-conventional usage of Vds-N collocations involves the use of intransitive Vds where transitive Vds is more appropriate. Examples for such case are listed below:

1) 中国依靠制造业来**增长**(*vi. to rise*)**国内生产总值**(*n. GDP*)。（提高[*vt. to raise*]国内生产总值, as replaced by native speakers）

China relies on manufacturing industry to increase its GDP.

2) **相比**(***vi. to compare with***)这两个国家的**政策**(***n. policy***)，我们发现… (对比[*vt. to compare*]政策/比较政策, as replaced by native speakers）

By comparing the policies of these two countries, we discover…

- Replacing Vms with Vms

Another type of replacement involves replacing a Vms in collocation with another more appropriate Vms. Such cases were less commonly found （total 11 cases）and there are also different linguistic feature concerning the Vms.

First, learners choose a general purpose Vms instead of a Vms with more specific meaning. Some examples include:

1) 你很快可以**做**(*v. to do*)**结论**(*n. conclusion*)。（下 v. [*to draw*]结论, as replaced by native speakers）

You will be able to draw a conclusion soon.

2) 每个人都需要**上**(*v. to attend*)**高考**(*n. college entrance exam*)。（考[*v. to test for*]高考, as replaced by native speakers）

Everyone needs to take the college entrance exam.

121

3) 中国春节的一个传统是**用**(to use, v.)**鞭炮**(fire crackers, n.)。（**放**[*v. to play*]鞭炮, as replaced by native speakers）

   One tradition of Chinese New Year is to play fire crackers.

Second, learners use a Vms that is similar in meaning with another more appropriate Vms.

1) 我们在会议上**给**(*v. to give*)了一个**报告**(*n. report*)。（做[v. to do]报告, as replaced by native speakers）

   We gave a report during the meeting.

2) 我**变**(*v. to change*)了一个**主意**(*n. idea*)。（改[*v. to change*]主意, as replaced by native speakers）

   I changed my idea.

3) 城市**盖**(*v. to build*)了很多新的**路**(*n. road*)。（修[*v. to build*]路, as replaced by native speakers）

   The city builds many new roads.

- Replacing Vms with Vds

   In addition to replacing Vds with another Vds and replacing Vms with another Vms, there were also a number of cases (9 cases) when a Vms was replaced by a better Vds and the following different linguistic situations are identified.

   First, some cases involve learners' choice of a Vms that is similar in both form (Both Vms share a character in common) and meaning with a more appropriate Vds. A few examples are shown below.

1) 我想**改**(*v. to change*)我的**看法**(*n. view*)。（改变[*v. to change*]看法, as replaced by native speakers）

   I want to change my view.

122

2) 教育制度不应该强迫所有的学生**追**(*v. to pursue*)一个教育**目标**(*n. goal*)。（<u>追求</u>[v. to pursue]目标, as replaced by native speakers）

   Educational system should not force all students to pursue the same educational goal.

Second, learners choose a Vms that is only similar in meaning with a more appropriate Vds. One example is:

1) 这就**丢**(to lose, v.)了一些多元文化的**优点**(*n. advantages*)。（<u>失去</u>[*v. to lose*]优点, as replaced by native speakers）

   This will cause (us) to lose some advantages of having a multicultural society.

5.3.3.2. Unconventional usage concerning the nouns

There are altogether 18 cases where native speakers replaced the nouns in the non-conventional V-N collocations. The relatively small number is probably due to the fact that native speaker raters were instructed to retain the intended meaning of the learner who made the utterance and nouns are often central to the meaning of a sentence.

The majority of unacceptable collocations concerning nouns are related to confusion of nouns with similar forms and/or meanings such as "相信"(*n. belief*) vs "信任"(*n. trust*)，"大战"(*n. war*) vs "战争"(*n. war*)，"斗争"(*n. flight*) vs "竞争"(*n. competition*) and "希望" (*n. hope*) vs "愿望" (*n. wish*). Some example sentences are listed below.

1) 企业想**得到**(*v. to obtain*)消费者的**相信**(*n. belief*)。（得到<u>信任</u>[*n. trust*]，as replaced by native speakers）

Companies want to obtain consumers' trust.

2) 这幅画的目的是**探索**(*v. to explore*)中国环境问题的**深度**(*n. depth*)。（探索<u>程度</u>[*n. level*], as replaced by native speakers）

The aim of this picture is to explore the level of environment problem.

3) 很多人**采取**(*v. to adopt*)这样的**主意**(*n. idea*)。（采取<u>观点</u>[*n. opinion*]/采取<u>想法</u>[*n. view*]）

Many people have adopted such view.

Besides the choice between synonyms, there are also a few cases where native speaker raters felt a Vds should be collocated with disyllabic nouns instead of monosyllabic noun and below are two examples.

4) 政府应该**修改**(*v. to correct*)过去的**错**(*n. mistake*)。（修改<u>错误</u>[n. *mistake.*]，as replaced by native speakers）

The government should correct their past mistakes.

5) 经济危机的时候公司**解雇**(*v. to fire*)**人**(*n. people*)。（解雇<u>员工</u>[*n. employees*], as replaced by native speakers）

During economic crises, the company fired some employees.

5.3.3.3. Unconventional usage concerning the entire collocations

In a total of 8 cases, native speaker raters changed the entire collocations to better express the intended meaning of the learner. Among these cases, raters either changed the word order of the collocations or provided alternative expressions to convey the meaning. Some examples of these replacements are as follows:

1) 我每天都要<u>见面</u>(***v. to meet***)十几个客户(***n. customers***)。（<u>跟十几个客户见面</u>[*to meet with more than ten customers*], as replaced by native speakers）

   I need to meet with more than ten customers every day.

2) 我们<u>做</u>(***v. to do***)了很多传统的风俗习惯(***n. traditions***)。（<u>尝试了很多传统活动</u>[*tried many different traditional activities*], as replaced by native speakers）

   We tried many different traditional activities.

3) 我们三个人<u>分享</u>(**to share, v.**)<u>了账单</u>(**bill, n.**)。（<u>分开结账</u>[*pay separately*], as replaced by native speakers）

   The three of us shared the bill.

4) 污染<u>恶化</u>(***vi. to worsen***)了气候(***n. climate***)。（<u>使气候恶化</u>[*make the climate worsen*], as replaced by native speakers）

   Pollution will make the climate become worse.

5) 假如不改变我们对环境的态度，<u>人会造成</u>(***v. to cause***)越来越不舒服的<u>世界</u>(***n. world***)。（<u>世界会越来越不适合人类生存</u>[*the word will become inhabitable for human beings*], as replaced by native speakers）

   If we don't change our attitudes toward the environment, we will create a world that is inhabitable for human beings.

5.3.4. Findings regarding the communicativeness of non-conventional collocations

In addition to judging the acceptability of non-conventional collocations, the 2 native experts were also given a task to evaluate the communicativeness of non-conventional collocations produced by the learners on a scale of 3: (1) non-communicative; (2) not sure; (3) communicative.

Ratings of the two raters show that among the 104 non-conventional collocations, 68 collocations (65%) were judged by both native experts as "communicative", 11 (15%)were judged by both experts as "not sure", and only 4 collocations (3%) were judged by both experts as non-communicative. This result indicates that the majority of non-conventional collocations seem not to hinder the meaning which is being conveyed by the advanced learners.

The 4 collocations judged as non-communicative by both experts are listed below. The non-communicativeness of these collocations seems to be caused by the vague or inappropriate semantic meanings of the collocation. For instance, in the first example, the collocation "刺激教育环境 (*v. to stimulate education environment*)" sounds rather vague in that the listener cannot grasp the meaning of "刺激(v. to stimulate)". The verb could indicate that government should invest more in education or it could mean the government should change some policies. In the second example that follows, "配合政府干涉(*v. to work with government interference*)" entails a somewhat contradictory meaning. The verb "配合(*v. to cooperate, to work with*)" generally has a positive connotation to indicate cooperative attitude. However, the noun "干涉 (*n. interference*)" entails a negative meaning of forceful interference. Thus when these two words are combined, it creates an inappropriate meaning. If we change the noun "干涉 (*n. interference*)" to "行动 (*v. to action*)", the collocation could be more communicative.

- 政府应该**刺激**(***v. to stimulate***)好的**教育环境**(***n. education environment***)来支持高考的改革。

126

The government should stimulate good education environment to support the reform of college entrance exam.

- 为了改善环保的情况，公民个人的改变也需要**配合**(*v. to cooperate, to work with*)政府的**干涉**(*n. interference*)。
  In order to improve the environmental protection situation, the changes of the citizens should work with government's interference.

- 我认为人民有权利**欺负**(*v. to tease*)**环境**(*n. environment*)**,** 因为环境短期看不会报复。
  I think people have the right to tease the environment, because in short term the environment will not take revenge.

- 假如不改变我们对环境的态度，人会**造成**(*v. to cause*)越来越不舒服的**世界**(*n. world*)。
  If we don't change our attitudes toward the environment, we will create a more and more uncomfortable world.

It should be noted that the attempt made by the current study to evaluate the communicativeness of collocations is quite preliminary, further studies with clearer definition and scale of "communicativeness" should be carried out to measure the intelligibility and processing of these collocations for listeners or interlocutors. Nevertheless, the preliminary findings here strikingly reveal how most of the so called "nonconventional" or "deviant" uses of learners actually do not interfere with the overall understanding of the listeners.

127

5.4. Summary of research findings

Based on both quantitative analysis and qualitative analysis of the three researcher questions, the current study generates the following major research findings:

1. Both CFL learners and CHL learners produce significantly less V-N collocations, both in terms of number and in terms of type, than CNSs. Also CFL learners and CHL learners tend to employ similar high-frequency V-N collocations in their spoken language, but CHL learners seem to produce more Vms-N collocations than CFL learners.

2. In discussing the two types of topics, daily topics and academic topics, all three groups prefer to use more Vms-N collocations in discussing daily topics and to use more Vds-N collocations in discussing academic topics. Again, CFL learners and CHL learners share more similarities in making frequent uses of some common V-N collocations whereas CNSs seem to apply a wider varieties of collocations in their language. In terms of the difference between the two learner groups, while CHL learners tend to use some verbs with more colloquial meanings in their collocations, CFL learners prefer to use more "standard" verbs.

3. Regarding the types and characteristics of the non-conventional V-N collocations produced by learners, learners in general produced a relatively small number of non-conventional uses. Most non-conventional uses seem to be related with confusion between verb synonyms or noun synonyms in the collocation. And as judged by native speaker experts, the majority of the non-conventional collocations are communicative and do not affect the message that the learners intend to convey.

CHAPTER 6: DISCUSSION

Chapter 6 discusses the findings of the current study with regard to the three research questions and the previous literature.

6.1. Discussion of Research Question 1

Research question 1 asks whether there is any difference in using Chinese V-N collocations across the three groups: Chinese as a foreign language learners (CFL learners), Chinese as a heritage language learners (CHL learners), and native Chinese speakers (CNSs). The discussion of the findings is divided into two parts: the quantitative findings and the qualitative findings.

6.1.1. Quantitative findings and discussion

The statistical analysis for research question 1 finds a significant difference in using V-N collocations across the three groups. Further analysis also shows that CNSs produced significantly more V-N collocations, in terms of both number (token) and range (type), than advanced CFL learners and CHL learners with only one exception: tokens of Vms by CHL learners (see the discussion below). This result is in line with the findings of many previous studies that language learners tend to use fewer and a narrower range of collocations (Fan, 2009; Nesselhauf, 2003; Siyanova & Schmitt, 2008).

