VARIATION IN THE BA CONSTRUCTION IN THE INTERLANGUAGE OF L2 CHINESE

Xiaoping Gao

ABSTRACT

Recent years have witnessed a growing interest in the variability of interlanguage in Second Language Acquisition. While a substantial body of research has focused upon the sources of this variation in relation to European languages, little has been done to explore variation in Chinese interlanguage, particularly in the use of Chinese syntactic structures. The current study fills this gap.

This study investigates the sources of intra- and inter-learner variation in the use of the Chinese *ba* construction (*BC*) by adult English and Korean native speakers. The learners participated in the study in both New Zealand (n = 20) and in China (n = 90), with 22 native speakers of Chinese providing baseline data. The entire learner sample consisted of 56 English and 54 Korean native speakers (56 females and 54 males) who had mixed proficiency in Chinese, ranging from late beginner to advanced levels. A battery of three tasks (i.e., an Oral Production Task prompted by Video clips [OPTV], an Oral Imitation Task [OIT], and an Untimed Grammaticality Judgement Task [UGJT] conducted orally) was used to elicit the target types of *BC*. These tasks afforded four measures - oral production, oral imitation, untimed grammaticality judgments and error correction. Learners' performance was assessed according to the accuracy of their use of *BC*. The effects of eight potential sources (i.e., linguistic difficulty, task type, starting age, number of years of study, setting, first language, self-rated proficiency, and gender) on the accuracy of use of *BC* were examined in order to explore the variability in *BC* in the learners' interlanguage.

The study adopted an innovative methodology to examine the variability of *BC* by considering not only a sociolinguistic factor (i.e., setting) and individual factors (e.g., starting age) but also a psycholinguistic factor (i.e., task). The statistical analyses demonstrated that both linguistic difficulty (i.e., *BC* type) and task type were significant sources of intra-learner variation in *BC*. The learners' accuracy scores for *BC1* – a (locative) nominal *BC* whose complement contains a NP were consistently higher than those for *BC2* – a directional verbal *BC* whose complement contains only a directional verb but no NP. The learners' *BC* scores in terms of the four measures followed the order: Judgments > Correction > Oral production > Oral imitation. Inter-learner variation was examined in terms of two measures: Oral scores (i.e., the average of the oral production and oral imitation scores) and metalinguistic scores (i.e., the average of the judgment and correction scores). The statistical analysis also showed that

the sources of the inter-learner variation were the number of years of study, setting, first language, and starting age. The number of years of study positively correlated with the oral and metalinguistic scores for both types of *BC*. The China group outperformed the New Zealand group in terms of both oral and metalinguistic scores. The Korean learners had higher metalinguistic scores than the English speaking learners, but there was no difference between them in oral scores. Late starters outperformed early starters but only in metalinguistic scores for *BC2*. In addition, self-rated proficiency significantly correlated with both oral and metalinguistic scores of *BC*, whereas gender had no effect on the accurate use of *BC*. This study suggests that multiple sources account for the variation in *BC*, and that the effects of these sources differ somewhat in terms of the learners' oral and metalinguistic competencies.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A	adverbial
AH	at home (group)
BA	the preposition, *ba*
Ba-NP	the nominal object of *ba*
BC	the *ba* construction
CFL	Chinese as a foreign language learners
CL	classifier
CN	China
CRS	currently relevant state (*-le*)
CSC	complex stative construction (*-de*)
CSL	Chinese as a second language learners
DE	particle (*-de*)
DUR	durative aspect (*-zhe*, *zai*)
EC	error correction
ESL	English as a second language learners
EXP	experiential aspect (*-guo*)
GJ	grammaticality judgement
GJT	grammaticality judgment test/task
P	preposition
PFV	perfective aspect (*-le*)
PP	prepositional phrase
IELTS	International English Language Testing System
LN	locative noun
L1	First language
L2	Second language
MKT	metalinguistic knowledge test/task
NEG	negation
NP	noun phrase
NZ	New Zealand
OIT	oral imitation task/test
ONT	oral narrative test
OPI	oral proficiency interview

OPTV	oral production task prompted by video clips
OSV	object-subject-verb word order
RC	resultative complement
RRC	restrictive relative clauses
RVC	resultative verb complement
SA	study abroad (group)
SLA	Second Language Acquisition
SOV	subject-object-verb word order
SFL	Spanish as a foreign language learners
SSL	Spanish as a second language learners
SVO	subject-verb-object word order
TCF/SL	Teaching Chinese as a Foreign/Second Language
TCSOL	Teaching Chinese to Speakers of Other Languages
TESOL	Teaching English to Speakers of Other Languages
TGJT	timed grammaticality judgment test/task
UGJT	untimed grammaticality judgment test/task
V_{di}	directional verb

CHAPTER 1: INTRODUCTION

1.1 Background

Chinese is the indigenous language originally spoken by the Han Chinese in China. Modern Standard Chinese is variously labelled. It is called Hanyu (lit. the language of the Han ethnicity) in the People's Republic of China, Huayu (lit. the language of the Hua ethnicity, Hua is an archaic term for Chinese) in Singapore, Guoyu (lit. the national language) in Taiwan, and Mandarin in western countries. The spoken form of Standard Mandarin Chinese is called Putonghua (lit. the common language) in P.R. China. The term Zhongwen (lit. Chinese language) is also widely used in educational systems inside and outside China.

According to a definition of *Xiandai Hanyu* 'Modern Chinese' [1](1993, p. 2), Modern Standard Chinese refers to the Beijing-dialect-based *lingua franca* used in modern times. Its vocabulary is largely drawn from the Northern Dialect (also called Mandarin dialects), a large and diverse group of Chinese dialects spoken across northern and south-western China (see Appendix 6 for a map of the area of Mandarin dialects). The grammatical norms are standardised to the body of modern literary works written in vernacular Chinese, which in practice follows the same tradition of the Mandarin dialects with some exceptions. Standard Mandarin is usually simply called Chinese. This thesis will follow this convention.

As a canonical word order structure in Chinese language, the *ba* construction (*BC*) has received most debate in Chinese linguistics (see, Y.-H. A. Li, 2001; Sun, 2006; D. Xu, 2006). Controversy over *BC* mainly concerns two issues. First, the existence of *BC* has led to a debate on whether the canonical word order of Chinese language is SVO or SOV (C. N. Li & Thompson, 1976; Sun, 1995; Sun & Givón, 1985; Tai, 1985). Second, the explanations of the function of *BC* vary. It is a 'disposal construction' (e.g., Song, 1979, 1981; H. Wang, 1981, 1985; L. Wang, 1943), 'a causative construction' (W. Hu, 2004; Ye, 2004), or 'a topicalisation structure' (cf. Hsueh, 1987; Tsao, 1987). The difficulty lies mainly in the fact

---

[1] The definition of Modern Chinese (Zhonghua renmin gongheguo guowuyuan guanyu tuiguang Putonghua de

              xiàndài hànyǔ shì yǐ běijīng fāngyán wéi biāozhǔn yīn   yǐ běifāng fāngyán wéi jīchǔ fāngyán
zhishi, 1956) is: 现代 汉语 是 以 北京 方言 为 标准 音，以 北方 方言 为 基础 方言，
(http://www.gov.cn/test/2005-08/02/content_19132.htm).

that linguists have attempted to find a single overarching function which covers all types of *BC*. As a corollary, the inconclusive descriptions of the function of *BC* have raised difficulties in teaching this structure to second language (L2) learners.

The challenge in teaching *BC* is partly derived from the inconclusiveness of the rules for *BC* in pedagogical grammar. Instructors have struggled to provide a consistent, explicit and clear explanation. In addition, there are no agreed instructional techniques for teaching *BC*. It is also notoriously difficult for L2 learners of Chinese to grasp *BC*. The difficulty is also attributed to its absence from other languages besides the lack of transparent and explicit rules. Considerable evidence for the difficulty in learning *BC* came not only from anecdotal experience but also the findings of acquisition studies. J. Shi (1998), for example, based on analyses of a written corpus of Chinese interlanguage, found that *BC* was one of the late-acquired grammar features among 21 basic Chinese sentence structures.

A few empirical studies have explored the developmental order of various types of *BC*. Although general agreement has been reached regarding the acquisition order, variation in the use of *BC* is clear from the studies. This raises the following questions – to what extent do L2 learners follow a definite acquisition order for *BC*? What factors will cause the variability in the use of *BC*? The current study attempts to answer these questions.

Variability in interlanguage has received growing attention in SLA since 1980. The significance of variability can be seen from the premise that 'an understanding of language is incomplete without an account of variability' (Regan, Howard, & Lemée, 2009, p.14) However, while a considerable number of studies have examined variability in the L2 learners' use of Indo-European languages such as English, French and Spanish, few studies have investigated variability in Chinese interlanguage, particularly, in structures which are especially prone to variable use. This study addresses this deficiency.

Instead of investigating learners' use of all types of *BC*, however, this study focuses on two types*: BC1* – a (locative) nominal *BC* whose complement contains a noun phrase (NP) and *BC2* – a directional verbal *BC* whose complement contains only a directional verb but no NP. These two types of *BC* were chosen for study on the basis of five criteria: high frequency of use by native speakers of Chinese, high productivity, prototypicality, their early presentation in the textbooks, and their early acquisition by L1 and L2 learners.

1.2 My experience of learning, teaching, and researching *BC*

1.2.1 My experience of learning *BC*

My interest in *BC* started in 1996. In my oral examination of the Graduate Admission Exam in Teaching Chinese as a Foreign [Second] Language at Peking University, one of the examiners asked me "*Weishenme women bu neng shuo 'Wo ba jiaozi chizai fanguan li'*?" [Lit. "Why cannot we say that 'I BA dumpling eat at restaurant inside'?"]. I failed to find an answer from my knowledge of Chinese grammar and was forced to provide a semantic explanation based on my personal understanding. Only after I became a graduate student did I realise that this was an open question for which there was no satisfactory and generally agreed answer.

I continued to seek an appropriate answer for this question during my three years' graduate study. In fact, there was no lack of linguistic explanations, but there were far too many of them (see Lu & Guo, 1998). Although I was satisfied with my explanation based on the interplay of syntactic, semantic and pragmatic features, when I applied this explanation of *BC* in tutoring L2 learners, my sense of achievement rapidly disappeared.

1.2.2 My experience of teaching *BC*

As part of my graduate study, I was involved in teaching Chinese as a second language to L2 learners. I remembered that after I briefly and excitedly addressed the properties of *BC* in my first class of teaching, I asked my students "*Nimen dong le ma?* [Do you understand?]". Most of them looked confused. One student nodded with confidence and said "*Wo dong le. Women ye keyi shuo 'ba fang shu zai zhuozi shang' chule 'fang shu zai zhuozi shang'* [lit. Yes, I understand. We could also say 'I BA put book on the desk' besides 'I put book on the desk]". Confronted with the frustration of trying to explain the grammar of *BC*, I tried to use meaning-focused questions to elicit *BC*. It seemed that this method was quite successful in class, but I seldom heard my students, particularly those at a relatively low level of proficiency, spontaneously using *BC* in their free conversations.

1.2.3 My experience of researching *BC*

In 1999 I explored the acquisition order of *BC* by examining 90 English, Japanese, and Korean learners, and compared this L2 acquisition order with the L1 acquisition order. At that time, studies on L2 acquisition in the field of Teaching Chinese to Speakers of Other Languages (TCSOL) were very welcome, but they were scarce. While doing my study in 1999, I worked towards finding a consistent 'natural order' for all types of *BC* irrespective of the learners' L1 and for both L1 and L2 acquisition (cf. Krashen, 1977, 1981, 1982, 1985). Variation in the data was so pervasive that I doubted the validity of the acquisition order I thought I had identified.

After I began teaching Chinese at The University of Auckland in 2005, I often asked myself the following questions: Will the learners in New Zealand follow the same developmental order of *BC* as those in China? To what extent will the acquisition order of *BC* be consistent and systematic? Is there variation in *BC* in the interlanguage of L2 learners? If so, what are the sources of variability?

1.3 Statement of the problem

Only a limited number of studies explored the 'natural order' of grammar features in Chinese during the 1990s when the morpheme studies of L2 English had passed their heyday (cf. R.Ellis, 2008c). However, in the past ten years, a large number of acquisition studies have examined the acquisition order of Chinese grammar features. In contrast, variability in the interlanguage of Chinese has received little attention (J. Shi, 2006), although considerable research has been done on variability in the interlanguages of other languages, especially English (Bayley, 1991, 1996, 2005; Bayley & Lucas, 2007; Bayley & Preston, 1996; Bayley & Schecter, 2003; R. Ellis, 1987c; Larsen-Freeman, 1975, 1976; Larsen-Freeman & Long, 1991; Tarone, 1982, 1983, 1985, 1987, 1988, 1989; Tarone & Liu, 1995; Tarone & Parrish, 1988; Wolfram, 1985; Young, 1989, 1991).

Why did Chinese researchers focus on systematicity rather than variability? There are two main reasons. First, there may have been an influence from the theoretical framework in linguistics. Chomsky's, universal grammar (UG) theory held a dominant and influential position in TCSOL as a western linguistic theory. This linguistic paradigm essentially ignores variation among language users (R. Ellis, 1994, 2008c). The second reason is lack of

methodological guidance. Acquisition studies and theories in TCSOL generally rely on translating SLA theories published in English. The lack of any Chinese translation of work on variability theories and methodologies limited this line of research.