Several reasons have been proposed regarding the finding that even "advanced" L2 learners can have difficulty in developing receptive and productive collocational competence. First, learning collocations is considered a complicated and cumulative process that requires siginificant language input in different contexts as well as ample opportunities for consolidation through language use and repetition (Henriksen, 2013).

Second, as discussed in Chapter 2, some researchers have claimed that L2 learners tend to focus on individual words rather than on recurring chunks in the input (Barfield, 2009; Wray, 2002). Third, many collocations used in a natural language context are not salient and do not cause comprehension problems. Therefore, learners may not notice the relationship between different constituents of the collocation (Warren, 2005). And finally, in terms of the language learning environment for L2 learners, classroom language instruction tends to focus on teaching individual words but does not provide effective materials for raising learners' awareness of collocations (Koya, 2005).

All of the above factors may also contribute to the finding of the current study; and among these factors, the very different condition of exposure to the Chinese language between learners and native speakers seems to be the most important. Similar to other L1 speakers, a native Chinese-speaking child in China begins to acquire the language through exposure within the family setting. After this child starts formal schooling, he/she will spend a large portion of both the school day and afterschool hours in reading and recognizing Chinese characters from various sources (textbooks, literary works, traditional poems, newspapers, and magazines, etc.) and in copying and writing thousands of Chinese characters. A recent report by the Chinese Ministry of Education has shown that Chinese elementary school students on average spend 2-3 hours each day on homework that consists mostly of writing Chinese characters and doing math problems. By the end of elementary school (age 12), most native Chinese students have acquired a total of 2,500-3,500 Chinese characters and are able to read literary works and newspapers with few problems. They will then go on to secondary schools and further

develop their reading and writing skills with all types of Chinese texts. During this process, most native Chinese speakers attain high levels of literacy and thus can make use of a wealth of sophisticated vocabulary and expressions, including collocations, in both their spoken and written language.

For most CFL learners, especially those who have alphabetical languages as their L1, learning Chinese is quite a different process. Compared with Indo-European languages such as English, Chinese has minimal morphological changes; but it has a rich tonal system and a sophisticated writing system (Zhang, 2004). Many earlier researchers commented on the difficulty in learning the Chinese orthographical system in relation to the spoken language. Everson (1988) noted that:

> One of the more challenging aspects of learning to read in a foreign language is the adjustment the learner must make in dealing with a different orthography. A significant aspect of orthography is that different writing systems have different script-speech relationships, and thus the acquisition of reading skills may in fact be hindered by how the spoken language is represented in print. Languages such as French and German do not present significant problems for American learners of these languages. For American learners of Chinese, however, the dissimilarity of the character set from English is so striking as to suggest potential problems for both the learning and teaching of this language. (p. 1)

According to the Defense Language Institute (DLI) and the Foreign Service Institute (FSI), which offer the widest range of foreign language courses in the U.S., languages such as Spanish and French have been categorized as "Group I" languages, and languages such as Chinese and Japanese have been categorized as "Group IV" languages. With regard to the time needed to achieve a specific level of language proficiency, the

FSI estimates that it takes L1 English-speaking American students approximately 480 contact hours of classroom instruction to reach Level 2 ("limited working proficiency") for Group I languages. In contrast, it takes students approximately 1,320 hours of the same type of instruction to reach a comparable level of proficiency in Group IV languages (Everson, 1994). What makes Chinese such a difficult language is its rather complex writing system, which adds a third dimension to the teaching and learning of Chinese (Guder, 2005). Therefore, it is easy to imagine that "the return to the learner for the hundreds of hours spent writing characters has a smaller payoff in terms of functioning as a participant in a Chinese society than the work he/she puts into any other of the skill areas" (Walker, 1989, p. 65).

Based on the U.S. context and my own teaching experience, most CFL learners begin learning Chinese in secondary schools or in college; and for the first several years, they tend to rely heavily on their Chinese textbooks, which contain mostly edited dialogues and shortened texts. Thus their focus of learning is usually on improving pronunciation/intonation, recognizing and writing basic characters, and practicing basic sentence patterns. CFL learners who persevere through this process and progress to the advanced level usually gain a solid foundation in pronunciation, common character reading, and grammar knowledge. In terms of language learning goals, many advanced learners tend to put more emphasis on their communicative skills (e.g., listening and reading skills) than on their literacy skills (e.g., reading and writing skills). As a result, for most advanced Chinese learners, their main source of input comprises spoken Chinese interaction both inside and outside of the classroom, Chinese textbooks, shortened and

132

often simplified Chinese news articles, adapted or even translated Chinese literary works, and some popular media sources (such as TV shows, popular music, etc.). Thus, these learners often do not have exposure to the wider variety of literary and sophisticated expressions (including collocations) found in traditional literary works, academic or scholarly papers, or professional-level seminars.

The language input condition for CHL learners differs somewhat from that of CFL learners, particularly in relation to their early exposure to Chinese during childhood. Based on the findings for research question 1, despite the general trend that CNSs use more collocations than learners, the only exception exists with CHL learners' use of Vms-N collocations. The statistical analysis fails to find any significant difference in the number of Vms-N collocations used by CHL learners and CNSs. Considering that Vms are commonly used in discussing daily topics, this finding suggests that CHL learners may have been exposed to many Vms and their different combinations in a childhood language context; and thus they can make greater use of such Vms-collocations in their spoken language. However, considering that the types of Vms-N employed by CHL learners are still significantly less than CNSs, the finding also shows that CHL learners may still not have sufficient exposure to the wide range of Vms-N collocations used in different Chinese contexts (e.g., academic and professional contexts). This is probably due to the fact that many CHL learners do not go through intensive formal schooling in Chinese; and thus their Chinese language input relies heavily on spoken language interactions on an interpersonal level.

In addition to the general finding of learners' production of V-N collocations, Figures 1 and 2 from Chapter 5 (section 5.1.1.1.) also reveal an interesting comparison between participants' use of Vms and Vds. In terms of Vms, the difference among the three groups in producing total numbers of Vms-N collocations is proportionate with their difference in producing different types of Vms-N collocations, showing that the frequency of the same types of Vms-N collocations is similar across the three groups. However, the trend is a little complicated for Vds-N collocation. Although both CFL learners and CHL learners seem to use less than half of the total types of Vds-N collocations, the mean tokens of Vds-N collocations by both groups of learners are more than two-thirds of the mean tokens of the Vds-N collocations produced by CNSs. These results demonstrate that learners tend to repeat the same type of Vds-N collocations more often than CNSs do. This may be due to the fact that advanced learners have relatively more exposure to the variety of Vms and their different combinations that appear often in spoken and informal contexts, yet they tend to have less exposure to the wide range of Vds and collocations that occur frequently in written language and formal contexts. Thus, while many of them are aware of the importance of using more Vds in discussing different topics, especially academic topics, their repertoire of different types of Vds-N collocations seems limited so they choose to use the same types more frequently.

6.1.2. Qualitative findings and discussion

A qualitative examination of the different types of collocations in the three datasets provides more details about the varying use of V-N collocations across the three groups.

134

First, several commonly used verbs--such as "做, *to do, to make*", "看, *to see*", "学, *to learn*", "了解, *to understand*" and "参加, *to participate*" --appear frequently in all three datasets. Just as in many other languages, general purpose verbs, especially those associated with commonly expressed topics, tend to be employed regularly by both native speakers and language learners.

Second, the range of Vms and Vds used in V-N collocations by CFL learners and CHL learners is less diverse than that of CNSs. For example, the total tokens of the five most frequently used Vms by CFL learners account for 34% of the total Vms produced by CFL learners in V-N collocations, compared to 29% for CHL learners, and only 10% for CNSs. This observation lends more support to the statistical finding of research question 1 and can be explained by the often restricted input condition experienced by learners of Chinese.

Third, there is a high degree of overlap in terms of the most frequently appearing Vms and Vds collocations between the CFL and CHL datasets. For example, of the five most commonly appearing Vms (in Vms-N collocations) in the CFL and CHL datasets, four were the same ("看, *to see*", "去, *to go*", "学, *to study*", and "做, *to do*"). Similarly, among the five most frequent Vds used by CFL learners and CHL learners, three are the same ("学习, *to learn*", "了解, *to understand*", and "解决, *to solve*"). The high degree of overlap between the two learner datasets could be due to several factors. First, the current study assigns similar and specific topics to learners and thus limits the range of vocabulary and expressions used. Second, most CFL learners and CHL learners are enrolled in the same academic program, and have also taken at least a few years of formal

135

Chinese classes at the college level. Therefore, they may have similar exposures and inputs regarding the use of collocations in the classroom.

In addition to the difference in verb variability between learners and native speakers, it is also worth noticing that even when using the same verbs, learners form different V-N collocations than native speakers do. Table 26 compares native speakers' and advanced learners' choice of noun objects with six different verbs.

Table 26. Comparison of six verbs' collocations produced by native speakers and advanced learners

Verb	V-N collocations produced by CNSs	V-N collocations produced by CFL and CHL learners
发展(to develop)	经济(economy), 国家(country), 兴趣 (interest), 爱好(hobbies)	能力(ability), 政策(policy), 办法(methods)
参加(to participate)	活动(activity), 工作(work), 会议 (conference), 运动(sports), 比赛(game)	学校(school), 比赛(game), 活动(activity), 过程 (process)
了解 (to understand)	情况(situation), 困难(difficulty), 项目 (project)	能力(ability), 影响(*influence*), 污染(*pollution*)
做(to do)	研究(research), 报告(report), 功课 (schoolwork), 决定(decision), 项目 (project) , 饭(meal)	饭(meal), 作业(homework), 比赛(game), 项目 (project), 结论(conclusion)
换 (to change)	经理(*manager*), 角度 (angle), 钱 (money)	飞机票(plane ticket), 想法(idea), 学校(school)
打(to hit, to play)	电话(telephone), 酱油(soy sauce), 比赛 (game), 名气(reputation)	网球(tennis), 篮球(*basketball*), 电话 (*telephone*)

This Table illustrates that, with some verbs, learners and native speakers choose to use similar noun objects. For example, in forming collocations with the verb "参加, *to participate*"， both native speakers and advanced learners used nouns such as "活动, *activity*" and "比赛, *game*". And with the verb "做, *to do*", both native speakers and advanced learners produce collocations such as "做项目, *to do a project*" and "做作业, *to do homework*". This demonstrates that advanced learners are familiar with the

136

common usage and combinations of some general purpose verbs that appear often in their language input.

Despite such similarities in using V-N collocations, Table 26 also indicates that, with many other verbs, learners and native speakers prefer to use quite different noun objects. For example, native speakers pair Vms "换, *to change*" with nouns such as "经理, *manager*", "角度, *angle*" and "钱, *money*" whereas advanced learners prefer to use "飞机票, *plane ticket*", "想法, *ideas*" and "学校, *schools*" with "换, *to change*".

The various preferences shown in Table 26 could be explained in a number of ways. For one reason, advanced learners of Chinese generally have a solid understanding of the semantic meanings of major verbs and thus can create collocations based on the meanings of verbs. For example, based on the basic meaning of "发展, *to develop*", learners produce collocations such as "发展学生的能力, *to develop students' abilities*" and "发展一个新办法, *to develop a new method*". While the uses of "发展, *to develop*" in these two collocations are semantically plausible decisions, both seem to be low frequency V-N collocations that do not appear often in a Chinese dictionary or corpus. Although some scholars may argue that learners need to master the conventional use of collocations to identify with the target language community (Wray, 2002), more researchers begin to view learners' nonconventional usage of collocations as positive indications of learners' efforts in achieving communicative competence (Howarth, 1998). The current study also supports the idea that attempts made by advanced learners of

Chinese in forming V-N collocations are meaningful processes for effective

communication (see the discussion below regarding research questions 2 & 3).

A second reason for the different collocational usage lies in the fact that many

Chinese verbs have a variety of semantic meanings. For example, Vms "打, *to hit*" has

more than ten different semantic meanings (e.g., to call, to beat, to deal with, to play,

etc.). Learners are usually familiar with the most common meanings of these verbs and

thus tend to create collocations based on their more familiar or more meanings of the

verbs. Table 26 shows that learners tend to use the verb "打, *to hit*" to mean "to play

(ball)" and "to call" in collocations that appear frequently in beginning Chinese language

textbooks. Previous studies (Wei, 2009; Xin, 2014) have also shown that in using

Chinese verbs with many meanings, learners tend to focus only on one or two of the

meanings in forming collocations. This is not surprising, considering the often limited

classroom instruction provided to explain the various meanings of the same verb in

collocating with different noun objects.

For a third reason, compared with native speakers who are immersed in a wide

variety of language contexts, from social media to academic seminars, learners' exposure

to the Chinese language is generally limited to a number of contexts (text books,

supplementary news articles and popular media resources, in-class conversations and

discussions, casual conversations with family or friends speaking Chinese, etc.).

Therefore they are unlikely to produce V-N collocations that appear in an unfamiliar

context. For example, the collocation "打酱油, *to buy soy sauce*" is a popular web phrase

used in describing something that is not important or relevant. Three of the native

speakers in the current study used this expression in both daily conversation (describing

hobbies) and in discussing an academic topic (describing the role of the government).

However, most advanced learners mostly likely have never heard of this expression, let

alone use it in their own language. This observation can also be explained by the

underlying mechanism of language acquisition described in usage-based models.

According to these models, people learn a language by observing and absorbing its use in

actual communicative events (Tyler, 2010). Thus a lack of context for observing actual

language usage may affect learners' receptive and productive language competence in

certain areas.

In summary, both the quantitative and qualitative analyses of research question 1

found significant differences among CFL learners, CHL learners and CNSs in using

Vms-N and Vds-N collocations. These differences can be attributed to a number of

factors, the most important of which seems to be the various language exposure and input

conditions among the three groups.

6.2. Discussion of Research Question 2

Research question 2 seeks to understand the influence of topic and context on learners' use of V-N collocations. The statistical findings and qualitative results are presented below.

6.2.1. Discussion of statistical findings for research question 2

Through a number of mixed ANOVA tests, this study shows that topic has a significant influence on language speakers' oral production of V-N collocations, whether in terms of type or token. Two specific findings are discussed below.

Finding 1: Both learners and native speakers produced significantly more Vms-N collocations in discussing daily topics than in discussing academic topics. This finding is in line with the research on the characteristics of the Chinese Vms and Vds. Chinese Vms are found to be used more extensively in daily communication and colloquial contexts, whereas Chinese Vds are used more commonly in written and formal contexts (Duanmu, 1999; Wei, 2009). Scholars have studied the proportion of Vms and Vds in a variety of contexts and found that the use of Vds is strongly related to formal and written contexts (Zhang, 2015). For example, among the 3,264 verbs that appear in the constitution of China, 96.14% are disyllabic. In comparison, in the traditional stage play "Thunderstorm", 90.1% of the verbs used are monosyllabic and only 9.9% are disyllabic (Zhang, 1989).

In addition, the category of Vms also contains more physical action verbs (such as "跑, *to run*", "打, *to hit*", and "走, *to walk*" that can collocate with many different types of noun objects. For example the verb "跑, *to run*" has five different semantic meanings

140

and can collocate with seven types of noun objects (Wei, 2009). In the current study, the discussion of daily topics includes narration of personal experiences, hobbies, and daily activities that require the use of many physical action verbs. Therefore, it is quite natural for both learners and native speakers to choose to use more Vms-N collocations in speaking about such topics.

Finding 2: CFL learners and CHL learners produce a greater number of Vds-collocations in discussing academic topics than in discussing daily topics. However, the range of Vds-N collocations employed by both groups is similar across topics. Overall, all three groups of speakers used significantly more (token) Vds-N collocations in discussing academic topics than in discussing daily topics. This is not surprising, given that Vds are commonly associated with academic and formal language situations (Zhang, 1989).

However, a close examination of the types of Vds-N collocations produced by the three groups shows that only CNSs used significantly more types of Vds-N collocations in academic discussion than in daily conversation (Figure 6.3). In other words, although learners use a far greater number of Vds in talking about abstract topics than daily topics, they tend to repeat the same types of Vds-N collocations rather frequently. This lack of variability in learners' language also corroborates with previous findings with regard to learners' overuse of certain types of collocations that are more frequent or "safe" to use (Chen & Baker, 2010; Tsai, 2015). In the current study, all the learners have been receiving formal instruction in advanced Chinese courses that emphasize spoken Chinese in academic and professional contexts and have achieved at least an advanced high rating

141

in OPI testing that measures their oral language proficiency in professional contexts. But in terms of collocations, these highly advanced learners still need to develop a richer repertoire of Chinese Vms-N collocations and Vds-N collocations, especially for discussing abstract topics. As usage-based language learning emphasizes, language learning and production always occur in context, and learners' choice of utterance depends on their prior learning, among many other factors (Tyler, 2010). For the current study, advanced learners who did not have prior exposure to a wealth of Vds-N collocations through extensive reading or other engagement with formal Chinese contexts may face challenges in retrieving or creating various Vds-N collocations when discussing abstract topics.

6.2.2. Discussion of qualitative analysis for research question 2

Qualitative analysis of the different uses of V-N collocations in discussing the two types of topics has drawn the following findings and observations.

In discussing daily topics, CFL learners and CHL learners can make proficient use of a variety of Vms in describing their experiences, activities, and hobbies. Overall, the types of Vms-N collocations used by learners and native speakers are very similar. More specifically, several observations have been made regarding patterns of usage by different groups.

One interesting observation about CHL learners is that they sometimes choose to use more colloquial or dialectal verbs in their collocations. In the first two examples that follow, Vms "煮, *to cook*" and "搞, *to do*" are both considered colloquial verbs that appear commonly in casual and dialectal spoken language. In comparison, they do not

142

appear at all in the CFL learners' dataset. To express a similar meaning, CFL learners

tend to choose general purpose verbs and form collocations such as "做菜, *to cook*

dishes", "做研究, *to do research*", and "上大学, *to attend college*". Although most of

the colloquial verbs are monosyllabic, in the last example, Vds "欺负, *to bully, to tease*"

is also a vernacular verb that occurs often in interpersonal communications.

- 我平常很喜欢自己煮(*v. to cook*)菜，会煮(*v. to cook*)各种各样的食物。
  I usually like to cook myself, and I can cook many different foods.
- 我爸爸以前在中国是搞(*v. to do*)科学研究的。
  My dad did some scientific research when he was in China.
- 我小时候常欺负(*v. to tease, to bully*)我弟弟。
  When I was a child, I often teased my brother.

Due to early exposure to the Chinese language in a family context, CHL learners

are more likely to pick up these verbs by listening to or speaking with their families.

Another observation regarding learners' preference in using V-N collocations for

daily conversation is that both CFL learners and CHL learners tend to over-generalize the

use of some Chinese "light" verbs. In the first example blow, the verb "做, *to do*" is often

considered a "light" verb which means "*to do*". In this example, the first "做, to do" is

used correctly as in "做练习, *to do some practice*". But the second "做, *to do*" is used in

a non-conventional manner to form "做比赛, *to do a game*". A more appropriate and

specific verb to use would be "打, *to play*" as in "打比赛, *to play a game*".

143

- 我们先做(*v. to do*)一些练习(*n. exercise*)，然后做(v. *to do*)几个比赛(*n. games*)。

  We did some exercises first and then played several games.

The term "light verb" has been used to refer to verbs in constructions such as "to make an appointment" and "to have a rest" (Miyamoto, 2000). A major characteristic of these constructions is that their semantic meaning is determined not by the verb, but by the object noun or other complement. In English, a number of light verbs such as "make", "take", "do", "have", and "give" are often used by learners. In Chinese, a set of light verbs such as "做, *to do*", "打, *to play*, *to hit*" and "进行, *to carry out*" have also been observed and studied. (Lin, 2001; Zhu, 2005). As several previous studies have shown, L2 learners tend to overuse collocations with light verbs because these verbs are often acquired early and thus can be accessed easily by the learners (Jiang, 2009; Lorenz, 1999).

One final observation regarding V-N usage in daily topic conversation is that, compared with learners, native speakers seem more open about using Vds-N collocations in talking about casual topics. For example, in discussing personal hobbies, learners tend to choose more action Vms such as "打, *to hit*", "玩, *to play*", "学, *to learn*", and "做, *to do*". However, native speakers employ many more abstract Vds in describing their hobbies such as "熟悉, *to become familiar with*", "保持, *to maintain*", "关注, *to pay attention*", and "改善, *to improve*". This may be due to the fact that native speakers have a broader vocabulary and thus can make use of different verbs in describing their hobbies. Another possible explanation might be the perception of language learners in

144

using different verbs. Whens studying Chinese, learners generally begin with simple

daily conversations and then move on to more abstract topic as their proficiency level

increases. Along the way, due to limited exposure to more complex and diverse reading

and speaking contexts, these learners may associate the use of certain vocabulary and

phrases with more casual situations or more formal situations. One final possibility could

be a psychological distance between the participants and me as the researcher. Because

most learners are my students, they may feel more relaxed in talking to me individually;

whereas all native speaker participants are new to the study and thus may feel a little

pressured to speak more formally.

While discussing and narrating academic topics, the three groups also exhibit a

few interesting features.

On one hand, both learner groups employ many sophisticated and highly relevant

V-N collocations (mostly Vds-N collocations) in narrating abstract academic topics.

Their use of V-N collocations has greatly enriched the content and depth of their

narration. In the following example, a CFL learner employs seven different types of V-N

collocations in discussing how human beings should face the challenges of pollution. All

of the collocations are highly relevant and appropriate in terms of semantics, genre, and

register.

- Excerpt from the CHL learner dataset
  在我看来，人类确实已经造成 (*v. to cause*)了一定的环境恶化 (*n.
  deterioration*)，造成 (*v. to cause*)了一定的全球变暖 (*n. global warming*)。但
  我们还有希望。我们每一个国家每一个人民必须团结起来为了解决 (*v. to*

145

solve)这个巨大挑战(*n. challenge*)，开始过(*v. to spend*)更环保的生活(*n. life*)。

各个政府与有影响力的私立公司也必须继续投资 (*v. to invest*)新的可持续发

展技术 (*n. technology*) 。如果大家能够一起面临 (*v. to face*) 挑战 (*n.*

challenge)，我们就可以成功地避免 (*v. to avoid*)环境恶化的后果 (*n.*

outcome)。

In my opinion, human beings have indeed caused some environment

deterioration, caused global warming. But we still have hope. Every country and

every people should unite to solve this huge challenge, to lead a more

environmental-friendly life. Every government and private company should

continue to invest in renewable technologies. If we could face the challenge

together, we can avoid the outcome of environment pollution.