To my knowledge, only a few studies have investigated variability in Chinese interlanguage. Two studies investigated phonological features – English speaking learners' acquisition of Chinese tones (B. Yuan, 1995) and Japanese learners' pronunciation of three consonants in Chinese syllables (*zh-*, *ch-*, and *sh-*) (Mei, 2005). One study investigated variability in English speaking learners' use of Chinese *wh*-words (B. Yuan, 2006, 2007) in *wh*-questions and another investigated the variable use of the morphological feature DE (Xiaoshi Li, 2010). Except for Mei's (2005) study, all of these studies were published in English. Language obstacles may have restricted the dissemination of this line of research. There were also a number of methodological limitations evident in these studies. For example, the sample size (n = 5 × 3 groups) in B. Yuan's study was too small to use a parametric statistical test.

1.4 Purpose of this study

Given the significance and complexity of variability in interlanguage and a lack of empirical studies on variability in L2 Chinese, the current research seeks to explore the sources of intra-learner variation (e.g., linguistic difficulty, task) and those of inter-learner variation (e.g., starting age, number of years of study, L1, setting, self-rated proficiency, and gender) in L2 learners' use of *BC*.

The current study can claim to be original in four main aspects. First, L2 Chinese has not been researched to the same extent as English and other Indo-European languages. Second, variability in Chinese interlanguage is still under-researched, particularly in the case of grammatical structures which are subject to variable use. Third, the choice of the target structures was based on actual use by native speakers of Chinese and the findings of previous acquisition studies. In addition, the current study compared two types of structures, one of which is subject to categorical rules and the other of which is subject to variable rules in L1. Fourth, variability in the use of *BC* was examined in terms of the effects of linguistic factors, psycholinguistic factors, individual factors and social factors.

This research is limited in the following ways. First, this research followed a cross-sectional rather than a longitudinal design to probe the sources of variability of *BC*. Second, this research was intended to investigate only two types of *BC* rather than all kinds of *BC*. Third, this research aimed to examine only a number of the major potential sources of variability.

1.5 Research questions

The research questions were as follows:

1. Is there intra-learner variation in the learners' use of *BC*?

    1a. Is there intra-learner variation in the learners' use of *BC* in terms of *BC* type?

    1b. Is there intra-learner variation in the learners' use of *BC* in terms of task type?

2. Is there inter-learner variation in the interlanguage of *BC*? If so, what are the sources of the inter-learner variation?

Research question 1a was examined by comparing the accuracy of use of two types of *BC*: *BC1* – a (locative) nominal *BC* whose complement contains a NP – and *BC2* – a directional verbal *BC* whose complement consists of a single or compound directional verb but no NP. Research question 1b was investigated by comparing the learners' accuracy scores in terms of four measures (i.e., oral production, oral imitation, untimed grammaticality judgements and error correction) in three tasks: an oral production task prompted by video clips (OPTV), an oral imitation task (OIT), and an untimed grammaticality judgement task (UGJT). To answer research question 2, the effects of six potential sources of inter-learner variation in the learners' use of *BC* (i.e., starting age, number of years of study, setting, L1, self-rated proficiency, and gender) were examined in term of oral scores (i.e., the average of oral production and oral imitation scores) and metalinguistic scores (i.e., the average of judgment and correction scores) for *BC1* and *BC2*. This decision was made based on the findings of the first research question. That is, the accuracy scores for *BC1* in the four measures differed significantly but those for *BC2* in the two oral measures did not. Additionally, the accuracy scores were significantly higher for the two metalinguistic measures than for the two oral measures. A number of parametric tests (e.g., *t*-test, one-way repeated ANOVA, stepwise regression, Pearson correlation coefficient) and non-parametric tests (e.g., Mann-Whitney analysis, Friedman Analysis) were employed. The qualitative data obtained in the follow-up interview aimed to provide rich and first-person explanations for the learners' performance in the tasks and to help interpret the quantitative findings.

1.6 Outline of the thesis

This thesis contains 9 chapters. Chapter 1 constitutes a broad introduction to the background of this study, a statement of the problem, the purpose of this research, and the research questions. This chapter ends with an overview of this thesis.

Chapters 2 to 4 review relevant literature to situate the current research in the context of what has been done, what is currently being researched and how research in L2 variability is conducted. Chapter 2 reviews theoretical and empirical perspectives on the potential sources of intra-learner variation in interlanguage. Chapter 3 provides a theoretical and empirical review of the potential sources of inter-learner variation in the interlanguage. Chapter 4 illustrates the linguistic characteristics of the target feature (*BC*), elaborates the criterion and procedure for selecting the target types of *BC*, and contrasts the syntactic and functional constraints on the target types of *BC*.

Chapter 5 reports the pilot study with native speakers of Chinese, including information on the purpose, the participants, the instruments and procedures, the results, reliability and validity of the instruments, and describes the changes made for the main study.

Chapter 6 details the methodology of the study, including information on research questions, the research design, the participants, the instruments, the procedures, the methods of analysis, reliability and validity, and methods of statistical analysis.

Chapters 7 and 8 report the results of the main study. Chapter 7 presents and discusses the results of the first research question regarding the sources of intra-learner variation in the interlanguage of *BC*. Chapter 8 presents and discusses the results of the second research question regarding the sources of inter-learner variation in the interlanguage of *BC*.

Chapter 9 summarises the main findings, concludes the thesis by considering some theoretical, methodological, and pedagogical implications and the limitations of this study, and provides suggestions for further research.

CHAPTER 2: SOURCES OF INTRA-LEARNER VARIATION

2.1 Introduction

The current chapter aims to provide contextual background for two sources of intra-learner variation in interlanguage: linguistic difficulty and task type. First, typologies of variation and potential sources of intra-learner variation are introduced, followed by a review of studies that have investigated these two sources. The criteria for determining objective learning difficulty are considered, along with a review of studies that have examined objective learning difficulty. Next subjective difficulty is considered. There follows a consideration of task-based variability. This chapter ends with a discussion of the potential relevance of the existing research to the current study.

2.2 Variation in interlanguage

2.2.1 Typologies of variation

In SLA it has been generally agreed that L2 learners' interlanguage is characterised by pervasive variation (e.g., Bayley, 1996; Bayley & Lucas, 2007; N. Ellis & Larsen-Freeman, 2006; R. Ellis, 1988, 1994, 2008c; Preston, 1996; Romaine, 2003; Tarone, 1982, 1983, 1985, 1988). However, taxonomies of variation vary according to the distinguishing criteria and research paradigm. Variation can be distinguished as internal variation or external variation depending upon whether the sources of variation are linguistic in nature or beyond the linguistic domain (e.g., social factors and individual factors) (Adamson, 1988; R. Ellis, 1985, 1994, 2008c; Preston, 1989; Tarone, 1988). Variation can also be categorised as either vertical variation or horizontal variation depending on whether variation in interlanguage occurs over time or at any one point in time. Further, it can also be classified as either systematic variation or free variation, the former being regular and predictable and the latter the opposite (R. Ellis, 1984, 1985, 1994, 2008c; Larsen-Freeman & Long, 1991; Tarone, 1988; Young, 1996). Taking learners as the focus of inquiry, the distinction between intra-learner variation and inter-learner variation becomes important (R. Ellis, 2008c). The former refers to variation that occurs in each individual learner's interlanguage, while the latter refers to the variation that exists among the interlanguages of groups of learners. For the purpose of this chapter, only the sources of intra-learner variation will be considered.

2.2.2 Potential sources of intra-learner variation

Intra-learner variation has received considerable attention since Dickerson (1974, 1975) examined variation in the pronunciation of Japanese ESL learners. To date, a general consensus has been achieved that intra-learner variation can be attributed to linguistic factors and task-related factors (N. Ellis & Larsen-Freeman, 2006; R. Ellis, 1985, 1987a, 1989, 1992, 1994, 2008c; Tarone, 1983, 1985, 1988). Linguistic factors can be further divided into linguistic difficulty and linguistic contexts. The former focuses on the inherent properties that make a linguistic structure difficult to learn, whereas the latter considers the external linguistic environments within which a linguistic structure is situated (e.g., preceding or following vowels or consonants or syntactic constituents). Due to the difficulty in identifying the linguistic contexts of *BC*, the current study limits itself to investigating the effect of inherent linguistic difficulty as one of the internal sources of intra-learner variation, and the effect of task type as an external source of intra-learner variation.

2.3 Linguistic difficulty

Studying the linguistic difficulty of grammar structures is appealing to SLA researchers. This is partly because one of the major goals of SLA research is to describe learner language and reveal how learners' competence in using a second language develops (R. Ellis, 2008c), and also because it has been shown that some structures are more difficult to learn than others. However, it is not easy to establish a clear classification of the factors which are responsible for the difficulty of grammatical structures. For instance, it is not yet clear to what extent the difficulty in learning a second language is derived from the inherent properties of grammar structures and to what extent it is determined by the learning process.

It is therefore necessary to make a distinction between linguistic difficulty and learning difficulty. The former refers to the inherent linguistic characteristics that make grammatical structures easy or difficult to learn, whereas the latter is concerned with the difficulty that learners actually experience in learning grammatical structures. Linguistic difficulty is determined by a set of objective criteria that can predict the level of learning difficulty of different structures, while learning difficulty is demonstrated by collecting data from learners to see which structures they find 'easy' and which 'more difficult'. In this sense, linguistic difficulty and learning difficulty are referred to as objective difficulty and subjective

difficulty respectively by DeKeyser (2003). A definition of objective difficulty and subjective difficulty is provided in R. Ellis (2006):

> Objective difficulty concerns the inherent difficulty of different grammatical features. It is determined by reference to some theory of grammar that allows predictions to be made about which features will be easy and which difficult to learn. Subjective difficulty refers to the actual difficulty that individual learners experience when learning a second language (L2) (p.431)

The current research is only concerned with the learning difficulty of *BC*. However, criteria of objective linguistic difficulty and the relevant studies will be briefly introduced to provide a suitable research context and establish a theoretical basis for interpreting the findings of this study.

2.3.1 Criteria for determining the objective difficulty of different grammatical structures

The study of linguistic difficulty of grammatical structures pre-dates the establishment of SLA as a field of enquiry in the 1960s (R. Ellis, 1994, 2008c). Some researchers and instructors who were engaged in teaching foreign languages at that time proposed the Contrastive Analysis Hypothesis (CAH) to predict the difficulty of the L2 grammatical structures (e.g., Lado, 1957; Weinreich, 1953). This hypothesis assumed that all difficulty in learning L2 grammatical features could be predicted and explained by the language distance (i.e., their similarities and differences) between the L1 and the L2. That is, the more similar the L2 was to the L1, the easier to learn, and vice versa. However, this hypothesis was empirically discredited and received considerable criticism during the 1960s and 1970s. Empirical studies found that the CAH either under-predicted (Hyltenstam, 1977; Whitman & Jackson, 1972) or over-predicted the interference from the L1 (Dulay & Burt, 1974). Most seriously, however, behaviourism, the theoretical underpinning of the CAH, was heavily criticised by Chomsky in 1959 through his critique of Skinner's *Verbal Behavior* (see, Larsen-Freeman & Long, 1991). Furthermore, a fatal theoretical flaw was the 'dubious assumption that one could depend solely upon an analysis of a linguistic product to yield meaningful insight into a psycholinguistic process' (Larsen-Freeman & Long, 1991, p.56 ; Long & Sato, 1984). Accordingly, the CAH was very quickly rejected as a useful explanatory

device. However, its failure motivated researchers to further examine learner language, thus encouraging the establishment of SLA as a field of enquiry.

The exploration of the difficulty of grammatical structures in SLA originated in studies of the L2 acquisitional order, which aimed to address whether learners acquire some target language features before others. Although there is considerable evidence to support the existence of systematic and universal developmental patterns in L2 acquisition, no consensus has been reached to explain the systematic order that was found (see, DeKeyser, 2005; Gass & Selinker, 2001; Larsen-Freeman & Long, 1991; Long & Sato, 1984). Nor has much attention been paid to exploring objective criteria for predicting the difficulty of grammatical structures.

2.3.1.1 Determinants of the difficulty of English morphemes (Goldschneider & DeKeyser, 2001)

Goldschneider and DeKeyser (2001) conducted a meta-analysis to explain the 'natural order' of L2 morpheme acquisition in English. The study examined six functors (i.e., present progressive *-ing*, regular past *-ed*, articles, plural *-s*, possessive *-'s*, and third singular *-s*) all of which were investigated in 12 morpheme studies reported between 1973 and 1996. Multiple regression analysis, based on the oral production data of 924 adults and children, showed that a combination of the following five determinants could explain the acquisition order. Table 1 lists the determinants, their definitions and associated predictions that have been tested.