On the other hand, learners in general tend to employ a number of high-frequency

Vds-N collocations in discussing abstract topics. For example, in the first utterance made

by a CHL learner, the Vds-N collocation "面临问题, *face a problem*" appears three

times. In comparison, the second example below shows how one CNS uses different

collocations to express similar ideas. To express the meaning that "China is facing big

problems in environmental pollution", the CNS employs three different collocations "面

临问题, *face a problem*", "应对困难, *deal with difficulty*", and "克服挑战, *overcome the*

challenge". Such variation shows that learners tend to have a smaller repertoire of ready-

made V-N collocations for use in their spoken language.

- Excerpt from the CHL learner dataset
  中国许多大城市正在面临(*v. to face*)大量空气污染的问题(*n. problem*)，特别
  是在中国东岸与东北部面临(*v. to face*)更大问题(*n. problem*)。就算环保者能

146

够处理好"大气污染"的问题，他还是会需要<u>面临</u>(*v. to face*)更大的一个<u>问题</u>(*n. problem*)，水污染。

Many large Chinese cities are facing huge air pollution problems. In particular the east coast and northeast China are facing greater problems. Even if the environmentalists can deal with "air pollution", they still need to face a bigger problem: water pollution.

- Excerpt from the CNS dataset

中国现在<u>面临</u>(*v. to face*)了很大的污染<u>问题</u>(*n. problem*)，不仅是在空气方面，也在水和土壤方面。为了<u>应对</u>(*v. to deal with*)这些<u>困难</u>(*n. difficulties*)，政府已经开始采取一些具体的办法，比如制定更严格的法律和加强执法。我认为大概 10 年左右，中国很有希望能<u>战胜</u>(*v. to defeat*)这些<u>挑战</u>(*n. challenges*)。

China is facing major pollution problems, not only in terms of air, but also in terms of water and soil. To cope with these difficulties, the government has begun to adopt some concrete measures, such as enacting stricter laws and strengthening the enforcement of the laws. I think in about 10 years, there is big hope that China will overcome these challenges.

To further understand learners' collocational competence as demonstrated by the above examples, we must first review the complex and cumulative process for developing L2 collocational competence. Henriksen (2013) argued that there are at least five aspects in acquiring L2 collocations: (1) notice and recognize collocations in the input, (2), grasp the meaning of collocations and create form-meaning associations, (3) comprehend the restrictions for using collocations, (4) select appropriate collocations based on context, and (5) develop fluency and automaticity in using the collocations. Such a complicated process requires intensive and diverse input and output conditions

147

over long periods of time. Advanced learners, such as the CFL learners and CHL learners in the current study, have gone through many years of formal and informal language interaction and training in a variety of settings and thus have proven themselves to be competent language users who draw on a large inventory of ready-made expressions to support their high-level language production. While the focus of the current study is on specific linguistic features in spoken language, I have noticed that in terms of language content, especially with academic topics, language learners are fully competent in expressing abstract ideas, making in-depth analyses, and providing persuasive arguments in a professional manner. As for the role of collocations in developing language proficiency, it is important to note that although learners possess a relatively small repertoire of different collocations due to variations in language exposure, learners have made use of a number of different communication strategies (e.g., experimentation, transfer, analogy, and repetition) to convey their meanings more precisely and clearly. Therefore, as argued earlier, the repetitions and nonconventional usages of collocations by L2 Chinese learners should be treated as meaningful attempts in achieving communicative competence.

In summary, the quantitative findings and qualitative observations of research question 2 demonstrate that both CFL learners and CHL learners employ different V-N collocations in discussing various topics: more Vms occur in daily conversations and more Vds occur in academic discussions. Also, learners seem to possess a relatively small repertoire of V-N collocations, especially regarding Vds, for their productive usage.

6.3. Discussion of Research Question 3

Research question 3 examines the types and characteristics of non-conventional collocations produced by advanced Chinese learners.

Among the over 1,600 total collocations produced by both CHL learners and CFL learners, only a small percentage (6%) was found to be non-conventional usage. This finding differs from the few previous studies that claim about one-quarter of the collocations produced by learners are deviant uses (Nessalhauf, 2005). This is likely due to the fact that all learners in the current study are highly advanced learners who have had rich experiences in and exposure to the target language. Another possible reason is that the present study focuses on the most direct form of V-N collocations (simple verb + simple noun), which can be relatively easy for learners to comprehend and produce.

A close examination of the non-conventional V-N collocations reveals that language learners tend to have more difficulty with Vds-N collocations than with Vms-N collocations, which are often used to discuss common daily topics. Vms- N collocations are usually acquired early in the learning process and are also reinforced through extensive input and output practice. Vds-N collocations are generally associated with academic and formal contexts. Learners can encounter a variety of them only by reading and listening to sources such as news reports, magazine articles, academic papers, academic presentations, documentaries, etc. Due to the difficulty of understanding and reading Chinese texts, most learners of Chinese do not have the literacy skills to read extensively in Chinese, even at the advanced level. As a result, their ability to produce a

wide range of Vds-N collocations is often hindered by the insufficient input of formal and academic Chinese.

In addition, several other factors relate to learners' unconventional use of V-N collocations.

6.3.1. Chinese lexical and semantic rules

The most important factor is related to Chinese lexical and semantic rules. As noted by several Chinese L1 acquisition studies, Chinese-speaking children on average acquire a greater proportion of verbs than English-speaking children do (Tardif, 1996; Tardif, Gelman, & Xu, 1999; Tardif, Shatz & Naigles, 1997). Thus, researchers have argued that the Chinese language places greater emphasis on verbs by applying a greater number of specific verbs to describe different situations. Unlike English-speaking children, who rely on only a few light verbs and more specific nouns to make meaning, Chinese-speaking children, by the age of 16 months old, have already acquired many "heavy" verbs such as "抱, *to hug*", "背, *to carry on back*", "打, *to hit*", "要, *to want*", and "给, *to give*" (Tardif, 2006). And their repertoire of verbs continues to expand as they grow up. Thus for L2 learners of Chinese, the process of learning and applying the vast number of Chinese verbs can be a particularly daunting task. For example, in the first example below, a CFL learner collocates the verb "穿, *to wear*" with both clothes and a belt. However, in modern Chinese, while clothes such as shirts, pants and shoes commonly collocate with "穿, *to wear*"; other clothing items such as hats and gloves are used with another verb "戴, *to wear*" which also means "to wear". In the second example,

150

learners used the verb "盖, t*o build*" with noun objects such as "house", "road" and "bridge". Yet in Chinese, there are a number of verbs that mean "to build" including "修, *to build, to fix*", "建, *to build*" and "建造, *to build*". Each generally collocates only with a certain type of noun. "盖, *to build*", for instance, refers only to building houses but not building bridges or roads.

- 潜水的时候需要<u>穿</u>(*v. to wear*)一种特别的<u>衣服</u>(*n. clothes*)，<u>穿</u>(*v. to wear*)一个重的<u>腰带</u>(*n. belt*)，还要<u>穿</u>(*v. to wear*)潜水<u>帽子</u>(*n. hat*)
  When you dive, you need to wear a kind of special clothes, wear a heavy belt, and wear diving hat.
- 政府到处<u>盖</u>(*v. to build*)<u>房子</u>(*n. house*)，<u>盖</u>(*v. to build*)<u>路</u>(*n. road*)，<u>盖</u>(*v. to build*)<u>桥</u>(*n. bridge*)。
  The government builds houses, builds roads, and build bridges everywhere.

Coupled with verbs that are similar in semantic meanings, there are also a large number of synonyms that are identical in forms. For examples, the verbs "转移", "转变" and "转型" all contain the character "转, *to change*" and all share the meaning of "to transform". Due to the subtle connotation meaning among these verbs, they seem particularly likely to cause confusion in forming V-N collocations. In the current study, the following groups of verbs with similar semantic meanings and/or similar forms seem particularly difficult for many learners:

- 控制 (*to control*), 限制 (*to limit*)
- 增长 (*to increase, to grow*), 增加 (*to increase*), 提高 (*to raise*)
- 降低 (*to reduce*), 减少 (*to reduce, to decrease*), 下降 (*to decrease*)

151

- 评价 (*to evaluate*), 评论 (*to comment*), 衡量 (*to measure, to judge*), 测量 (*to measure*)

- 相比 (*to compare*), 对比 (*to contrast*)

- 转移 (*to transfer*), 转变 (*to transform*), 转型 (*to transform*)

- 改 (*to change*), 变 (*to change*), 改变 (*to change*)

In the above list, verbs such as "降低, *to reduce*" and "下降, *to decrease*" seem to be very fundamental for language users. Nevertheless, the similarity of these verbs in both form and meaning could make it difficult for even advanced learners to learn and apply, especially in spontaneous spoken language.

6.3.2. Influence from L1

Another factor that relates closely to learners' unconventional use of V-N collocations is the influence from the learners' L1. About one-third of the non-conventional collocations in the current study seem to be direct translations from learners' L1 (English). This influence was found to be particularly strong when learners tried to express a culturally specific meaning or some complex and abstract ideas, as shown in the following two examples. In the first example, learners tried to express the meaning of "share the bill" in Chinese but could not locate the commonly used expression. In fact, such an expression is not used widely in Chinese as this particular action does not normally occur in the Chinese culture. In the second example, the learner was expressing his opinion that "the educational system should aid learners in realizing their highest potential." This seems to be a very important view on the part of the, who made the effort to express himself clearly and in a sophisticated manner. Therefore, he

chose to produce the word-for-word translation of "to realize one's highest potential" in Chinese to make his meaning clear. However, the noun "潜力, *potential*" co-occurs more commonly with another verb, "发挥, *to exert*" and does not appear naturally with the verb "实现, *to realize*". The third example is similar in that the learner chose to translate an abstract concept, "to manage the diversity (of students)" into Chinese without understanding that the word "管理, *to manage*" often collocates with persons or companies, but not with abstract ideas.

- 我们在一家饭馆吃了晚饭，我们三个人分享(*v. to share*)了账单(*n. bills*)。
  We ate dinner at a restaurant and the three of us shared the bill.
- 制度应该认识学生们的多样性而且帮助各个学生实现(*v. to realize*)他们自己最高的潜力(*n. potential*)。
  The (educational) system should acknowledge the diversity of students and facilitate them in realizing their high potential.
- 我们不能用一个标准体制管理(*v. to manage*)这样的多元化(*n. diversity*)。
  We cannot use one standard system to manage such diversity.

Regarding the influence of L1 on the acquisition of collocations, a number of scholars have examined the issue from different perspectives. Generally speaking, collocations are often cross linguistic, because "a collocation in one language usually has a counterpart in another language except when culture-specific concepts are involved" (Yamashita et al., 2010, p. 649). For example, both Chinese and English have the identical collocation of pay the bill. But when it comes to the culturally specific concept of sharing the bill, Chinese does not have an equivalent translation. Thus, a collocation is considered congruent in L1 and L2 if there is word-for-word translation in the two

languages; whereas a collocation is considered incongruent if there is no direct translation between the two languages.

It has been well documented that L2 learners have more difficulty in processing and producing incongruent collocations than congruent collocations (Altenberg & Granger, 2001; Nesselhauf, 2003). This is because learning congruent collocations in L2 is facilitated by L2 learners' understanding of the L1 counterpart. However, with incongruent collocations, the learning process involves obtaining meaning from each component word and establishing long-term associations between different components of the collocation in the mental lexicon (Yamashita et al., 2010). For advanced learners, although the process of acquiring many incongruent collocations is usually quite lengthy and requires intensive input, scholars have found support that advanced L2 learners can be successful in developing direct links between different types of L2 collocations and learners' conceptual representations (Jiang, 2004; Kroll & Stewart, 1994).