Table 1: Determinants of the Natural Order of L2 Morpheme Acquisition in English (Goldschneider & DeKeyser, 2001)

Determinant and subfactor	Definition	Assumption/ Prediction
1. Perceptual salience	How easy it is to hear or perceive a given structure.	The more perceptually salient a functor is, the earlier it will be acquired.
a. phonetic substance	The number of phones in the functor	The more phones in a functor, the more perceptually salient it should be
b. syllabicity	Presence/ absence of a vowel in the surface form	Functors containing a vowel in the surface form should be more perceptually salient than those without a vowel
c. sonority	Numerical values based on sonority hierarchy (Laver, 1994)	Functors that are more sonorous should be more salient
2. Semantic complexity	A measure of how many meanings are expressed by a particular form	Forms with more meanings should be more difficult to learn and later acquired than forms with fewer meanings
3. Morphophonological regularity	The degree to which the functors are (or are not) affected by their phonological environment	The more phonologically regular a functor is, the earlier it should be acquired
a. number of phonological alternations	The number of phonological alternations found for each functor	Functors with more alterations should be acquired later
b. homophony with other grammatical functors	Whether there is homophony with other grammatical functors	A negative correlation is expected.
4. Syntactic category	The characteristics of each functor from the perspective of Functional Category theory. The syntactic category consists of lexical/functional groups which are further subdivided according to the free/bound distinction (Zobl & Liceras, 1994)	The category with the higher point score was assigned to the forms predicted to be acquired earlier. The acquisition order for syntactic category is: Lexical Free> Lexical Bound > Functionally Free> Functionally Bound
5. Frequency in input	The number of times a given structure occurs in speech addressed to the learner	The more frequent a grammatical item is in the input to the learner, the more easily and quickly that item should be acquired.

In this study, Goldschneider and DeKeyser reiterated that it was difficult to attribute the 'natural order' (i.e. the nearly same order of acquisition of certain grammatical features by ESL learners at different ages, with different types of exposure to English, and different L1s) to a single determinant (cf., Larsen-Freeman, 1976; Pienemann, 1998; Zobl, 1995; Zobl & Liceras, 1994). They argued that 'these five factors are not a completely heterogeneous set, but can all be seen as aspects of salience in a broad sense of the word, and that this salience at various levels (phonological, morphological, syntactic semantic, and numerical) facilitates the process of induction of grammatical structure from elements of the input' (p.37). However, this meta-analysis was restricted by the limited information in the original studies. For example, the most likely potential predictor, L1 transfer, could not be examined because of the absence of original data. In addition, the final criteria were reached based on only L2 morpheme studies in English, and thus cannot be generalised to other languages.

2.3.1.2 Determinants of grammatical difficulty (DeKeyser, 2005)

DeKeyser (2005), drawing on empirical evidence from acquisition studies on a wide variety of L2s, proposed five determinants of grammatical difficulty. These determinants and their definitions and predictions are summarised in Table 2.

Table 2: Determinants of Grammatical Difficulty (DeKeyser, 2005)

Determinant	Definition	Prediction
1. Complexity of form	Ease or difficulty in picking up the right forms (e.g., morphemes and allomorphs) to express certain meanings and putting them in the right place (e.g., inflection is relatively difficult to acquire)	Especially in richly inflected languages, agglutinative, polysynthetic, or inflectional structures are more difficult to learn.
2. Complexity of meaning	The extent to which a grammatical feature involves novelty, abstractness, or a combination of both in meaning.	The features expressing highly abstract notions that are extremely hard to infer, implicitly or explicitly in input, are hard to acquire and even strongly resistant to instructional treatments.
3. Complexity of meaning-form relationship	The extent to which the link between form and meaning is transparent	The less transparent the link between form and meaning, the more difficult the grammatical features.
a. redundancy	Whether the form is not semantically necessary because the meaning is also expressed by at least one other element of the sentence (e.g., a verb ending, third person -s)	
b. optionality	Presence/ absence of alternatives for expressing the same meaning (e.g., null subjects in Spanish or Italian, and case marking in Korean)	
c. opacity	A complex form of the problem of low form-meaning correlation (e.g., -s in English can be the third person singular of the verb, the plural of the noun, or the genitive of the noun and in each case has the same three allomorphs)	
4. Frequency	According to N. Ellis (2002, 2003), the typical route of acquisition of grammatical structures is form formulae through low-scope patterns to constructions. The abstraction of regularities within these constructions is frequency-based.	In principle, the importance of frequency is independent of semantic transparency, but how important frequency is depends to some extent on the transparency of the mapping.
5. Salience	Salience prevails over markedness in acquiring pied piping and preposition standing in ESL (Bardovi-Harlig, 2000); there are multiple components that contribute to salience (combination of phonological salience, semantic complexity, morphological regularity, and frequency) strongly intercorrelated in English morphology (Goldschneider & DeKeyser, 2001)	Lack of salience plays an important role in acquisition difficulty.

As DeKeyser (2005) points out, the first three determinants – complexity of form, complexity of meaning, and the complexity of form-meaning mapping – 'leave out the difficulty of grasping the form-meaning relationship while processing a sentence in the L2'. He asserts that:

> [I]t is the transparency of form-meaning relationships to a learner who is processing language for meaning that determines the difficulty of acquisition, at least for learners who are left to their own resources instead of presented with a reasonably complete set of rules about form-meaning relationships (p. 3).

DeKeyser, drawing on N. Ellis's (2002, 2003) work, identified frequency as an important factor that contributes to the difficulty of learning form-meaning mappings. He also argued that the effect of frequency depends to some extent on the transparency of the form-meaning mapping. For instance, 'if the mapping is very obscure, the structure may will never be acquired by adults' (p.11).

DeKeyser, on the basis of previous empirical evidence, inferred that a lack of salience played an important role in determining the difficulty of acquisition. Following a scrutiny of research into acquisition orders (Bardovi-Harlig, 1987; Goldschneider & DeKeyser, 2001) and ultimate-attainment relations (DeKeyser, Alfi-Shabtay, Ravid, & Shi, 2005; DeKeyser, Ravid, & Alfi-Shabtay, 2005), he argued that much remained to be done, such as operationalising salience in syntax as opposed to morphology or phonology.

Despite drawing on substantial past empirical evidence, DeKeyser's hypothesis has not been directly tested by empirical studies. More work is required, particularly meta-analysis studies of the acquisition of typologically different target languages.

2.3.1.3 Determinants of grammatical difficulty as implicit knowledge and explicit knowledge (R. Ellis, 2006)

To predict the learning difficulty of grammatical structures, R. Ellis (2006) examined the psycholinguistic characteristics of implicit knowledge and explicit knowledge (R. Ellis, 2004, 2005a). Drawing on the previous work (N. Ellis, 1996a, 1996b; Goldschneider & DeKeyser, 2001; J.H. Hulstijn & de Graaff, 1994; Pienemann, 1998), he suggested five determinants to

explain the acquisition of different grammatical structures as implicit knowledge (see Table 3 for their definitions and predictions).

Table 3: Determinants of the Difficulty of Learning Grammatical Structures as Implicit Knowledge (R. Ellis, 2006)

Determinant	Definition	Prediction
1. Frequency	How frequently does the grammatical feature occur in the input?	Features that occur frequently in the input will be easier to acquire than features that occur infrequently (based on N. Ellis, 1996b, 2002).
2. Saliency	Is the grammatical feature easy to notice in the input?	Phonologically salient features are easier to acquire than those that are less phonological salient.
3. Functional value	Does the grammatical feature map onto a clear, distinct function?	The forms that realise a single function and that are typically non-redundant cater to the learner's One-to-One Principle (Andersen, 1984) are easier to learn than forms that realise multiple functions or that are always or often redundant.
4. Regularity	Does the grammatical feature conform to some identifiable pattern? The regularity can be distinguished according to two aspects: scope which concerns the number of cases that a rule can cover, and reliability which concerns the extent to which a rule holds true (J. H. Hulstijn & de Graaff, 1994).	Regular features will be easier to acquire than irregular features
5. Processability	Is the grammatical feature easy to process according to the hierarchy of processing procedures proposed by Pienemann (1998)?	The features at higher stages of the hierarchy of processing procedures are relatively more difficult to learn

R. Ellis (2006) also proposed determinants for learning grammatical structures as explicit knowledge drawing on the distinction between analysed knowledge and metalinguistic knowledge (R. Ellis, 2004) and the distinction between structural complexity and the complexity of the accompanying explanation (Robinson, 1996). He identified two principal factors that determine the difficulty of declarative rules of grammar: conceptual clarity and metalanguage. Conceptual clarity can be considered from four perspectives, while metalanguage concerns the extent to which technical terms are required to describe the declarative rules (see Table 4 for detailed definitions and predictions).

Table 4: Determinants of the Difficulty of Learning Grammatical Structures as Explicit Knowledge (R. Ellis, 2006)

Determinant	Definition	Prediction
1.Conceptual clarity	Conceptual clarity can be considered from four perspectives:	The structures with conceptual clarity are easier to learn
	a. Are the form and function complex?	The structures with both formal and functional complexity are more difficult to learn
	b. Is there a transparent, general rule or is there an expressible rule, which is related to the distinction between 'rule learning' and 'item learning' (J.H. Hulstijn & de Graaff, 1994)?	Structures for which clear (or true) rules can be formulated can more easily be learned as explicit knowledge than structures that necessarily involve item-learning
	c. If rules are expressible, how many different formal or functional grammatical features contribute to the specific form of a target structure and the specific function it performs (de Graaff, 1997, p.41) ?	A pedagogic rule will be relatively simple if it addresses fewer forms which transparently related to fewer functions.
	d. Is the rule prototypical or peripheral (G. Hu, 2002) ?	Prototypical functions/ structures are easier to learn than peripheral functions/structures.
2. Metalanguage	The extent to which technical terms are required to describe the declarative rules	The more technical terms required, the greater the difficulty of learning the explicit knowledge.

R. Ellis further tested these criteria by examining the learning difficulty of 17 grammatical structures in L2 English. A detailed review of Ellis's (2006) empirical study is provided in Section 2.3.3.

2.3.1.4 Processability theory (Pienemann, 1998, 2005)

Pienemann (1998, 2005) has proposed a universal framework for predicting L2 development – Processability Theory (PT) – based on language processing research such as Levelt's (1989) psychological model of language production and Lexical-Functional Grammar which brings linguistic theory and psycholinguistic theory closer to each other (Bresnan, 1982; Kaplan & Bresnan, 1982). The principles of processability involve a hierarchical set of processing procedures and routines. Table 5 presents the definition of each procedure, as well as examples from both English and Chinese.

Table 5: Hierarchy of Processing Procedures with Examples in English and Chinese

Hierarchy	Definition	Example			
		In English		In Chinese	
		Syntax	Morph-ology	Syntax	Morphology
5. Matrix / subordinate clause procedure	To process the word order of subordinate structures	Embedded questions		$(S)V_1OV_2O$; (S) *ba* OVC (DC/RC/PP)	
4. S-procedure and Word Order rules	Exchange of information between heads of different phrases (either without or with saliency)	INV	S-V agreement	OSV; (S) ADVO/ ADSVO; (S) *ba*O V *le*	
3. Phrasal procedure (head)	Exchange of information is required to check if the value of diacritic feature of one lexical entry matches that of another in order to produce a structural phrase	ADV Articles (NP)	Plural agreement (on Det)	ADV ; PP ; *de* P (XP *de* N)	Aspectual particles: e.g. *zhengzai-* Particle :- *de*
2. Category procedure (lexical category)	Access to diacritic features of lemmata, which are within a single constituent and matched with the underlying conceptual content of a message	SVO	'Number' on nouns; 'Tense' on verbs;	SVO;	Compound directional verbs (-DC); Aspectual suffixes : -*le* -*zhe*, -*guo*
1. Word/lemma access	Access to L2 words, which are in invariant forms and single-constituent utterances	Single words		Single words	

Note. $(S)V_1OV_2O$ = a verbs-in- series construction (e.g. *qu xuexiao shang ke*, 'go to school to attend a class'; (S) *ba* OVC (DC/RC/PP) = the *ba* construction with a complement such as a directional complement (DC), resultative complement (RC), and prepositional phrase (PP) (e.g., *ba shu na chulai/ zou/ dao xuexiao* 'take the book out / away / to school'); (S)AVO/ ASVO= variable adverbial placement in the sentential procedure (e.g. *wo mingtian qu Zhongguo* / *mingtian wo qu Zhongguo* 'I will go to China tomorrow'). PP = a prepositional phrase (e.g. *zai fangjian li* 'in the room'); de P (XP de N) = any phrase+ de +noun (e.g., *wo de shu* 'my book(s)')

As Pienemann points out, there is a basic difference between the first three procedures and the last two. That is, structures in the early stages involve 'a direct mapping of argument structure onto functional structure' (Pienemann, 2005, p.14) while those in the later stages are

constructed in accordance with constituent structure rules. The PT makes predictions about the emergence of grammatical features and their learning difficulty in relation to the nature of the processing procedures required to produce these forms.

Predictions based on PT have received general support from empirical studies using both longitudinal and cross-sectional designs (Pienemann, 1987, 1989, 1998) and across a range of typologically diverse target languages (for Chinese see e.g., Charters, 2005; Xiaodan Gao, 2005; Y. Zhang, 2001, 2002, 2004, 2005; for Arabic see e.g., Mansouri, 2005; for Japanese see e.g., Kawaguchi, 2005). However, discrepancies do exist. For example, Charters (2005) argues that PT over-simplifies the processing procedures in the acquisition of Chinese nominal syntactic and morphological features. Dewaele and Veronique (2001) demonstrate that PT is not a suitable conceptual framework for predicting the acquisition of French adjectives. According to Pienemann's (2005) explanation, gender is a lexical feature and thus has to be acquired for every lexical item. Such a requirement may have 'rendered the application of PT to this phenomenon in French superfluous' (p. 62). Most importantly, PT does not make any predictions about levels of accuracy, but instead utilises the emergence as the criterion of acquisition.

However, the assumption of PT that formal interventions cannot alter a postulated sequence of acquisition does inevitably encounter difficulties in two aspects: a) developmental gaps correlating with certain defined aspects of interlanguage variation, and b) form-function relationships or the complexity of form-meaning mappings (i.e., complexity in many-to-many mappings).

PT struggles to deal with variation in L2 acquisition. In particular, variationists and language testers in SLA question the extent to which the universal developmental procedures that underlie the acquisition of grammar structures are sensitive to, and affected by, individual variation that may result from different elicitation contexts (e.g., Bachman, 1988; Crookes, 1989; Crookes & Gass, 1993; Douglas, 1986; Selinker & Douglas, 1985; Tarone, 1983, 1985). In response, Pienemann (1998) suggested a 'steadiness hypothesis' which states that the basic nature of the grammatical system of an interlanguage does not vary when tasks are based on the same skill type in language production. This hypothesis received strong support from Pienemann's (1998) empirical study involving 6 ESL learners (4 Indonesian, 1 Chinese, 1 Korean). These learners' performance on 6 communicative tasks (i.e., Habitual Actions,

Story Completion, Informal Interview, Picture Sequencing, Picture Differences, and Meet Partner) was tested for steadiness using well-defined emergence criteria in the area of syntax and morphology. In the case of syntax, 'all samples were perfectly consistent across all participants, all structures and all tasks. In morphology, the consistency was 99.1%. This constitutes overwhelming support for the steadiness hypothesis' (p. 308).

To deal with the issue of complex form-function mapping, Pienemann (2005) added a second set of principles that contribute to the formal modelling of levels of processability, namely, the mapping of argument-structure onto functional structure and the mapping of constituent structure onto functional structure (Pienemann, Di Biase, & Kawagichi, 2005). By doing this, linguistic non-linearity can be accounted for by including discourse functions and Lexical Mapping Theory (LMT). LMT systematically explains how the conceptual representation of thematic roles is mapped onto the grammatical functions mediated by a structure which has both a semantic aspect (i.e., an aspect 'that specifies the core participants in events') and a syntactic aspect (i.e., an aspect 'that provides the minimal information required to identify the dependents of an argument-taking head') (Bresnan, 2001 cited from Pienemann et al. 2005, p. 212) .

As a universal framework, PT hypothesises that 'L1 transfer is constrained by the capacity of the language processor of the L2 learner irrespective of the typological distance between the two languages' (Pienemann, et al., 2005, p. 85). Research has shown that when learning a language typologically close to the L1, transfer does not necessarily occur with learning features at the initial stages of the processability hierarchy, but does when learning those located further up the processability hierarchy. This is because the interlanguage has developed the necessary processing prerequisites.

2.3.1.5 Comments

All four sets of criteria described above can be used to predict learning difficulty of grammatical structures. However, they do differ in a number of respects. First, they were different in the theoretical bases. Goldschneider and DeKeyser's (2001) determinants of grammatical difficulty were based on a meta-analysis of studies of the 'natural order'of acquisition of L2 English morphemes. DeKeyser's (2005) criteria for determining grammatical difficulty were developed on the basis of empirical evidence from a variety of

acquisition studies of typologically diverse languages. The criteria for predicting grammatical difficulty in R. Ellis (2006) were developed to explain the learning difficulty of implicit knowledge and explicit knowledge. In Pienemann's (1998, 2005) PT, grammatical difficulty was predicted according to the emergence procedures of psycholinguistic language processing. In addition, PT is the only theory that has attempted to predict learning difficulty in terms of language-specific rules and the role of L1 transfer.

A second difference is in the relationships among the proposed determinants. In the PT, the processing procedures of the PT can neither be reversed nor skipped (Pienemann, 1998; 2005). In contrast, in the other three proposals the objective determinants predicting the difficulty of grammatical structures are simply listed without clearly indicating how they are combined to predict learning difficulty. In addition, in DeKeyser's studies (Goldschneider & DeKeyser, 2001; DeKeyser, 2005), salience alone, broadly understood, accounts for all five determinants of the grammatical difficulty of English morphemes.

A third difference lies in the extent to which the theoretical predictions have been empirically tested. The PT has received extensive support from empirical studies of a wide range of typologically diverse languages. R. Ellis (2006) has also empirically tested the objective criteria he identified (see Section 2.3.3 for a detailed review). Nevertheless, little has been done to test the objective criteria raised by Goldschneider and DeKeyser (2001) and DeKeyser (2005), although the objective criteria used by Goldschneider and DeKeyser (2001) were derived from an attempt to explain the order of acquisition reported in the morpheme studies.

Because in the current study the acquisition order was measured according to accuracy order rather than emergence order, a detailed review of research into the PT will not be provided. Instead, Goldschneider and DeKeyser's (2001) study and R. Ellis's (2006) investigation of the learning difficulty of 17 English grammatical structures will now be considered in detail.

2.3.2 Studies that report on the learning difficulty of different grammatical structures (subjective learning difficulty)

This section reviews studies that report on the subjective learning difficulty of different grammatical structures.

21

The learning difficulty of grammatical structures is closely related to the developmental order of grammatical features in learner language. Actual morpheme studies can serve as evidence for the subjective learning difficulty of different grammatical structures, as in Goldschneider and DeKeyser (2001).

The investigation of the 'natural order' of acquisition of L2 morphemes in English began in the 1970s and was a notable milestone in SLA research. A considerable number of empirical studies have found that English morphemes manifest a universal order of acquisition or accuracy. That is, the acquisitional order of certain English morphemes was nearly identical for L2 learners irrespective of their L1s (e.g., Dulay & Burt, 1973, 1974), age (i.e. whether the subjects are adults or children) (Bailey, Madden, & Krashen, 1974), and, more controversially, types of data (i.e., whether the study was based on oral or written data) (Krashen, Butler, Birnbaum, & Robertson, 1978).

However, the 'natural order' of English morphemes found by Dulay and Burt has been replicated in only a few studies. Other studies (e.g., Hakuta, 1974; Rosansky, 1976) have reported different orders. These discrepancies have been attributed to the fact that in the longitudinal studies accuracy levels were sometimes calculated using fewer than ten obligatory occasions (Krashen, 1977). The differences 'do cast doubt on the validity of equating cross-sectional accuracy with acquisition orders' since mean scores of cross-sectional accuracy can be misleading (R. Ellis, 2008, p. 86). Furthermore, the so-called 'natural order' of acquisition was found to vary with the elicitation tasks in some studies (e.g., Larsen-Freeman, 1975; LoCoco, 1976). A detailed review of the effect of tasks on the variation in acquisition orders is considered in Section 2.4.

'Natural order' studies also have other methodological problems. First, the differential difficulty of grammatical structures has not been considered. All target structures were placed in different ranks according to relative order irrespective of the magnitude of the differences between their accuracy of use. To overcome this problem, Dulay and Burt (1975) and Krashen (1977) proposed grouping the grammatical features with similar accuracy. Second, the accuracy order was calculated based on obligatory occasion analysis which did not take into account overuse (i.e., misuse in inappropriate contexts) of target features. To deal with this flaw, Pica (1983) employed target-like use analysis to calculate accuracy of use and concluded that the acquisition order was not affected even when overuse was considered.

Finally, the acquisition order of morphemes was restricted to a small set of morphemes, so it could only provide a partial picture of acquisition.

It should be noted that, although morpheme studies ceased in the early 1980s, the study of the 'natural order' of grammatical features remains active. In effect, research into the developmental order of L2 Chinese grammatical structures began in the 1990s in the field of Teaching of Chinese as a Foreign/Second Language (Xiaoping Gao, 1999; H. G. Jin, 1993; Qian, 1997; J. Shi, 1998). The studies that have investigated the developmental orders of the *ba* constructions will be reviewed in Chapter 4 (see Section 4.3.3.5.2 for more details).

2.3.3 Studies that relate predictive criteria to actual learning difficulty

The criteria for predicting the learning difficulty of grammatical feature**s** have been tested in a few empirical studies by examining actual learning difficulty. R. Ellis (2006), for example, investigated the relative difficulty of 17 English grammatical features in terms of implicit knowledge and explicit knowledge. The subjects consisted of 220 learners with mixed language proficiency (n1 = 147 ESL learners in New Zealand with mixed L1s, mainly Chinese; n2 = 28 Japanese learners in Japan; n3 = 54 TESOL students in Malaysia). The data were elicited using a battery of five tests (i.e., an Oral Narrative Test [ONT], an Oral Imitation Test [OIT], a Timed Grammaticality Judgement Test [TGJT], an Untimed Grammaticality Judgement Test [UGJT], and a Metalinguistic Knowledge Test [MKT]). A Principal Component Analysis yielded a two-factor solution, the three tests (i.e., the ONT, the OIT and the TGJT) loading on one factor and the other two tests (i.e. the UGJT and the MKT) loading on the other. The first factor was interpreted as implicit knowledge and the second factor explicit knowledge. Stepwise regression analyses showed that structures that were easy in terms of implicit knowledge were often difficult in terms of explicit knowledge and sometimes vice versa. That is, the accuracy orders of the 17 grammatical structures in terms of the two types of knowledge were not correlated. The study also found that the structures varied as to whether implicit or explicit knowledge of them was related to general language proficiency as measured by the academic version of the International English Language Testing System (IELTS). The scores of four components (i.e. listening, reading, speaking, and writing) and their sum were considered. Both implicit and explicit measures of grammatical structures predicted a substantial amount of the variance in general proficiency

scores. R. Ellis's study suggests that the learning difficulty of grammatical structures should be considered separately in terms of implicit knowledge and explicit knowledge.

This argument was reinforced by R. Ellis's (2008) study of four grammatical structures. R. Ellis (2008) investigated whether data elicited by instruments designed to provide separate measures of implicit and explicit second language afforded a valid basis for determining what learners had learned. The data was elicited from the same groups of learners using the same instruments as in R. Ellis (2006) except for the Oral Narrative Test, which was excluded. The study tested predictions derived from Pienemann's Processability Theory regarding the learning difficulty of four grammatical structures (i.e. possessive *-s*, *since/for*, 3rd person *-s*, and question tags) chosen to represent the four procedures of PT: category procedure, phrasal procedure*, -s* procedure, and subordinate clause procedure, respectively). The results showed that the predictions were borne out for implicit knowledge but not for the data obtained from the tests of explicit knowledge. The study suggests that experimentally elicited data can be used to examine interlanguage development (i.e., how learners' implicit knowledge develops) and to make statements about learners' grammatical proficiency.

2.3.4 Relevance to the current study

The four types of objective criteria for predicting grammatical structures reviewed above not only provide a research context for examining the learning difficulty of the target features in the current study, but also serve as a theoretical basis for interpreting the results of the current research.

2.4 Task effects

2.4.1 Definition of tasks

Tasks have been recognised as one of the sources of variation in interlanguage (Bygate, 1999, 2001; Bygate, Skehan, & Swain, 2001; N. Ellis & Larsen-Freeman, 2006; R. Ellis, 1985, 1987d, 1992, 1994, 2008c; Foster & Skehan, 1996; Larsen-Freeman, 1975, 1976; Larsen-Freeman & Long, 1991; LoCoco, 1976; Skehan, 1998; Sorace, 1985; Tarone, 1983, 1985, 1987, 1988, 2007). However, definitions of tasks vary considerably. One of the difficulties is that the word 'task' is used to describe different things. It has a general meaning (i.e., a task is any activity that elicits some performance from learners including metalinguistic performance

and production). It also has a more specific meaning in language teaching as 'a language-teaching activity where meaning is primary, there is some kind of gap, students are required to use their own linguistic resources, and there is an outcome other than the display of language for its own sake.' (R. Ellis 2008, p. 919). I will use the task with its general meaning in the study reported in this thesis.

2.4.2 Theoretical perspectives on task effects

The issue of what data is most appropriate for examining interlanguage has received debate among researchers since the early days of SLA (Corder, 1973; Larsen-Freeman, 1975; Selinker, 1972; Tarone, 1979, 1982). While some researchers (e.g., Adjemian, 1976, 1981; Arditty & Perdue, 1979) suggest that intuitional data is appropriate to study interlanguage, others (e.g., Selinker, 1972) assert that intuitional data should be ruled out because it only provides information about learners' intuitions of the target language system. Instead, utterances produced by L2 learners are more suitable data to investigate acquisition. Similarly, while some researchers are inclined to emphasise the plausibility of using a certain type of task (e.g. intuitional data, spontaneous speech, or elicited imitation data) to collect learner samples (e.g., Adjemian, 1976, 1981; Naiman, 1974), others advocate the importance of employing multiple types of tasks to get a full description of interlanguage (e.g., Corder, 1973, 1981; Selinker, 1972; Tarone, 1982, 1983). The theoretical perspectives that lend support to the second argument will be introduced below.