The findings and examples from the current study provide further evidence that CFL learners and CHL learners who have had repeated exposure to collocations in different contexts can make ready use of both congruent collocations (e.g., "解决问题, *to solve a problem*") and incongruent collocations (e.g., "进行思考, *to carry out thinking*"). In terms of Vms-N collocations and Vds-N collocations, it also seems that there are more incongruent Vms-N collocations than incongruent Vds-N collocations. Chinese is traditionally a monosyllabic language, and many Vms carry strong cultural conceptions and connotations. In contrast, many Vds were introduced into the language as direct translations of Western concepts and ideas (Duanmu, 1999). If this is case, Vds-N

154

collocations should be easier to acquire for L2 Chinese learners than Vms-N collocations. However, as the findings for research question 2 show, the learners' main challenge in learning collocations lies in their knowledge of only a small range of Vds-N collocations. These results again seem to indicate the important roles of exposure and input in acquiring collocations: Even advanced language learners do not have sufficient exposure to different genres of texts in both spoken and written contexts.

6.3.3. Collocations and communicative competence

Despite the complicated lexical and semantic rules regarding Chinese V-N collocations, advanced learners are able to produce a large number of conventional and highly communicative collocations. Among the 104 non-conventional collocations, more than 60% were judged by both native speaker experts as communicative. This shows that advanced learners are successful in employing different strategies to communicate effectively. As Howarth (1998) pointed out, non-conventional usage of collocations by L2 learners should be viewed more positively as indications of learners' risk-taking behavior for better communication.

More importantly, learners' investment in learning collocations is not simply driven by the goal of sounding more native-like. Barfield (2009) interviewed four learners as they developed their collocational competence. These learners expressed their motivation to acquire more precision in their language use and to be able to discuss more complex issues by learning various collocations. Moreover, they also wanted to express their individual identity and take on different social roles by functioning as confident second language users in the target language.

155

Both the CFL learners and CHL learners in the current study constitute a highly motivated and dedicated group of learners. They have invested considerable amount of time and energy in learning and reinforcing their Chinese language skills. Thus it is important to consider their learning motivations to attain sophisticated and professional language proficiency in the Chinese language. Thus far, there has been some evidence regarding the correlation between collocational knowledge and overall language proficiency (Koya, 2005; Pei, 2008), and some researchers have found significant processing advantages for collocations and other formulaic sequences for both L1 and L2 speakers (Columbus, 2010; Siyanova & Schmitt, 2008). Therefore, with the increasing evidence of certain benefits of collocations for language processing and production, advanced learners should be made more aware of the relationship between learning collocations and developing their overall communicative competency.

6.4. Summary of Discussion

In chapter 6, I discuss the findings of the current study as related to previous literature. The significant difference in using V-N collocations among the three groups, the influence of spoken topic on speakers' production of colocations and learners' unconventional use of collocations have been analyzed from a number of different perspectives: (1) CFL learners and CHL learners' Chinese input conditions; (2) the different use of Vms and Vds in V-N collocations across different contexts; (3) the relationship between collocational knowledge and communicative competency. Overall, it seems that the development of collocational knowledge is a slow and complicated process that demands intensive exposure as well as ample opportunities for practice and reinforcement. For advanced learners of Chinese, despite their ability in employing diverse Vms-N and Vds-N collocations for discussion of a variety of topics, their repertoire of certain types of V-N collocations still needs to be expanded to facilitate higher-level language proficiency. More importantly, findings of the current study demonstrate the efforts and achievements made by advanced learners' to reach communicative competence in the target language through acquiring and applying a variety of collocations.

CHAPTER 7: IMPLICATION AND CONCLUSION

This last chapter proposes theoretical and pedagogical implications of the findings, addresses some limitations of the study, and suggests areas for future research.

7.1. Implications

The current study has been designed to explore difference among Chinese language learners with different backgrounds in producing V-N collocations and the factors that may contribute to such differences. Findings of the study hold a number of implications for SLA theory and language teaching.

7.1.1. Theoretical implications

Findings of this research have implications for second language acquisition of collocations. A number of researchers have claimed that, unlike first language acquisition of children that involves the storage and process of language chunks, adult second language learners tend to separate fixed phrases and expressions into single lexical items (Wray, 2002; Wray, 2008; Gyllstad, 2007).

However, the current study shows that both CFL learners and CHL learners produce a large number of native-like conventional V-N collocations in their spoken language. Considering the relatively short processing time allowed in spontaneous spoken conversations, it is only reasonable that learners do have a way of storing and retrieving combinations of words in their memory. Also, regarding the underlying mechanism of collocation production, Sinclair (1991) made an influential distinction between two modes of language production: the idiom principle (retrieval of chunks) and

158

the open-choice principle (creating new combinations of words in keeping with syntactic and semantic rules). It has been hypothesized that learners may rely more on the open-choice principle than the idiom principle in formulating collocations. Nevertheless, findings from a number of studies do not support such hypothesis. Learners have been found to use the idiom principle to a comparable extent with the native speakers (Weinert 1995; De Cock, Granger, Leech, & McEnery, 1998). In the current study, both CFL learners and CHL learners smoothly incorporated many appropriate and sophisticated collocations into their spoken language to express highly abstract ideas. And they do not seem to rely much on creating different combinations out of single words. Recent studies (Siyanova-Chanturia & Martinez, 2015) on the processing and production of collocations claim that although collocations tend to be represented more in the mental lexicon of native speakers than language learners, there is no clear evidence that collocations are stored and processed differently by native speakers than by language learners.

To date, frequencies and intensity of language exposure or "engagement" have been found to be major factors related to language speakers' collocational knowledge (Hoey, 2005; Schmitt, 2010). Usage-based models of language acquisition claim that the frequent co-occurrence of two words in linguistic input will lead to their becoming associated in long-term memory (Ellis, 2002). The current study shows that CHL learners seem to have an advantage over CFL learners in using a greater number and variety of Vms-N collocations, especially in discussing daily topics. Considering CHL learners' early and frequent exposure to Chinese in childhood, it is likely that CHL learners have established stronger links between the constituents of some frequently occurring V-N

collocations. CFL learners, on the other hand, seem to be able to produce more Vds-N collocations than CHL learners, especially in discussing abstract academic topics. This finding could also be explained by the fact that CFL learners generally received longer periods of academic and formal language training in Chinese and thus had more repeated exposure to and practice of the different types of collocations in formal contexts. Although a number of CHL learners did mention their experience attending community Sunday Chinese schools in elementary school, most CHL learners enrolled in the current study did not take formal Chinese classes or attend Chinese immersion programs for extended periods of time. Therefore, the links between the different constituents of Vds-N collocations seem weaker in the mental lexicon of CHL learners than CFL learners.

Overall speaking, the results found in the current study can be taken to support a number of usage-based (Goldberg, 2006; Bybee, 2006; Tomasello 2003) approaches to language acquisition, processing, and use.

7.1.2. Pedagogical implications

The results presented by the current study also have important pedagogical implications, especially regarding consciousness-raising pedagogy, explicit teaching of certain V-N collocations, and design of foreign language curriculum.

There are different explanations to why collocations become a major challenge for language learners. For one reason, collocations are largely transparent and thus they do not usually constitute comprehension problems. As a result, collocations are often neglected in the process of foreign language teaching and learning. Another reason regarding the difficulty in collocation lies in the fact that collocations are so pervasive in

160

language input that identifying what to learn and learning them effectively seems such a daunting task (Yang & Hendricks, 2004).

The current study found that both CFL learners and CHL learners tend to have some difficulty differentiating the meanings of synonyms, especially verbs, in forming collocations and they also seem to rely heavily on a number of commonly-used and general meaning verbs in making V-N collocations. Thus a first step in effectively teaching collocations should be to make learners more aware of the existence of a large number of diverse collocations. A number of scholars have advocated awareness-raising pedagogy in teaching collocations (Howarth, 1996; Hill, 2000; Nesselhauf, 2005). Studies have shown that learners may benefit from guided learning approaches that emphasize encouraging learners to attend to syntagmatic structures in the input, try different learning strategies in acquiring and producing collocations, and reflect on their own learning process (Ying & O'Neill, 2009).

Findings of the current study also points to the necessity of teaching certain types and rules of collocations explicitly to advanced language learners. A number of recent studies have suggested that explicit vocabulary activities, in which learners are provided with repeated opportunities to encounter and practice target collocations, seem to be effective in making initial form-meaning links in learners' mental lexicon (Peters, 2014; Webb, Newton, & Chang, 2013). Firstly, language learners need to be taught the subtle differences among many synonyms, especially synonyms that share the same character, in forming appropriate collocations. For example, it is important to reinforce the use of synonyms such as "下降, *to reduce*" vs. "降低, *to decrease*" and "控制, *to control*" vs. "

161

限制, to limit". Although learners may already have a good command of the meaning of these words, they need to practice using these words in various appropriate combinations. Secondly, language learners need to be taught explicitly the "rhythmic principle" of creating collocations in Chinese. As discussed in Chapter 2, the "rhythmic principle" stipulates that for a V-N collocation, if the verb is disyllabic, the noun object can only be disyllabic. Such principle could help learners in differentiate the use of some synonymous Vms and Vds such as "学, *to learn*" vs. "学习, *to learn*" and "找, *to look for*" vs. "寻找, *to look for*". In language classrooms, we tend to notice that learners often use these synonyms interchangeably, so it is important to remind learners that they should still follow certain rules in creating V-N collocations. Lastly, in addition to explicitly teaching certain non-salient V-N collocations, teachers should also focus on helping advanced learner build a large repertoire of more sophisticated and abstract vocabulary and expressions suitable for different language contexts. As shown by findings to research question 3, the V-N collocations produced by learners are considered mostly communicative by native speaker experts. Therefore, it does not seem urgent or essential to correct all the "unconventional" usage by learners. Rather, attention should be given to facilitate advanced learners' efforts to express their more abstract and critical ideas in a clear and concise manner.

The third pedagogical implication concerns curriculum design for learners of Chinese and other foreign languages. The participants in the current study included a combination of traditional foreign learners of Chinese and learners who had early exposure to Chinese as a heritage language. They represent the diverse learner population

162

in postsecondary Chinese programs in the United States. Therefore the findings in the current study also provide invaluable implications for Chinese language education, particularly in language curriculum development and classroom instruction.

With the presence of more CHL learners in post graduate level Chinese courses, their special characteristics and needs have been taken into consideration in designing language learning curriculum. More colleges and universities are now offering dual or separate track programs in which CHL learners have the chance to develop essential Chinese literacy skills while improving their overall language competency. Findings of the current study indicate that CHL learners have a good command of V-N collocations in daily topic conversation, but their production of V-N collocations in academic topic discussion could be further enhanced. A number of previous studies also found that HL learners tend to have highly developed language competence in informal and vernacular language varieties, acquired primarily through interpersonal interactions at home. However, they still need to expand their linguistic competence in formal and professional language varieties and registers (e.g. Campbell, 2000; Fairclough, 2001; Valdés, 1995).

Regarding CFL learners, those who have achieved advanced Chinese proficiency are generally highly motivated learners who want to achieve the ability of functioning successfully in a Chinese-speaking professional context. Barfield (2009) pointed out that many advanced learners had the goal of being able to express individual identity and to attain more precision and sophistication in their language use. In other words, learning complex linguistic constructions such as collocations enable learners to gain more freedom in linguistic choices and to function more confidently as a second language user

163

(Wray, 2002; Barfield, 2009). Thus for this group of learners, language instruction should provide them more opportunities (through intensive reading, listening and other interactive activities) to be exposed to many different varieties of collocations or other formulaic constructions in academic and professional contexts.