2.4.2.1 Corder's three types of data

Corder was one of the earliest researchers to advocate the use of multiple types of data. In expounding error analysis, Corder (1967) stated that the nature and quantity of errors is likely to vary depending upon whether the data reflects natural and spontaneous use or careful elicitation. Although natural samples are generally preferred, learners often do not produce sufficient spontaneous data. This led Corder (1973) to argue that three kinds of data are necessary to investigate learner language. These were naturally occurring data, clinically elicited data, and experimentally elicited data (Corder, 1976; reproduced in Corder, 1981). In contrast to naturally occurring data, Corder distinguished between clinical elicitation and experimental elicitation according to its purpose. The former attempts to get 'the informant to produce data of any sort', the latter is designed to elicit data incorporating the particular

linguistic features that the researcher intends to investigate at that moment (cited from R. Ellis 2008, p. 46).

Naturally occurring samples of learner language are usually produced in real-life situations and have to be collected through observation. Clinically elicited data can be collected by means of general interviews or asking learners to write a composition, while experimentally elicited data can be obtained by asking learners to perform tests such as the Bilingual Syntactic Measure (see Burt, Dulay, & Hernández- Chávez, 1975), an oral imitation task, translation, fill-in-blanks and a grammaticality judgement task. A detailed account of the types of tasks and the data they provide is included in Section 2.4.3.

Selinker (1972) pointed out that data collected from different sources might provide access to L2 learners' different underlying linguistic systems. For this reason, researchers have increasingly used multiple tasks to elicit different kinds of data in order to achieve a complete description of learner language. Explanations for using multiple types of tasks to collect data vary according to the research paradigm.

2.4.2.2 Stylistic continuum and sociolinguistic accounts

Tarone (1979, 1982, 1983), in her seminal research argued that interlanguage should be viewed as a system which systematically varies according to the linguistic environment or the data elicitation task. Learners' performance in these tasks creates a style-shifting continuum, ranging from the 'vernacular style' to the 'careful style' depending upon the extent to which they pay attention to form. Drawing on Labov's (1969) Observer's Paradox, she further argued that the vernacular style is the most systematic style.

Tarone's style-shifting model is based on a distinction between capability and competence. According to Tarone (1983), competence generally refers to the linguistic knowledge underlying learners' use of the target language, whereas capability refers more broadly to the regularities underlying all language behaviours 'in learner production and perception, writing and reading, as well as in making judgements on grammaticality' (p. 151). On the ground of this distinction, the investigation of capability requires data which can reflect learners' different speech styles when performing multiple tasks.

Tarone's (1983) interlanguage continuum is shown in Figure 1. The vernacular speech style is the most systematic style because it provides learners with minimal opportunities to pay attention to form, whilst the superordinate (i.e. careful) style is the least systematic speech style because it allows learners to pay the most attention to form; this style is therefore permeable to invasion from other rule systems (i.e. target language). Data elicited by various tasks reflect the varying styles that fall between the two poles depending upon the amount of attention paid to form. Thus, attention is considered the cause of style-shifting.

vernacular style (more pidgin-like)	style 2	style 3	style 4	style n	careful style (more TL/NL Like)
unattended speech data	attended speech data	various elicitation tasks: elicited imitation, sentence-combining, etc.			grammatical intuition data
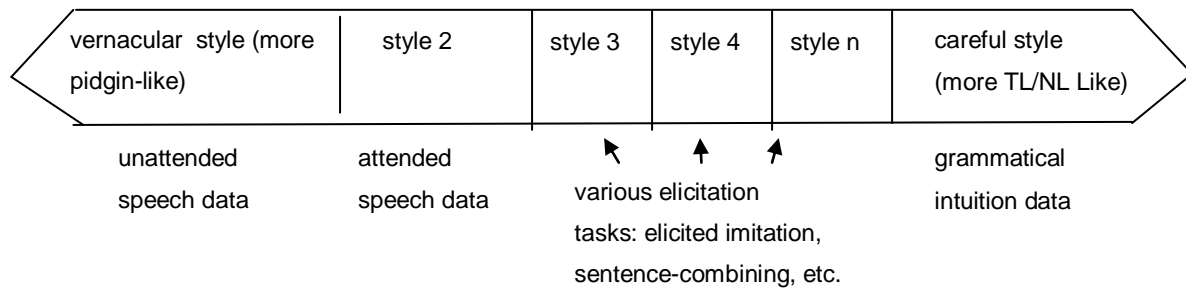
Figure 1: Tarone's interlanguage continuum (1983, p. 152)

Tarone's model suggests that it is necessary to collect data from a range of tasks reflecting different styles to obtain a complete description of learners' capability. For instance, the vernacular style reflected in the observation and analysis of utterances can provide the most accurate evidence of the systematic regularities underlying a learner's interlanguage, but some structures may not be attempted in spontaneous speech. Elicited and intuitional data are easier to obtain, but cannot be used to construct an accurate description of interlanguage alone. Thus, tasks that can be used to elicit data representing the continuum of styles include at least naturalistic speech (either monologue or dialogue), elicitation tasks (e.g., sentence combination, elicited oral imitation), and grammaticality judgement tasks.

Following the Labovian sociolinguistic paradigm, Tarone first attributed style-shifting to the extent to which learners are able to pay attention to linguistic forms. However, Tarone (1985) further demonstrated that it is not appropriate to attribute variation to attention alone as multiple factors are involved. This is considered further later in this chapter.

Although Tarone's style shifting model belongs to a sociolinguistic paradigm, attention, the key variable in this model, is an important construct in cognitive or psycholinguistic theories. As R. Ellis (1994, 2008) puts it, attention 'underlies the variability resulting from performing

different tasks, planning and monitoring'. Therefore, 'a full account of variability in learner language must consider psycholinguistic sources' (p. 150).

2.4.2.3 Psycholinguistic accounts of task effects

The psycholinguistic paradigm views variation in learners' performance in terms of the effect of task conditions on language processing. This section examines variation in terms of the following:

1) learners' attention to linguistic forms (J. H. Hulstijn & W. Hulstijn, 1984; Tarone, 1985).

2) learners' monitoring of their language processing and output (Kormos, 2000; Krashen, 1980, 1981; M. Schmidt, 1980).

3) the distinction between explicit knowledge and implicit knowledge.

4) the Speech Generation Procedures (Levelt, 1989).

5) the Dual-mode System (Skehan, 1998).

6) the Cognition Hypothesis (Robinson, 2001a).

2.4.2.3.1 Attention

Attention is not only a key source of learners' variable performance within the sociolinguistic paradigm (Tarone, 1985, 1988), but it is also widely discussed as a psycholinguistic source of task-induced variation. In psychology, attention refers to the mechanism that controls access to awareness. It is assumed that attention is limited, selective, partially subjective to voluntary control, and essential for action control and learning. The objects of attention and noticing are limited to elements of the surface structure of utterances rather than underlying abstract rules or principles. In the field of research, verbal reports are usually taken as a method of assessing the allocation of attention (R. Schmidt, 2001).

Variability may be induced by task constraints and instructions placed on learners' attention. For example, J. H. Hulstijn and W. Hulstijn (1984) investigated the effects of time pressure, the focus of attention (i.e., whether on information or linguistic form), and metalinguistic knowledge on the accuracy with which two Dutch word order rules (i.e., inversion and verb-end) were used. The analysis of storytelling data showed that attention to form increased accuracy in both structures, but time pressure and metalinguistic knowledge had no effect on their own. It is possible that the different nature of the tasks and task demands can affect learners' performance. As VanPatten (1990) indicates, learners, particularly low proficiency

learners, have difficulty in focusing on meaning and form at the same time due to limited working memory.

2.4.2.3.2 Monitoring

Monitoring occurs when native speakers and learners try to correct mistakes they have made by attending to vocabulary, grammar, phonology, or discourse (R. Ellis, 2008).

Krashen (1981) uses the term 'Monitoring' to refer to the way learners use learned knowledge to edit utterances generated by means of acquired knowledge. He argues that 'successful Monitor users edit their second language output when it does not interfere with communication' and that 'this editing results in variable performance, that is, we see different types and amounts of errors under different conditions. Monitoring generally improves accuracy levels' (pp. 12-13).

M. Schmidt (1980), in a study of second-verb ellipsis in sentences (e.g., 'Mary is eating an apple and Sue a pear'), found that learners all used the second verb in an oral picture description task, but omitted the second verb in proportion to the degree of monitoring that was allowed by the tasks (i.e., elicited imitation, written, sentence-combining and grammaticality judgements). See Section 2.4.4 (Table 7) for a further discussion of this effect.

Kormos (2000) investigated how Hungarian learners monitor their use of English through self-repairs by employing a role-play task and a retrospective interview. The results showed that whereas native speakers focused more on correcting the information content, L2 learners monitored more and focused on correcting linguistic errors. In addition, high-proficiency learners were more likely to focus on appropriateness than low-proficiency learners due to the extra attention available.

2.4.2.3.3 Explicit knowledge versus implicit knowledge

The psycholinguistic paradigm also accounts for task-induced variation in terms of the types of knowledge underlying learners' performance in different tasks. In this paradigm, the knowledge underlying learners' interlanguage is viewed as dichotomous or continuous. Terminologies used to label these dichotomous categories include declarative vs. procedural knowledge (cf., Anderson, 1983, 1990; Færch & Kasper, 1981, 1984, 1986; Paradis, 2009); acquired versus learned knowledge (Krashen, 1982); and implicit versus explicit knowledge

(Bialystok, 1978, 1979). When underlying knowledge is seen as continuous, it is called analysed knowledge and control over the knowledge according to the Analysis-Control model (e.g., Bialystok, 1994; Bialystok & Ryan, 1985). Alternatively, implicit knowledge and explicit knowledge are seen as constituting two poles of a continuum (Færch & Kasper, 1986).

Implicit knowledge			explicit knowledge
learner uses but does not reflect on rule	learner can reflect on rule by relying on intuition	learner can describe rule in own words	learner can describe rule in metalinguistic terms
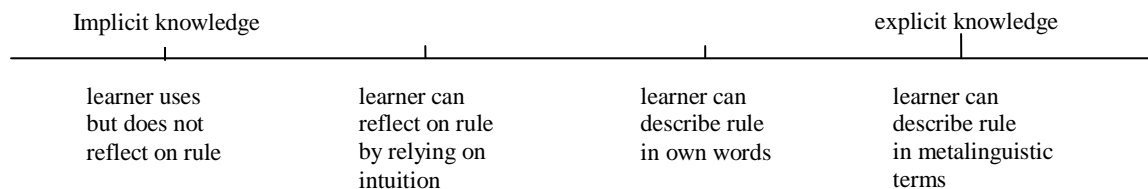
Figure 2: Types of underlying interlanguage knowledge (from Færch, Haastrup, & Philipson, 1984, p. 202)

Færch and Kasper's (1986) illustration of the continuum of types of interlanguage knowledge is shown in Figure 2. Implicit knowledge is the 'knowledge which underlies language use only'. In moving towards the right, learners gradually utilise more explicit types of knowledge. 'The right-most type of explicit knowledge is intimately related to literacy and schooling, as its presence requires the existence of a meta-language in the learners' mind.' (pp. 212-213)

Regarding the interface between implicit knowledge and explicit knowledge, there are generally three positions in SLA (R. Ellis, 2005a, 2008a).
1) The non-interface position (Krashen, 1981) claims that implicit knowledge and explicit knowledge are processed (i.e. acquired or learned) in distinct mental mechanisms, stored separately.
2) The strong interface position (DeKeyser, 1998) advocates that L2 knowledge can be converted into implicit knowledge through communicative practice.
3) The weak interface position (R. Ellis, 1993) posits that explicit knowledge functions as a facilitator of the processes involved in acquiring implicit knowledge.

As R. Ellis (2005a, 2008b) stresses, it is necessary to obtain separate measures of the two types of knowledge to examine these positions. A fundamental hypothesis underlying them is that different tasks processing activate different types of knowledge, and a series of studies have attempted to investigate this. For example, Y. Han and R. Ellis (1998) and R. Ellis

(2005a) conducted psychometric studies to examine separate measures of the two types of knowledge (see Section 2.4.3 for a detailed review). R. Ellis (2005a) claims that an oral imitation test taps into implicit knowledge, while an untimed grammaticality judgement test taps into explicit knowledge. Erlam (2006) also accepts that an elicited imitation task can serve as a measure of learners' implicit knowledge. Furthermore, a number of studies suggest that a grammaticality judgement test is more likely to measure implicit knowledge in judgement of correct sentences but explicit knowledge in judgement of incorrect sentences (Bialystok, 1979; R. Ellis, 2005a, 2008b, 2009a; R. Ellis et al., 2009; Loewen, 2009).

Based on these findings, it can be inferred that different tasks draw on the different types of knowledge.

2.4.2.3.4 Levelt's Speech Generation Procedures

Levelt (1989) proposed a psycholinguistic model to explain speech processing. This model informs theoretical models in SLA such as Pienemann's Processability Theory and pre-task and within-task planning (R. Ellis & F. Yuan, 2004; F. Yuan & R. Ellis, 2003).
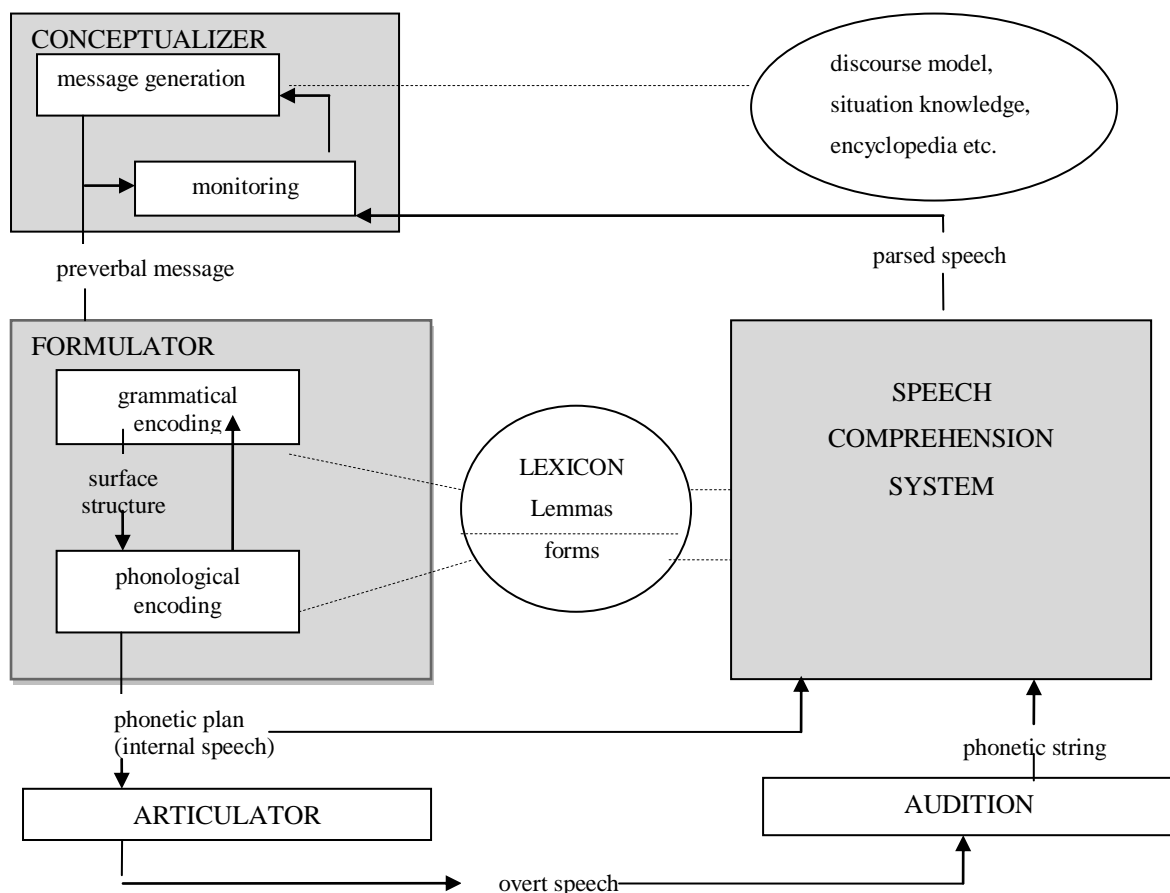
Figure 3: Levelt's model of language generation (1989, p.9)

31

Levelt's (1989) model of language generation (Figure 3) is concerned with both comprehension and production, but the focus in this discussion will be on production. The model views the process of speech generation in three phases: conceptualisation, formulation, and articulation. Conceptualisation involves the macro-planning of speech. The conceptualiser draws on general knowledge and discourse knowledge to help construct message content. The formulator, drawing on lexical, grammar, and syllabary stores, expresses the intended message through accessing, grouping and ordering linguistic features. This process involves three types of encoding rules: grammatical encoding, morpho-phonological encoding, and phonetic encoding. By the end of this phase, a phonological plan is generated. In the final phase, the physically produced strings are assembled by selecting patterns of stress, rhythm and intonation. The resulting articulation is fed into the auditory mechanisms.

This model also stipulates that the whole process of speech generation is supervised by a monitor and completed under time pressure. Because this process is sensitive to the amount of time available for learners to access their linguistic resources, variation may occur in learners' oral production under different task conditions.

2.4.2.3.5 Skehan's (1998) Dual-mode System

Skehan (1998) developed a dual-mode system (i.e., a rule-based system and an exemplar-based system) to account for the cognitive mechanisms underlying learning and performance. The rule-based system is parsimonious, 'elegantly organised' and 'generative', containing rules that are 'compactly structured', 'creative in their application', and 'precise in the meanings that they can express'. This system draws 'in turn upon lexical elements (themselves well organised in a lexicon)' (p. 88). In contrast, the exemplar-based system 'lacks parsimony and has only a limited generative potential'. It is 'heavily based on the operation of a redundant memory system in which there are multiple representations of the same lexical elements' (p. 89). The rule-based system 'leads to the development of an open, form-oriented system, while the exemplar-based system emphasizes meaning and is less appropriate for underlying system change' (p. 89).

Skehan (1998) further argues that the two systems are not ideally separate but work harmoniously in language organisation procedures (i.e., Lexicalisation > Syntacticalisation > Relexicalisation). These procedures can be mobilised as appropriate for different communicative contexts and goals. Whether particular L2 learners access their representation of rules or exemplars or both depends upon a range of factors such as the context of learning, the nature of the instruction and individual differences. 'When time is pressing and the contextual support high, memory-based communication is appropriate. When there is more time, and precision is important, the rule-based system can be accessed.' (pp. 90-91)

Skehan claims that tasks need to be analysed in terms of their design features and implementation conditions. Further, he proposes that there is competition between the three aspects of performance (i.e. fluency, accuracy, and complexity) (Skehan & Foster, 1997). Certain task qualities and conditions lead to an emphasis on one or two of them, and so cause variation in learners' task performance. Skehan (1998) believes that it is not only important to understand the separate factors that influence task performance, but also to examine how the factors interact. Skehan's model of oral test performance, which centres around 'task' is shown in Figure 4. This shows the various factors that impact on how a task is performed such as task qualities and task conditions.
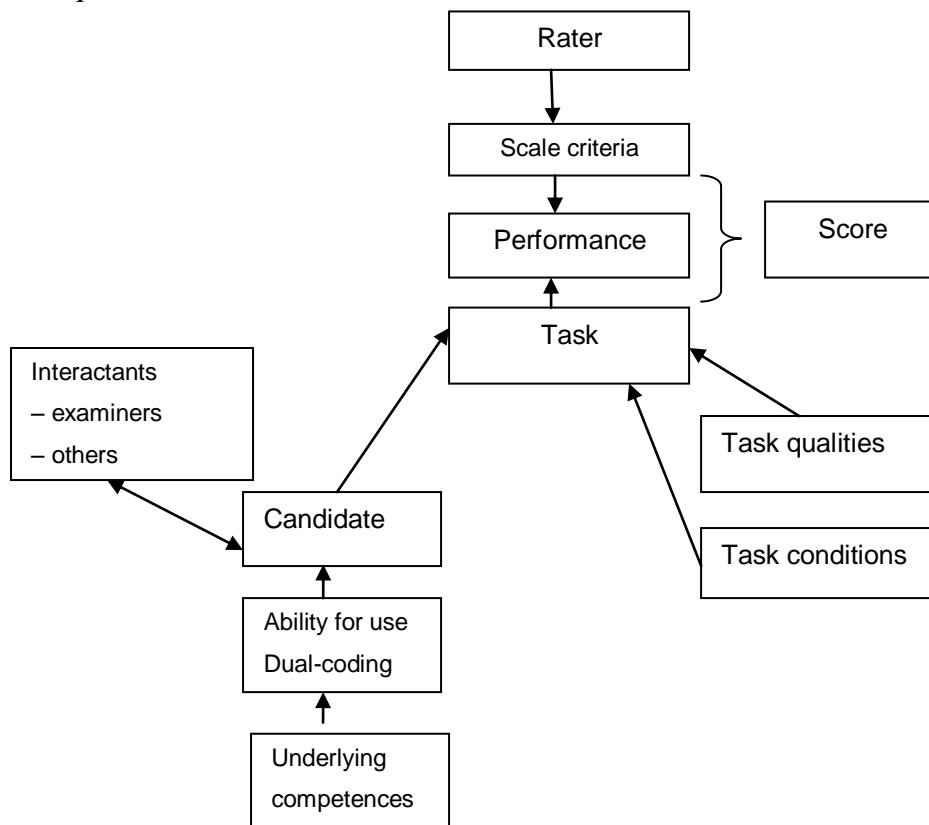
Figure 4: Skehan's model of oral test performance (1998, p. 172)

2.4.2.3.6 Robinson's (2001a) Cognition Hypothesis

Robinson proposed the Cognition Hypothesis (Robinson, 2001a, 2002, 2003, 2005; Robinson & Ellis, 2008; Robinson & Gilabert, 2007) based on work in functional/cognitive linguistics (Talmy, 2000a, 2000b; Tomasello, 2003) and in developmental psychology. This hypothesis claims that L2 tasks should be sequenced for learners on the basis of cognitive complexity rather than linguistic complexity. It predicts that increasing the cognitive demands of tasks can 'push learners to greater accuracy and complexity of L2 production to meet the consequently greater functional/communicative demands' placed on them, and 'promote heightened attention to and memory for input' as well as 'longer term retention of input'. In addition, 'performing simple to complex sequences will also lead to automaticity and efficient scheduling of the components of complex L2 task performance' (http://www.cl.aoyama.ac.jp/~peterr/hf/index.html).

Robinson (2001b) also distinguished three groups of factors: *task complexity*, *task conditions*, and *task difficulty*, which interact to influence task performance and learning. This triadic framework is shown in Figure 5.

Task complexity (cognitive factors)	Task conditions (interactive factors)	Task difficulty (learner factors)
a) resource-direction e.g., +/- few elements +/-here-and-now -/+ no reasoning demands	a) participation variables e.g., one-way/two-way convergent/ divergent open/closed	a) affective variables e.g., motivation anxiety confidence
b) resource-depleting e.g., +/- planning +/- single task -/+ prior knowledge	b) participant variables e.g., gender familiarity power/solidarity	b) ability variable e.g., aptitude proficiency intelligence
Sequencing criteria Prospective decisions about task units		Methodological criteria On-line decisions about pairs and groups

Figure 5: Robinson's triadic componential framework (2001b, p. 294)

Task complexity is concerned with the increase of demands on learners' cognitive resources along two dimensions: resource-directing and resource-depleting. Resource-directing

dimensions direct learners' resources to aspects of the language code that can be utilised in completing the task (e.g., using present or past tense forms to describe things happening now versus in the past). Resource-depleting dimensions make extra resource demands (relative to tasks simpler along these dimensions) which cannot be met through the use of any particular features of the language code. A task can be made more complex by removing prior knowledge support, or by making it a dual rather than a single task. Each of these dimensions consists of 'design features of tasks and their implementation which can be manipulated to increase or lessen the cognitive demands tasks make on the learner during task performance' (p. 294). These features can be thought of as on a continuum, along which relatively more of a feature is present or absent.

Task conditions concern the nature of the participation required for tasks (e.g., whether the information is one-way or two-way) and participant variables (e.g., whether participants are the same/different gender). Robinson suggests that these factors are unlikely to be a useful basis for a priori sequencing decision.

Task difficulty concerns learners' perceptions of the demands of the task. It is determined by both affective variables (e.g., motivation to complete the task) and ability factors (e.g., aptitude).

Robinson points out that task complexity should 'help explain intra-learner variation in the performance on any two tasks', and that task difficulty should help explain variation in task performance between any two learners performing the same task' (p. 295), though affective variables cannot inform a *priori* decisions about task sequencing because of the difficulty of diagnosing in advance learners' engagement with the task.

Although Robinson's Cognition Hypothesis was intended to provide a set of sequencing criteria for classroom tasks, its explanatory power is not limited to this. The triadic framework can also be used to explain task-induced variation in a general sense.

2.4.3 Types of tasks

In this section I will focus on the two types of tasks most relevant to the current study: clinically elicited oral production tasks and metalinguistic tasks.

2.4.3.1 Clinically elicited oral production tasks

Tasks used to elicit clinical oral production include focused interviews, picture description or narrative tasks (e.g., R. Ellis, 2005a; J. White & Ranta, 2002), video description tasks (e.g., Bardovi-Harlig, 2000; Chafe, 1980; Tarone, 1985) and 'spot-the-difference' tasks (e.g., Charters, 2005; Loschky & Bley-Vroman, 1993). Clinically elicited tasks are of particular importance in SLA because they not only provide data 'which reflect an essential quality of naturally-occurring data', but also elicit structures that rarely occur in a naturally occurring situation and thus make the process of data collection 'practical and less arduous' (R. Ellis & Barkhuizen, 2005, p. 36).