In curriculum design, advanced Chinese program for both CFL learners and CHL learners should incorporate the goal of assisting HL learners to acquire the appropriate use of sophisticated, academic and professional vocabulary (including collocations and other formulaic sequences), registers as wells as genres.

One last pedagogical implication is the incorporation of the National Standards for Foreign Language Education into curriculum design. The National Standards for Foreign Language Education were first published in 1996 as a collaborative product of 10 foreign language associations. And the different standards are grouped around five themes: Communication, Cultures, Connections, Comparisons, and Communities. This document has been one of the most far-reaching and encompassing documents in the field of foreign language teaching and learning. The design and implementation of the current study is also based on the proficiency guidelines underlying the ACTFL OPI test which aims to evaluate how well students meet the National Standards. Although the national Standards are not a curriculum, but their specific organization can help language teachers and educators analyze our curriculum by looking closely at how we are addressing and implementing the Standards in our classes. And the result of analysis provides a clear picture of the areas that may be under-represented in the curriculum and leads to the development of a better balance in the future.

164

7.2. Limitations and Future Research

7.2.1. Limitations

Having discussed the potential contributions of the findings to the field of second language acquisition and heritage language instruction both theoretically and pedagogically, I will point out some limitations of the current study from the following two aspects: the generalizability of the current findings and the design.

First of all, the current study was carried out with a small number (20) of highly advanced learners of Chinese enrolled in the same graduate-level Chinese course. Also, data collected for the current study includes 1-hour long interviews between me and the learners which constitute relatively small datasets compared with larger corpus. Topics of the interview were also controlled to include only a number of daily conversation topics and abstract topics. Thus the findings and the implications may not be expanded to studies conducted in other settings, where learners may have a wider range of background and learning experiences. Another limitation that might weaken the generalizability of the current findings concerns the participants of the study. Due to the difficulties of recruiting adequate numbers of participants, the current study only included CHL learners who grew up in a Mandarin-speaking family. Learners who grew up in a home where a Chinese dialect other than Mandarin was spoken (e.g., Cantonese, Shanghainese, etc), were excluded from the current study. It is not clear whether CHL learners with different dialect backgrounds would have different collocational knowledge.

Secondly, in terms of the design, the current study only collected production data, but did not include any receptive knowledge data or processing data. A number of studies

have employed more complex research design to measure both receptive and productive knowledge (Siyanova & Schmitt, 2008; Webb et al., 2013) and their findings provided a more holistic view of learners' collocational competence. Also, the present study collected spoken language data at a single time which did not reflect the development of learners' collocational competence, so there needs to be long-term studies focusing on learners' progress in understanding and applying collocations over time.

7.2.2. Future research

In light of the findings and limitations of the current study, future research could be carried out to investigate theoretical and pedagogical aspects of learning collocations.

First of all, researchers need to carry out more fundamental studies regarding the language acquisition mechanism underlining the storage, processing and production of collocations. Moreover, specific models of collocation acquisition should be proposed in relation to general SLA theory and in relation to theories of vocabulary acquisition.

Second, future research should also address methodological problems involved in capturing and assessing internal learner processes of learning collocations. A variety of different methodologies from controlled laboratory studies to exploratory case studies should be carried out to examine the development of L2 collocational knowledge from different perspectives.

Third, learners' use of collocations needs to be examined further in a variety of different contexts. The relationship between collocational competence and overall language proficiency needs to be further investigated in terms of different languages. More research is also necessary regarding other types of collocations, other types of

166

language contexts (spoken vs. written, casual vs. professionals, academic vs. daily), and other types of language learners (e.g. bilingual speakers and multilingual speakers).

Last but not least, educators should work together to develop pedagogical tools for raising learners' awareness of the different types of collocations in various contexts, enhancing learners' ability to analyze and consolidate recurring patterns in the input, and developing a refined L2 collocational competence to achieve their individual goals with language learning.

7.3. Concluding Remarks

The present study explored the use of V-N collocations in discussing daily topics and academic topics by advanced CFL learners, advanced CHL learners and CNSs. The results showed that advanced CFL learners' and CHL learners' usage of V-N collocation is different from native speakers in that they produce less numbers and types of V-N collocations. Furthermore, learners' productive knowledge of collocations is also affected by different spoken topic: in discussing daily topics, advanced learners tend to use more Vms-N collocations whereas in discussing academic topics, advanced learners prefer to employ more Vds-N collocations. In addition, CFL learners and CHL learners exhibit both similarities and differences in using Vms-N collocations and Vds-N collocations under different spoken contexts, indicating their varying prior language learning experience and input conditions.

Results of the current study shed some light on how L2 learners with diverse language background incorporate sophisticated linguistic expressions, such as collocations, into their spoken language. More importantly, language researchers and

language educators can draw implications from the current study regarding (1) the underline mechanism of acquiring V-N collocations by Chinese L2 learners; (2) the relationship between collocational competence and communicative competence for advanced Chinese learners; and (3) the teaching strategies and materials of collocations for CFL learners and CHL learners.

Finally, due to the limitations in design and methodology, further research is needed to increase our understanding of the acquisition of collocational competence by CFL learners, CHL learners and other bilingual/multilingual learners of Chinese.

APPENDICES

Appendix A: Language learning questionnaire

Mandarin Chinese Learning Background Questionnaire

Please answer the following questions to the best of your knowledge.

1. Age: _____ ; Country of residence _____

2. If you have moved to the United States from a different country, what is your birth country and at what age did you move?

3. At what age did you start learning Chinese?

4. If you started learning Chinese as a child in your family, what dialect was spoken to you and how often?

5. How did you learn Chinese up to this point?
(Mainly Mostly Occasionally) through formal classroom instruction.
(Mainly Mostly Occasionally) through interacting with people.
A mixture of both, but (More classroom More interaction Equally both)

6. How many years of formal instruction have you received in learning Mandarin Chinese?

7. Have you lived or worked in a Chinese-speaking country? If so, for how long?

8. In the past year, how many hours do you spend in learning and using Chinese per week?
Speaking: _____(hrs)
Listening: _____(hrs)
Reading: _____(hrs)

Appendix B: List of interview questions

Part 1: Daily conversation topics

- 请简单介绍一下你自己。
  Could you introduce yourself?
- 你从小到大的教育背景是怎样的呢？
  Can you share with me your educational background?
- 你有怎样的工作经历呢？
  Can you tell me your work experience?
- 可以谈谈你学习英文（中文）的经历吗？
  How did you learn English/Chinese?
- 你为什么选择到宾大来学习？
  What did you come to study at Penn?
- 你在费城的生活是怎样的？每天你都做什么呢？
  What does your daily life look like? What do you do everyday?
- 你平常有什么爱好吗？可以向我具体谈谈你的爱好吗（方法，规则，装备等）？
  Do you have some hobbies? Can you describe to me your hobby in details?
- 可以谈谈你最近一次旅行的经历吗？
  Can you tell me about your recent travel experience?
- 可以谈谈你喜欢的电影或书吗？
  Can you tell me about a movie or a book that you really like?

Part 2: Academic topics

- 请介绍一个你从事过的研究项目，包括这个项目的背景，过程和意义等方面。
  Please describe a research project that you were involved in, including its background, procedure and significance.
- 请你看看这幅漫画，具体描述一下漫画的内容主题。
  Please read this cartoon carefully and describe the theme of this cartoon.
- 你对漫画的主题有怎样的观点呢？
  What is your opinion on the theme of this cartoon?
- 对于这个问题，你有任何的建议吗？
  Do you have any suggestions for this problem?

Appendix C: Cartoon picture prompts for academic topic discussion

漫画：我们的教育体制 (*Cartoon 1: our educational system*)

漫画：代价(*Cartoon 2: cost*)

BIBLIOGRAPHY

Au, T. K., Knightly, L. M., Jun, S., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science, 13*(3), 238-243.

Bahns, J., & Eldaw, M. (1993). Should we teach EFL students collocations? *System, 21*(1), 101-114.

Barfield, A. (2009). Following individuals' L2 collocation development over time. In A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 208-223). Basingstoke: Palgrave Macmillan.

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes, 26*(3), 263-286.

Brinton, D., Kagan, O., & Bauckus, S. (2008). *Heritage language education: A new field emerging*. New York: Routledge.

Bruhn de Garavito, J. (01). Verb raising in Spanish: A comparison of early and late bilinguals. *Proceedings of the 26th annual Boston university conference on language development* (pp. 84; 84) Cascadilla.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*, 711-733.

Campbell, R. (2000). Heritage language. In J. W. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 165-184). Mahwah, NJ: Lawrence Erlbaum.

Cao, P. (2008). Review of empirical studies on collocation in the field of SLA. *Chinese Journal of Applied Linguistics, 31*(6), 72-81.

Carreira, M. (2004). Seeking explanatory adequacy: A dual approach to understanding the term "Heritage language learner". *Heritage Language Journal, 2*(1), 1-25.

Cenoz, J., & Hornberger, N. H. (2008). *Knowledge about language* (2nd ed.). New York: Springer.

Chang, C., Haynes, E., Rhodes, R., & Yao, Y. (2008). A tale of two fricatives: Consosnant contrast in heritage speakers of mandarin. *University of Pennsylvania Working Papers in Linguistics, 15*, 37-43.

Chapelle, C., & Rosengarten Family Fund. (2013). *The encyclopedia of applied linguistics*. Chichester, West Sussex, UK: Wiley-Blackwell.

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14*(2), 30; 30.

Columbus, G. (2010). Processing MWUs: Are MWU subtypes psycholinguistically real? In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 194-212). London/New York: Continuum.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics, 29*(1), 72-89.

Cook, V. J. (1997). Monolingual bias in second language acquisition research. *Revista Canaria De Estudios Ingleses, 34*, 35-50.

Cook, V. J. (2009). Multilingual universal grammar as the norm. In I. Leung (Ed.), *Third Language Acquisition and universal grammar* (pp. 55-70). Bristol: Multilingual Matters.

Cowie, A. P. (1988). Stable and creative aspects of vocabulary use. In R. Carter, & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 126-139). London: Longman.

Cowie, A. P. (1994). Phraseology. In R. E. Asher (Ed.), *The encyclopedia of language and linguistics* (pp. 3168-3171). Oxford: Pergamon.

Cowie, A. P. (1998). *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press.

Cowie, A. P. (1981). The treatment of collocations and idioms in learners' dictionaries. *Applied Linguistics, 2*, 223.

Crossley, S., & Salsbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *International Review of Applied Linguistics in Language Teaching, IRAL, 49*(1), 1-26.

Dai, j. -. E., & Zhang, L. (2008). What are the CHL learners inheriting? Habitus of the CHL learners. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 37-51). Honolulu: University of Hawai'i Press.

David, L., & Ricardo, O. (2010). A multilevel statistical analysis of changes in language use among first-generation immigrants in a bilingual setting. *International Journal of the Sociology of Language, 203*, 83-100.

De Cock, S. (2007). Routinized building blocks in native speaker and learner speech: Clausal sequences in the spotlight. In M. C. Campoy, & M. J. Luzón (Eds.), *Spoken corpora in applied linguistics* (pp. 217-233). Bern: Peter Lang.

De Cock, S., Granger, S., Leech, G., & McEnery, T. (1998). *An automated approach to the phrasicon of EFL learners. Learner English on computer*. London: Longman.

DeKeyser, R. M. (1997). Beyond explicit rule learning. *Studies in Second Language Acquisition, 19*(02), 195-221.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In D. Long, & M. Long (Eds.), *The handbook of second language acquisition* (pp. 589-630). Oxford: Blackwell.