2.4.3.2 Metalinguistic knowledge tasks

Metalinguistic knowledge tasks are tasks where the nature of the task invites or encourages learners to access their analysed knowledge of the L2. Variants of this group of tasks widely used in SLA include preference decision tasks, acceptability or grammaticality judgement or rating tasks, error identifying or correcting tasks, and rule explanation tasks (see, Chaudron, 2003; R. Ellis & Barkhuizen, 2005; Loewen, 2009). Some researchers argue that metalinguistic tasks can be viewed as falling on a continuum of increasing use of 'analyzed knowledge' (e.g. judging grammaticality, locating errors, correcting errors, and explaining ungrammaticality (Bialystok & Ryan, 1985). In other words, metalinguistic ability constitutes a continuum of abilities ranging from 'sporadic insights into aspects of language' at one end, to the ability to generate the elaborate metalinguistic explanations that linguists enjoy at the other (Sharwood Smith, 1991, p. 20; cited from J. White & Ranta, 2002, p. 261).

The grammaticality judgement task (GJT) is perhaps the most widely used metalinguistic knowledge task (Bialystok, 1979; Chaudron, 2003; R. Ellis, 1991; Leow, 1996; Loewen, 2009; Sorace, 1985) . This is partly because GJTs can be used to examine acquisition of target features that learners seldom use in naturally occurring conversations. However, there is disagreement between researchers regarding what GJTs measure. Some researchers argue

that GJT data measure learners' intuitions of the target language (Corder, 1981; Kellerman, 1978; Tarone, 1985) or learners' linguistic knowledge (Birdsong, 1989; Chaudron, 1983; Davies & Kaplan, 1998; Gass, 1994; Hedgcock, 1993). Others claim that GJT data may only tap learners' analysed knowledge (Bialystok & Ryan, 1985; R. Ellis, 1991; Selinker, 1972). Furthermore, while some researchers believe that GJTs are a reliable method to collect data about learners' L2 performance (see Birdsong, 1989; Chaudron, 1983; R. Ellis, 2008a, 2008c; Gass, 1994; Kellerman, 1986; Leow, 1996; Loewen, 2009; Sorace, 1996) , others challenge the validity of GJTs on the grounds that learners rely on translation or explicit knowledge rather than implicit knowledge when performing GJTs (Birdsong, 1989; Christie & Lantolf, 1992; R. Ellis, 1989, 1991; Goss, Zhang, & Lantolf, 1994).

Differences in the results obtained by studies using GJTs can be attributed to the following factors: 1) proficiency of learners (R. Ellis, 1991, 1994); 2) nature of the target feature (Goss, et al., 1994); 3) task condition (timed/untimed) (R. Ellis & Yuan, 2004); and 4) grammaticality of stimuli (grammatical/ ungrammatical) (R. Ellis, 1991; R. Ellis & Yuan, 2004; Loewen, 2009). As Loewen (2009) concludes,

> [f]eatures of GJTs may be manipulated to predispose L2 learners to draw on different types of L2 knowledge. GJTs with limited response times seem to limit the ability of L2 learners to access their explicit knowledge in making a judgement, while ungrammatical sentences on an untimed test appear to encourage learners to access explicit L2 knowledge (p. 111).

2.4.3.3 Oral production versus metalinguistic task performance

Like metalinguistic knowledge tasks, oral production tasks are widely used in SLA. However, whether the two types of tasks afford the same or different data is controversial. Some studies have demonstrated that learners' metalinguistic judgements match their use of the L2 in oral production tasks (Arthur, 1980), while others have found the opposite (R. Ellis & Rathbone, 1987; Liceras, 1983), particularly when the grammaticality judgement task was operationalised by asking learners to verbalise grammatical rules (Green & Hecht, 1992; J. H. Hulstijn & W. Hulstijn, 1984; Sorace, 1985). A theoretical explanation for differences in learners' performance of the two types of tasks is that they make differing demands on two qualitatively different kinds of knowledge (Lightbown, 1985). The two types of knowledge

are viewed as either independent (Krashen, 1978) or interdependent by different theorists (Bialystok, 1978). Further, different labels are used to describe the knowledge underlying the two types of tasks. A summary of these labels (cited from J. White & Ranta, 2002, p. 260) is shown in Table 6.

Table 6: Distinctions between Two Kinds of Linguistic Knowledge in SLA (J. White & Ranta, 2002, p. 260)[2]

Knowledge underlying production/comprehension	Knowledge underlying metalinguistic task performance	Author
Acquisition	Learning	Krashen (1978)
Implicit	Explicit	Bialystok (1978)
Automatic processing	Controlled processing	McLaughlin (1978)
Communicative system	Cognitive system	Lamendella (1977)
Language specific structure	Problem solving structure	Felix (1981)
Basic interpersonal communication skills (BICS)	Cognitive academic language proficiency (CALP)	Cummins (1980)
Conversational proficiency	Academic proficiency	Cummins (1991)
Procedural	Declarative	Anderson (1983)
Submeta (default) mode	Meta mode	Sharwood Smith (1993)

A number of studies have explored the underlying knowledge that these two types of tasks elicit (Bialystok, 1982; J. White & Ranta, 2002). However, there has been no such study of this with learners of L2 Chinese. This is one of the motivations of the current study.

2.4.4 Studies reporting on variation according to task type

Task-related variation (Tarone, 1985, 1988), or task-induced variation (R. Ellis, 1992, 1994, 2008c), has generally been acknowledged as one of the crucial types of intra-learner variation in SLA. A large number of studies have examined a variety of types of tasks and their features. However, the following review will be limited to the studies that have involved both oral production tasks and metalinguistic judgement tasks, as these tasks are the focus of the current thesis. Ten relevant studies are summarised in Table 7.

---

[2] Disputably, this table could be labeled as 'differences in the constructs measured by oral production and metalinguistic tasks' because 'automatic processing vs. controlled processing' and 'conversational proficiency vs. academic proficiency' should not be considered as linguistic knowledge underlying the two types of tasks.

Table 7: Studies That Have Examined Task Effects Based on Oral Production and Metalinguistic Tasks

Study	Participants	Target feature	Tasks	Main findings
M. Schmidt (1980)	9 subjects in 5 L1s (i.e., Japanese, Chinese, Finnish, German, and Arabic)	Second verb ellipsis in English coordinate clauses	1) oral picture description; 2) elicited imitation; 3) sentence combination; 4) grammaticality judgement	The second verb was used in free speech but deleted increasingly on tasks (supposedly) allowing more and more attention to form.
Tarone (1985)	20 ESL learners Japanese (n=10) Arabian (n=10)	4 English grammatical structures: third person singular present tense verb -*s*; the article; the noun plural -*s*; and third person singular direct object pronounces	1) written grammaticality judgement task; 2) oral narrative task; 3) oral interview	Learners perform differently in grammaticality judgement tasks and in oral production tasks. In some cases, spontaneous oral accuracy scores are better than grammar judgement scores. The multiple tasks can be ordered in terms of degree of attention to form. Learners' styles in response to these tasks may be ranged along a continuous dimension.
Tarone & Parrish (1988)	20 ESL learners Japanese (n=10) Arabian (n=10)	English article associated with noun phrases	1) written grammaticality judgement task; 2) oral narration task; 3) oral interview	Task-related variability in interlanguage is due to a complex range of variables including the communicative demands and discourse characteristics of the tasks, and not only 'attention to form'.

J.H. Hyltenstam (1983)	33 adult SL learners of Swedish with various L1s (e.g. Polish)	2 Swedish syntactic features (i.e., pronominal copies in relative clauses, and sentence negation)	1) elicited written production; 2) elicited oral production (picture identification task); 3) imitation; 4) grammaticality judgement task (spoken & written)	In the case of relative clause, regular pattern presented on the elicited oral production but not on oral and written grammaticality judgement tasks. In the case of negation, a similar patterning presented on different tasks but lack of comparability. Different tasks may be appropriate for different linguistic phenomena and for learners at different phases of acquisition.
Sorace (1985)	17 English speaking college students of Italian in UK; Beginners (n=9) Intermediate group (n=8)	6 grammatical structures of Italian (perfect, imperfect, use of indirect pronoun with 'piacere', use of indirect pronoun with 'dire'; choice of auxiliary in compound tenses; past participle agreement in compound tenses)	1) written judgement test (including judging, making correction, and stating grammar rules); 2) oral picture description task; 3) informal conversation with the interviewer	There was a significant task effect among the tasks. Picture Description Task was easier than the Conversation for both groups. No difference was found between the two groups on the Judgement Test.
Leow (1996)	30 university students of Spanish in UK	Agreement in Spanish between nouns/ non phrases and adjectives or past particles (functioning like adjective in Spanish)	1) grammaticality judgement task (including judging, making correction, and stating grammar rules); 2) oral production task; 3) written production task	Learners' grammaticality judgement scores had a substantially stronger relationship with written production performance than with oral production for the participants at two stages

Y. Han & R. Ellis (1998)	48 ESL university students in the US	English verb complement	1) timed oral production test; 2) timed grammaticality judgement test; 3) delayed grammaticality judgement test; 4) an interview	Delayed grammaticality judgement test loaded on one factor, other tests loaded on the other. So the two factors were interpreted as explicit knowledge and implicit knowledge.
R. Ellis (2006)	220 Learners in mixed L1s (N1= 147 ESL learners in NZ; N2= 28 Japanese learners in Japan; N3= 54 TESOL learners in Malaysia)	17 English grammatical structures	1) oral imitation test involving grammatical and ungrammatical sentences; 2) oral narration test; 3) timed grammaticality judgement test; 4) untimed GJT with the same content; 5) metalinguistic knowledge test	Two oral tests and untimed GJT loaded on one factor, and the UGJT and MKT loaded on the other. The findings suggest that tests could be designed to provide relatively separate measure of implicit and explicit knowledge.
R. Ellis (2009b)	111 (91 ESL learners, 20 native speakers)	17 English grammatical structures	1) oral imitation test involving grammatical and ungrammatical sentences; 2) oral narration test; 3) timed grammaticality judgement test; 4) untimed GJT with the same content; 5) metalinguistic knowledge test	Two oral tests and untimed GJT loaded on one factor, and the UGJT and MKT loaded on the other. The findings suggest that tests could be designed to provide relatively separate measure of implicit and explicit knowledge.

Note. ESL= English as a second language learners; GJT= grammaticality judgment test; MKT= metalinguistic knowledge test; TESOL= Teaching Chinese to Speakers of Other Languages; UGJT= untimed grammaticality judgment test

All these studies reported a significant variation in output according to the tasks used, although the explanations for this variation vary. The main findings of these studies are synthesised below.

2.4.4.1 Monitoring or attention

Three studies (M. Schmidt, 1980; Tarone, 1985; Tarone & Parrish, 1988) accounted for task variation in terms of the psycholinguistic processing required by the tasks. M. Schmidt (1980), for instance, argued that the subjects' variable use of ellipsis in the second verb in coordinate constructions was caused by the task conditions which allowed for more or less attention to form and more or less opportunity for monitoring. M. Schmidt suggests that the learners may have two different competencies: a receptive competence which allows them to comprehend the meaning of a deleted verb structure and a productive competence which allows them to produce such an utterance.

Tarone (1985) claims that the learners' performance on different tasks may form a style-shifting continuum, ranging from a vernacular style to a careful style depending upon the extent to which learners are able to pay attention to language form. She speculates that the style-shifting may also be caused by the nature of the discourse which the tasks require.

2.4.4.2 Multiple sources of variation

Two of the studies attributed task-induced variation to multiple sources. Tarone and Parrish (1988), after analysing the relationship between the form and function of English articles in the discourse of three tasks, argue that task-induced variability may be due to a complex of variables, including the communicative demands and discourse characteristics of the tasks, and not only 'attention to form'. Hyltenstam (1983) found different patterning for different linguistic structures based on data from different tasks. He suggested that task-induced variability should be considered normal and that may be attributed to multiple factors including linguistic phenomena and individual variables such as learners' level of proficiency.

2.4.4.3 Knowledge versus use of language

These studies examined task variation in terms of the relationship between learners' metalinguistic knowledge and their use of the language. Sorace (1985) found that the

correlation between metalinguistic knowledge scores and oral production scores was significant, positive and strong for non-beginners but not significant, negative and weak for beginners. Her statistical analysis suggests that the picture description task she used was easier than the conversation task for both groups of subjects, but, unexpectedly, the judgement test was not easier than the production tasks. A definite developmental pattern for metalinguistic knowledge was revealed, but it is hard to explain the increasing relationship between the metalinguistic knowledge of the two groups of learners and their productive use of the target language. Sorace suggested a further investigation of the psycholinguistic processes underlying the relationship.

Leow's study (1996) demonstrated significant relationships between the learners' GJTs and their performance on both oral and written production tasks, with a stronger correlation with the written production. He claims that grammaticality judgements can be used to predict production and provide a reliable measure of learners' performance. In particular, the modality of production (oral vs. written) needs to be considered when examining the relationship between GJTs and production.