Doughty, C., Long, M. H., & Class of 1891 Department of Arts Fund. (2003). *The handbook of second language acquisition*. Malden, MA: Blackwell Pub.

Duanmu, S. (1999). Stress and the development of disyllabic words in Chinese. *Diachronica, 16*(1), 1.

Durrant, P., & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics in Language Teaching, 47*(2), 157-177.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163-188.

Ellis, N. C. (2003). Constructions, chunking and connectionism: The emergence of second language structure. In C. Doughty, & M. Long (Eds.), *The handbook of second language acquisition.* (pp. 63-103). Oxford: Blackwell.

Ellis, N. C. (1994). *Implicit and explicit learning of languages*. London/San Diego: Academic Press.

Ellis, N. C. (2002). Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies. *Studies in Second Language Acquisition, 19*(02), 145-171.

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly, 42*(3), 375-396.

174

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29-62.

Everson, M. (1988). Speed and comprehension in reading Chinese: Romanization vs. characters revisited. *Journal of the Chinese Language Teachers Association, 23*(2), 1-20.

Everson, M. (1994). Toward a process view of teaching reading in the second language Chinese curriculum. *Theory into Practice, 33*(1), 4-9.

Fairclough, M. (2001). *Language and power (2nd ed.).* (2nd ed. ed.). London: Longman.

Fan, M. (2009). An exploratory study of collocational use by ESL students – A task based approach. *System, 37*(1), 110-123.

Fillmore, L. W. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly, 6*(3), 323-346.

Firth, J. R. (1957). Modes of meaning *Papers in linguistics 1934-1951* (pp. 190-215). London: Oxford University Press.

Fishman, J. A. (2001). 300-plus years of heritage language education in the united states. In J. K. Peyton, D. A. Ranard & S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 81-98). McHenry, IL: Center for Applied Linguistics.

Fishman, J. A. (1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon/Philadelphia: Multilingual Matters.

Flores, C. M. M. (2015). Understanding heritage language acquisition. Some contributions from the research on heritage speakers of European Portuguese. *Lingua, 164, Part B*, 251-265.

Flores, C., & Barbosa, P. (2014). When reduced input leads to delayed acquisition: A study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism, 18*(3), 304-325.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gilquin, G. (2007). To err is not all: What corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift Für Anglistik Und Amerikanistik, 55*(3), 273-291.

Gilquin, G., Decock, S., & Granger, S. (Eds.). (2010). *LINDSEI: Louvain international database of spoken English interlanguage*. Louvain, Belgium: UCL Presses.

Gitsaki, C. (1996). *The development of ESL collocational knowledge.* Unpublished Ph.D thesis, University of Queensland,

Godson, L. (2004). Vowel production in the speech of western Armenian heritage speakers. *The Heritage Language Journal, 2*, 1-26.

Goldberg, A. (2006). *Constructions at work*. Oxford: Oxford University Press.

Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and application* (pp. 145-160). Oxford: Oxford University Press.

Granger, S. (2009). Learner corpora: A window onto the L2 phrasicon. In A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 60-65). Basingstoke: Palgrave Macmillan.

Guder, A. (2005, August). Struggling with Chinese: New dimensions in foreign language teaching. *International and Interdisciplinary Conference,* University of Mainz in Germersheim, Germany.

Guijarro-Fuentes, P., Pires, A., & Nediger, W. (2015). Delay in the acquisition of differential objectmarking by Spanish monolingual andbilingual teenagers. *International Journal of Bilingualism*, 1-19.

Gyllstad, H. (2007). Testing English collocations: Developing receptive tests for use with advanced Swedish learners. (Doctor, Lund University).

Gyllstad, H., & Wolter, B. (2015). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning.*

Håkansson, G. (1995). Syntax and morphology in language attrition: A study of five bilingual expatriate swedes. *International Journal of Applied Linguistics, 5*(2), 153-171.

Halliday, M. A. K. (1961). Categories of the theory of grammar. *Word, 17*(3), 241-292.

Halliday, M. A. K. (1966). Lexis as a linguistic level. In C. E. Bazell, C. Catford, M. A. K. Halliday & R. H. Robbins (Eds.), *In memory of J. R. firth* (pp. 148-162). London: Longmans.

Han, Y. (2010). *The effect of implicit and explicit feedback: A study on the acquisition of mandarin classifiers by Chinese heritage and non-heritage language learners.* Unpublished Doctoral dissertation, University of Florida,

He, A. W. (2008). Chinese as a heritage language: An introduction. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 1-12). Honolulu: University of Hawai'i Press.

He, A. W. (2004). Identity construction in Chinese heritage language classes. *Pragmatics: Quarterly Publication of the International Pragmatics Association, 14*(2-3), 199-216.

He, A. W. (2006). Toward an identity theory of the development of Chinese as a heritage language. *Heritage Language Journal, 4*(1), 1.

He, A. W., & Xiao, Y. (Eds.). (2008). *Chinese as a heritage language: Fostering rooted world citizenry*. Honolulu: University of Hawai'i Press.

Hendryx, J. D. (2008). The Chinese heritage language learners' existing linguistic knowledge and abilities. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 53-66). Honolulu: University of Hawai'i Press.

Henriksen, B. (2013). Research on L2 learners' collocational competence and development - a progress report In C. Bardel, B. Laufer & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 29-56) EUROSLA.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 49-60). London: Language Teaching Publications.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Hoffmann, S., & Lehmann, H. M. (2000). Collocational evidence from the British national corpus. In J. M. Kirk (Ed.), *Corpora galore: Analyses and techniques in describing English*. Amsterdam: Rodopi.

Hornberger, N. H. (2003). *Continua of biliteracy: An ecological framework for educational policy, research, and practice in multilingual settings*. Clevedon, UK; Buffalo: Multilingual Matters.

Hornberger, N. H., & Wang, S., C. (2008). Who are our heritage language learners? identity and biliteracy in heritage language education in the united states. In D. Brinton, O. Kagan & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 3-35). New York: Routledge.

Howarth, P. (1996). *Phraseology in English academic writing: Some implications for language learning and dictionary making.* Tübingen: Max Niemeyer.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*(1), 24-44.

Hulsen, M. (2000). Language loss and language processing. Three generations of Dutch migrants in New Zealand. (Doctoral dissertation, University of Nijmegen, The Netherlands).

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

James, W. (1890). *The principles of psychology*. New York: Holt.

Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing, 12*(4), 377-403.

Jia, G. (2008). Heritage language development, maintenance, and attrition among recent Chinese immigrants in New York City. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 189-203). Honolulu: University of Hawai'i Press.

Jia, L. (2009). Contrasting models in literacy practice among heritage language learners of mandarin. *Journal of Asian Pacific Communication, 19*, 56-75.

Jiang, X. (1998). Ci hui xi de yan jiu ji qi zai jiao xue shang de yi yi [research of vocabulary acquisition and implications for teaching]. *Yu Yan Jiao Xue Yu Yan Jiu, 3*, 65-73.

Jiang, W. (2009). *Acquisition of word order in Chinese as a foreign language*. Berlin ; New York: Mouton de Gruyter.

Johansson, S. (1979). American and British English grammar an elicitation experiment. *English Studies, 60*(2), 195.

Jones, S., & Sinclair, J. (1974). English lexical collocations: A study in computational linguistics. *Cahiers De Lexicologie, 24*(1), 15-61.

Jun, L. (2008). Ci yu da pei wen ti yu yu yan yan jiu (Chinese collocations and research). *Zhong Hua Wen Hua Lun Tan, 4*, 73-78.

Kagan, O., & Dillon, K. (2004). Heritage speakers' potential for high-level language proficiency. In H. Byrnes, & H. Maxim (Eds.), *Advanced foreign language learning: A challenge to college programs* (pp. 99-112). Boston: Heinle/Thomson.

Kagimoto, E., & Webb, S. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 32*(3), 259.

Kang, S., & Dong, M. (1998). Xian dai han yu wei bin dong ci fen lei tong ji yan jiu [research on the categories of Chinese verbs in verb-object construction]. *Liao Ning Shi Fan Da Xue Xue Bao, 1*, 15-25.

Kanno, K., Hasegawa, T., Ikeda, K., Ito, Y., & Long, M. (2008). Prior language-learning experience and variation in the linguistic profiles of advanced English-speaking learners of Japanese. In D. Brinton, O. Kagan & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 3-35). New York: Routledge.

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

Kazubski, P. (2000). *Selected aspects of lexicon, phraseology and style in the writing of polish advanced learners of English: A contrastive, corpus-based approach.* Unpublished PhD Thesis. Adam Mickiewicz University,

Ke, C. (1998). Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals, 31*(1), 91-102.

Kim, D. H. (2009). *A study on the use of lexical collocations of Korean heritage learners: Identifying the sources of errors.* Unpublished master thesis, University of Southern California,

Kim, J., Montrul, S., & Yoon, J. (2009). Binding interpretations of anaphors by Korean heritage speakers. *Language Acquisition, 16*(1), 3-35.

Kirk, J. M., & Class of 1924 Book Fund. (2000). *Corpora galore: Analyses and techniques in describing English: Papers from the nineteenth international conference on English language research on computerized corpora (ICAME 1998).* Amsterdam; Atlanta, GA: Rodopi.

Knightly, L. M., Jun, S., Oh, J. S., & Au, T. K. (2003). Production benefits of childhood overhearing. *The Journal of the Acoustical Society of America, 114*(1), 465-474.

Koda, K., Lv, C., & Zhang, Y. (2008). Effects of print input on morphological awareness among Chinese. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 125-135). Honolulu: University of Hawai'i Press.

Koya, T. (2005). *The acquisition of basic collocations by Japanese learners of English.* Unpublished Doctoral dissertation, Waseda University, Japan,

Leung, Y. I. (2009). *Third language acquisition and universal grammar*. Bristol, UK ; Buffalo, NY: Multilingual Matters.

Li, D. (2005, August). Attitudes, motivations and identities in learning Chinese as a heritage language. . *Paper Presented at the 14th World Congress of Applied Linguistics,* Madison, WI.

Li, D., & Duff, P. A. (2008). Issues in Chinese heritage language education and research at the postsecondary level. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 13-36). Honolulu: University of Hawai'i Press.

Li, F. (2003). *Han yu yu yi jie gou yan jiu [research on Chinese semantic constructions]*. Wuhan: Wuhan University Press.

Li, J., & Schmitt, N. (2010). The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 22-46). New York: Continuum.

Li, L. (1983). Bin yu shi yong qing kuang kao cha [research on the use of Chinese objects]. *Zhong Guo Yu Wen, 2*

Lin, T. (2001). Light verb syntax and the theory of phrase structure. (PhD Dissertation, University of California, Irvine).

Lorenz, G. R. (1999). *Adjective intensification-learners versus native speakers: A corpus study of argumentative writing* Rodopi.

Lu, X., & Li, G. (2008). Motivation and achievement in Chinese language learning: A comparative analysis. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 89-108). Honolulu: University of Hawai'i Press.

Lynch, A. (2008). The linguistic similarities of Spanish heritage and second language learners. *Foreign Language Annals, 41*(2), 252-381.

Meng, Z., Zhen, H., Meng, Q., & Cai, W. (Eds.). (1987). *Dongci yongfa cidian [A dictionary of verb usage]*. Shanghai: Shanghai Lexicographic Publishers.

Meunier, F., & Granger, S. (2008). *Phraseology in foreign language learning and teaching*. Amsterdam; Philadelphia: John Benjamins Pub. Co.

Miyamoto. (2000). *The light verb construction in Japanese. The role of the verbal noun* John Benjamins.

Montrul, S. (2012). Is the heritage language like a second language? *EUROSLA Yearbook, 12*, 1-29.

Montrul, S. (2010). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics, 31*(01), 167-207.

Montrul, S., & Bowles, M. A. (2008). Negative evidence in instructed heritage language acquisition: A preliminary study of differential object marking. In M. Bowles, R. Foote, S. Perpiñán & R. Bhatt (Eds.), *Proceedings of the 2007 second language research forum* (pp. 252-262). Somerville, MA: Cascadilla Proceedings Project.

Montrul, S., & Ionin, T. (2012). Dominant language transfer in Spanish heritage speakers and second language learners in the interpretation of definite articles. *The Modern Language Journal, 96*(1), 70-94.

Montrul, S. (2010). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics, 30*, 3-23.

Montrul, S., & Foote, R. (2014). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism, 18*(3), 274-303.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning, 58*(3), 503-553.

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

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223-242.

Oh, J. S., Jun, S., Knightly, L. M., & Au, T. K. (2003). Holding on to childhood language memory. *Cognition, 86*(3), B53-B64.

Paradis, M. (2004). *A neurolinguistics theory of bilingualism.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Paradis, M. (2009). *Declarative and procedural determinants of second languages.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

Partington, A. *Patterns and meanings: Using corpora for English language research and teaching.* Amsterdam: John Benjamins.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. C. Richards, & R. W. Schmidt (Eds.), *Language and communication* (pp. 191-226). London: Longman.

Pei, C. (2008). Review of empirical studies on collocations in the field of SLA. *Celea Journal, 31*(6), 72-81.

Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research, 18*(1), 75-94.

Peyton, J. K., Ranard, D. A., & McGinnis, S. (2001). *Heritage languages in America: Preserving a national resource*. McHenry, IL: Center for Applied Linguistics.

Pires, A., & Rothman, J. (2009). Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism, 13*(2), 211-238.

Polinsky, M. (2008). Russian gender under incomplete acquisition. *Heritage Language Journal, 6*, 40-71.

Polinsky, M. (2008). Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage Language Journal, 6*(1), 40-71.

Qi, C. (2005). Dui wai han yu jiao xue zhong de ci yu da pei yan jiu [research of collocations in teaching Chinese as a foreign language]. *Yun Nan Shi Fan Da Xue Xue Bao, 2*, 18-23.

Revier, R. L., & Henriksen, B. (2006). Teaching collocations. Pedagogical implications based on a cross-sectional study of Danish EFL.In M. Bendtsen, M. Björklund, C. Fant & L. Forsman (Eds.), *Språk, lärande och utbilding i sikte* (pp. 191-206)

Rinke, E., & Flores, C. (2014). Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingualism: Language and Cognition, 17*(04), 681-699.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced and instructed conditions. *Studies in Second Language Acquisition, 19*(02), 223-247.

Robinson, P. (2001). *Cognition and second language instruction*. Cambridge; New York: Cambridge University Press.

Rothman, J., & Treffers-Daller, J. (2014). A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics, 35*(1), 93-98.

Saadah, E. (2012). *The production of Arabic vowels by English L2 learners and heritage speakers of Arabic*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke, UK: Palgrave Press.

Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use*. Amsterdam; Philadelphia: John Benjamins Pub.

Schmitz, K. (2016). Differential object marking in Spanish speaking heritage speakers in Germany. In P. Guijarro-Fuentes, M. Juan-Garau & P. Larrañaga (Eds.), *Acquisition of romance languages: Old acquisition challenges and new explanations from a generative perspective* (). Berlin/New York: Mouton De Gruyter.

Serrano, R., Stengers, H., & Housen, A. (2015). Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research, 19*(1), 89-106.

Shao, j. (1995). Shuang yin jie V+N jie gou de pei jia fenxi [Analysis of the valency properties of Chinese disyllabic V+N constructions]. In Y. Shen, & D. Zhen (Eds.), *Xian dai han yu pei jia yu fa yan jiu [research on the valency grammar of modern Chinese]* (pp. 168-191). Beijing: Beijing University Press.

Shei, C. C. (1999). *A brief review of English verb-noun collocation.* Unpublished manuscript.

Shin, D., & Nation, P. (2008). Beyond single words: The most frequent collocations in spoken English. *ELT Journal, 62*(4), 339-348.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford ; New York: Oxford University Press.

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*(3), 429-458.

Siyanova-Chanturia, A., & Martinez, R. (2015). The idiom principle revisited. *Applied Linguistics, ,* 549-569.

Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27*(2), 251-272.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning, 63*(1), 121-159.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*(1), 1-33.

Stubbs, M. (2003). Two quantitative methods of studying phraseology in English. *International Journal of Corpus Linguistics, 7*(2), 215-244.

Taeko, K. (2005). The acquisition of basic collocations by Japanese learners of English. (Doctoral dissertation, Waseda University, Japan).

Tardif, T. (1996). Nouns Are Not Always Learned Before Verbs: Evidence from Mandarin Speakers' Early Vocabularies. *Developmental Psychology, 32*(3), 492-504.

Tardif, T., Gelman, S.A., & Xu, F. (1999). Putting the "Noun Bias" in Context: A Comparison of English and Mandarin. *Child Development, 70*(3), 620-635.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's' use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24*(1997), 535-565.

Tesnière, L. (1959). *Lucien, elements of structural syntax* (T. Osborne, S. Kahane Trans.). Amsterdam: John Benjamins.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Treffers-Daller, J., Daller, M., Furman, R., & Rothman, J. (2016). Ultimate attainment in the use of collocations among heritage speakers of Turkish in Germany and Turkish–German returnees. *Bilingualism: Language and Cognition, FirstView*, 1-16.

Tsai, K. (2015). Profiling the collocation use in ELT textbooks and learner writing. *Language Teaching Research, 19*(6), 723-740.

Tse, L. (1998). Ethnic identity formation and its implications for heritage language development. In S. Krashen, L. Tse & J. McQuillan (Eds.), *Heritage language development* (pp. 15-29). Culver City,CA: Language Education Associates.

Tse, Y. K. (2010). Parataxis and hypotaxis in modern Chinese. *International Journal of Arts & Sciences, 3*(6), 351-359.

Tsimpli, I. M. (2004). Features in L1 and L2 acquisition: Evidence from greek clitics and determiners. In H. Hendricks (Ed.), *Analyse comparative des processus d'acquisition*

en L1 et L2, special issue of AILE (acquisition et interaction en langue étrangère) (pp. 87-128)

Tyler, A. (2010). Usage-based approaches to language and their applications to second language learning. *Annual Review of Applied Linguistics, 30*, 270-291.

Valdés, G. (1995). The teaching of minority languages as academic subjects: Pedagogical and theoretical challenges. *Modern Language Journal, 79*, 299-328.

Valdés, G. (2001). Heritage language students: Profiles and possibilities. In J. K. Peyton, D. A. Ranard & S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 37-80). McHenry, IL: Center for Applied Linguistics.

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal, 89*(3), 410-426.

Waibel, B. (2008). *Phrasal verbs: German and Italian learners of English compared* . Saarbrücken, Germany: VDM.

Walker, G. (1989). The less commonly taught languages in the context of American pedagogy. In H. Lepke (Ed.), *Shaping the future: Challenges and opportunities* (pp. 111-137). Middlebury, VT: Northeast Conference.

Wang, S. C. (2010). Chinese language education in the United States: A historical overview and future directions. In J. Chen, C. Wang & J. Cai (Eds.), *Teaching and learning Chinese: Issues and perspectives* (pp. 3-32). Raleigh, NC: Information Age Publishing.

Wang, Y., & Shaw, P. (2008). Transfer and universality. *ICAME Journal, 32*, 201; 201.

Warren, B., & Erman, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning, 63*(1), 91-120.

Weger-Guntharp, H. (2006). Voices from the margin: Developing a profile of Chinese heritage language learners in the FL classroom. *Heritage Language Journal, 4*(1), 29-46.

Wei, H. (2009). *Mian xiang han yu xi de de chang yong dong ci dai bin yu qing kuang yan jiu [research on the common verb-object constructions in the acquisition of Chinese as a foreign language]*. Beijing: People Press.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics, 16*, 180-205.

185

Wen, X. (1997). Motivation and language learning with students of Chinese1. *Foreign Language Annals, 30*(2), 235-251.

Wiley, T. G. (2001). On defining heritage languages and their speakers. In J. K. Peyton, D. A. Ranard & S. McGinnis (Eds.), *Heritage languages in America: Preserving a national resource* (pp. 29-36). McHenry, IL: Center for Applied Linguistics.

Wiley, T. G. (2008). Chinese "dialect" speakers as heritage language learners: A case study. . In D. Brinton, O. Kagan & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 3-35). New York: Routledge.

Wood, D. (2010). *Perspectives on formulaic language: Acquisition and communication*. London; New York: Continuum.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge; New York: Cambridge University Press.

Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford; New York: Oxford University Press.

Wu, S. (2008). Robust learning for Chinese heritage learners: Motivation, linguistics and technology. In K. Kondo-Brown, & J. D. Brown (Eds.), *Teaching Chinese, Japanese, and Korean heritage language students: Curriculum needs, materials, and assessment* (pp. 271-297). New York: Lawrence Erlbaum Associates.

Wu, M. (2013). Re-imagining education for linguistically, culturally, and racially diverse students in a changing era: One U.S. urban school's alternative vision. (Ph.D., University of Pennsylvania). *ProQuest Dissertations and Theses.*

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

Xiao, Y. (2008). Home literacy environment in CHL development. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 151-166). Honolulu: University of Hawai'i Press.

Xin, P. (2014). *Mian xiang dui wai han yu jiao xue de chang yong dong ci V+N da pei yan jiu [research on common V+N collocations in teaching Chinese as a foreign language]*. Beijing: World Books Publishing.

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44*(4), 647-668.

Yang, Y., & Hendricks, A. (2004). Collocation awareness in the writing process. *RELT, 3*, 51-78.

Ying, Y., & O'Neill, M. (2009). Collocation learning through and 'AWARE' approach: Learner perspectives and learning process. In A. Barfield, & H. Gyllstad (Eds.), *Researching collocations in another language: Multiple interpretations* (pp. 181-193). Basingstoke: Palgrave Macmillan.

Yip, P. C., & Rimmington, D. (2004). *Chinese: A comprehensive grammar*. New York: Routledge.

Zhang, G. (1989). *Dan shuang yin jie dong ci gong neng cha yi yan jiu [functional difference between monosyllabic verbs and disyllabic verb].* Unpublished Master thesis, Shanghai Normal University,

Zhang, L. (2015). Xian dai han yu yin jie dan shuang yu kou yu / shu mian yu de guan lian yan jiu [research on the relationship between word syllabi and colloquial/written language]. *Qing Chun Sui Yue, 13*

Zhang, Y. (2004). *Xian dai han yu shou shi bin yu yan jiu [research on modern Chinese objects]*. Shanghai: Xuelin Publisher.

Zhou, X., Ye, Z., Cheung, H., & Chen, H. (2009). Processing the Chinese language: An introduction. *Language and Cognitive Processes, 24*(7-8), 929-946.

Zhu, X. F. (2005). Qing dong ci he han yu bu ji wu dong ci dai bin yu xian xiang [light verbs and the phenomenon of Chinese intransitive verbs bearing an object]. *Xian Dai Wai Yu, 28*(3), 221-231.