2.4.4.4 Measurement of underlying knowledge

The findings of a number of studies that oral production tasks and untimed metalinguistic knowledge tasks can be used to provide relatively separate measures of implicit and explicit knowledge. Y. Han and R. Ellis (1998), found that learners' scores on oral production tasks and the untimed GJT loaded on two different factors, while those on the timed GJT loaded on the same factor as the oral tasks. They interpreted these two factors as implicit knowledge and explicit knowledge, claiming that this finding lends support to Krashen's (1981) non-interface position and R. Ellis's (1994) claims that implicit knowledge and explicit knowledge are separate. R. Ellis's (R. Ellis, 2005a, 2006, 2008b, 2009a,b) seminal studies of 17 English grammatical structures with large samples of learners also yielded the same result. He explained that timed GJTs and oral production tasks were more likely to make learners tap implicit knowledge while untimed GJTs was more likely to allow learners to access their explicit knowledge.

2.4.5 Relevance to the current study

The study of task effects is of central importance in SLA. Methodologically, tasks serve as key devices for collecting learner data in SLA research and thus it is crucial to address their

construct validity. Also, the study of tasks contributes to theory development. The theoretical explanations and empirical exploration of the task effect elaborated above provide a framework for designing tasks and interpreting the data they provide in the current study. For example, my decision to use multiple tasks to collect data draws on the key finding that different tasks measure different underlying language competencies.

2.5 Conclusion

This chapter has provided the theoretical background and empirical grounds for investigating two sources of intra-learner variation in interlanguage. With respect to linguistic difficulty, four types of objective criteria that predict grammatical difficulty were considered by reviewing relevant empirical studies that have examined subjective learning difficulty. With respect to task effects, the definition and different types of tasks were discussed, various theoretical perspectives on task effects were illustrated, and relevant empirical studies that have attested to differences in oral production and metalinguistic knowledge tasks were reviewed. The next chapter turns to a consideration of the factors that account for inter-learner variation in L2 performance.

CHAPTER 3: SOURCES OF INTER-LEARNER VARIATION

3.1 Introduction

Chapter 2 provided a review of sources of intra-learner variation in L2 performance. This chapter considers the factors that account for inter-learner variation. Theoretical perspectives on six potential sources (i.e., starting age, the number of years of study, setting, native language (L1), self-rated proficiency, and gender) are documented, and empirical evidence for the effects of these sources on the accuracy of use of grammatical structures in oral production and on metalinguistic knowledge are presented.

3.2 Age-related factors

This study focuses principally on two age-related factors: starting age and the number of years of study. This decision was made because the investigation of starting age requires a consideration of the number of years that learners have been learning the L2. The effect of biological age is not a focus of this study because the majority of learners in this study were adult learners at a lower level of Chinese proficiency. This study of the variable use of *BC*, therefore, cannot provide a test of the critical period.

3.2.1 Critical Period Hypothesis

The role of age is one of the most interesting and perennial issues in SLA. L2 researchers' interest in age effects have been motivated largely by theoretical hypotheses and relevant empirical findings in L1 acquisition (Lenneberg, 1967; Penfield & Roberts, 1959). In particular, Lenneberg's (1967) Critical Period Hypothesis (CPH) has become contentious. This hypothesis predicts that completely successful acquisition of a language must occur before the age of puberty (around 10-12), before cerebral lateralisation is complete.

The findings of empirical studies that have tested the CPH continue to be controversial. Some researchers found support for a critical period, arguing that there is a critical period after which most people cannot achieve native-like proficiency in the target language (DeKeyser, 1990; Johnson & Newport, 1989, 1991). Others doubt the existence of such a critical period, finding no discontinuity between starting age and developing proficiency (Bialystok & Hakuta, 1994, 1999; Birdsong, 1992; Hakuta, Bialystok, & Wiley, 2003).

45

The key issue regarding the CPH is whether it is possible to identify a cut-off age, and, if so, what this cut-off age is. The expectation in the CPH is that, those learners start learning prior to the cut-off age will perform in markedly different ways from those who start after it. Researchers have claimed to have identified different cut-off ages depending on the aspect of the target language being considered. For instance, 6-8 years (e.g., Walsh & Diller, 1981), 13 years (e.g., Curtiss, 1977) and 10-12 years (e.g., Lenneberg, 1967) have been considered critical ages for mastering specific linguistic categories of a first language. Age 6 (Long, 1990), age 12 (Flege, Munro, & MacKay, 1995) and age 13 (Scovel, 1988) are considered crucial for achieving native-like L2 phonology. Age 7 (Hyltenstam, 1992) and age 15 (DeKeyser, 2000; Johnson & Newport, 1989; Long, 1990; Patkowski, 1980) have been considered as cut-off ages for achieving native levels of L2 morphology and syntax. Due to the complexity of the age issue, some L2 researchers simply prefer to talk about a *sensitive period* (e.g., Bornstein, 1989; Colombo, 1982; DeKeyser, 2000; Oyama, 1976, 1978, 1979; Patkowski, 1980) or an *optimal age* (Cummins, 1980; Patkowski, 1994) instead of a *critical period.*

Other researchers argue, however, that it is impossible to establish a clear cut-off point (Bialystok & Hakuta, 1994, 1999; Birdsong, 2006). The evidence is that the slope of the linear relationship between proficiency and starting age is a gradual one rather than an abrupt degradation after a certain age. Furthermore, Bialystok and Hakuta (1994), after re-evaluating Johnson and Newport's (1989) data, demonstrated that if the cut-off age for the end of the critical period was moved to 20 years, then the age of the older group in Johnson and Newport's study was found also to be correlated with performance (cited in R. Ellis, 2008). Further evidence for rejecting the CPH can be found in studies that show it was not impossible for adult learners to achieve a native-like proficiency (Birdsong, 1992; L. White & Genesee, 1996). It seems that this position is now strongly favoured in the literature (R. Ellis, 2008c; Kellerman, 1995).

Another issue related to the CPH involves examining whether younger L2 starters have superior L2 skills than older starters. On the one hand, some studies demonstrate that younger beginners are generally more successful than those beginning as adults both in the acquisition of morphosyntactic structures in a second language setting (Krashen, Long, & Scarcella, 1979) and in the acquisition of vocabulary in a foreign language instructional setting (Yamada, Takatsuka, Kotabe, & Kuruse, 1980). Counter-evidence remains (Singleton &

Ryan, 2004), however. Evidence exists that older children have superior L2 skills than younger children (Ekstrand, 1976; Fathman, 1975; Snow & Hoefnagel-Höhle, 1978), and that adults excel child learners in the rate of acquisition during the initial stages of L2 acquisition (Krashen, et al., 1979). Other studies limit the explanation of age effects to certain aspects of L2 acquisition, claiming that younger learners outperform older ones as a result of their biologically endowed capacity when learning phonetic or phonological features (Flege, et al., 1995). However, a consensus has been achieved to some extent – the earlier the exposure to a target language begins, the higher the level of L2 proficiency that can be achieved (Cook, 1991; Harley, 1985; Long, 1990, 2005; Singleton, 2005; Singleton & Ryan, 2004), but only providing there is ample exposure to the target language.

Because in my study most of the learners began learning Chinese after age 15, I will not be able to directly test the prediction of the CPH. However, I will focus on whether early starters demonstrate a higher level of acquisition of *BC* than late starters. Thus, the following review will concentrate on the empirical studies that have examined the effects of starting age on the acquisition of grammatical features.

3.2.2 Starting age

Starting age refers to the age at which a learner starts learning an L2. Several variants of the term 'starting age' have appeared in the literature such as *age of arrival (AoA), age of exposure, age of initial learning*, *age of onset, age of learning,* and *age of immersion* (Abrahamsson & Hyltenstam, 2009; Leather, 2003). The findings in the literature on the effects of starting age will be introduced separately in terms of second and foreign language contexts.

3.2.2.1 Research investigating the effect of starting age on L2 proficiency of learners in second language contexts

The investigation of the effect of starting age on the acquisition of grammatical features has generally focused on the following questions: 1) If there is an age effect, is the relationship between performance and age discrete or continuous? 2) If there is age effect, is it possible for learners who start learning an L2 after the presumed critical period to achieve native proficiency? 3) Do early starters outperform late starters? 4) Is there evidence for no effect of

starting age? My study is related to questions 1), 3), and 4), so the following review will focus on answering these questions.

Answers for the first question are mixed. A number of studies have showed that there is a determining point for the acquisition of L2 grammar, and that learners who started learning prior to this period outperformed those who started after it. One of the most cited studies is Johnson and Newport's (1989) investigation of 12 morphological and syntactic rules in English. Forty-six Korean and Chinese native speakers who had arrived in the United States between the ages of 3-39 years and had lived there for between 3 and 26 years were asked to judge the grammaticality of 276 spoken sentences. Statistical analyses demonstrated a stronger negative correlation between starting age and test scores for the earlier arrivals (AoA $\leq$ 15) ($r = -.87$, $p<.01$) than for the later arrivals (AoA $\geq$17) ($r = -.16$, $p > .05$). This result was interpreted as strong support for the existence of a critical period, but received criticism from Bialystok and Hakuta (1994) and Kellerman (1995) in that the starting ages of the subjects were grouped arbitrarily and the same age effects would be observed if an older cut-off age were selected.

Johnson and Newport's (1989) study has been replicated in a number of ways. Johnson (1992), for example, adopted both an auditory and written version of the grammaticality judgement test. A strong negative correlation between age of arrival and test performance for subjects of all ages of arrival ($r = -.54$) and for the earlier arrivals (AoA $\leq$ 15) ($r = -.73$) was observed. However, performance was higher on the written version than on the auditory version only for the older arrivals. DeKeyser's (2000) replication with 57 Hungarian-speaking immigrants found a significant negative correlation between age of arrival and grammaticality judgments of 200 sentences ($r = -.61$). Very few adult immigrants (age of arrival >16) scored within the range of the child arrivals (age of arrival <16) in their judgments of morpho-syntactic rules in English. Interestingly, length of residence which was related to age of arrival was not correlated with test scores. McDonald's (2000, 2006) replication of this study with Spanish early acquirers (age of arrival $\leq$ 5) (n = 14) and late acquirers of English (age of arrival $\geq$ 14) (n = 14) as well as Vietnamese early acquirers (n = 14) and child acquirers of English (age of arrival = 6-8) also revealed a significant negative correlation between age of arrival and judgment scores. However, the performance patterns of the two L1 groups differed. The L2 levels of Spanish early acquirers were not distinguishable from native English speakers, while Spanish late acquirers had difficulty with

all aspects of the grammar tested except word order. Vietnamese early acquirers had difficulty with those aspects of English that differ markedly from Vietnamese. McDonald (2006) further attributed the poor grammaticality judgments of late L2 learners (age of arrival ≥ 12) to processing difficulties caused by low L2 working memory capacity, poor L2 decoding, and/or inadequate L2 processing speed.

Whereas the above findings attest to the CPH, other studies cast doubt on the CPH by showing that age affected performance in a continuous way and so no critical period could be identified. Bialystok and Miller's (1999) study with three groups of participants (i.e., 33 Chinese, 28 Spanish and 38 native speakers of English) found that starting age only affected Spanish learners' judgements on five English grammar structures presented in both an oral and written form, and influenced the proficiency achieved through all ages tested. Flege, Yeni-Komshian, and Liu (1999) examined the acquisition of English phonological and morpho-syntactic properties by 240 Korean learners who had arrived in the United States between the ages of 1 and 23. They found that the correlation between age of arrival and level of proficiency was both significant and continued even after the age of 12 and that a few learners who had started learning after puberty scored in the range of native speakers. The rate of accuracy in the acquisition of L2 phonological and morphosyntactic properties declined with age according to age of arrival. The above evidence leads some researchers to cast doubt on the CPH.

With respect to the third question about whether early starters are superior to late starters, findings are also mixed. Some studies found that younger starters perform better than older starters. Oyama (1978), for example, examined 60 Italian learners aged 6-20 years in the U.S. for their ability to imitate English sentences. They reported that starting age in the U.S. explained the results but that other potential sources (e.g., years in the U.S., attitudinal variables including motivation, self-consciousness about speaking and cultural identification) did not (cited in Johnson & Newport, 1989; Krashen, *et al.,* 1979). Patkowski (1980) investigated 67 immigrants who had arrived in the U.S. between 5 and 50 years and resided there for between 6 and 61 years. It was demonstrated that the prepubescent learners (AoA $\leq$ 15) outperformed their postpubescent counterparts (AoA $\geq$ 15). Starting age was a strong predictor of syntactic proficiency in English based on oral interview data, while other independent variables such as *years in the target language environment*, *informal exposure to English*, and *formal instruction in English* had little effect.

49

In contrast, some studies demonstrated that older children outperformed younger ones. Ekstrand (1976), for example, after examining 2189 learners of Swedish aged 8-17 years who had resided in the second language environment for up to 2 years, reported that older children outperformed their younger counterparts in listening comprehension, reading, free writing, pronunciation and speaking. Fathman (1975) investigated the performance of 200 learners of English aged 6-15 years who had resided between 1 and 3 years using the SLOPE test and a picture description task. The results showed that the 11-15 year olds were superior to the 6-10 year olds for morphology and syntax, but that the 6-10 year olds were better at pronunciation. Snow and Hoefnagel-Höhle (1978) investigated how approximately 90 learners of Dutch aged 3-15 who had lived in the target language environment for between 1 month and 1 year performed in pronunciation, morphology, imitation, and translation tasks. The results showed that the 12-15 year olds were best in morphology and syntax, with 8-10 year olds next best, but that the differences diminished over time (they were strongest at 1-3 months).

Moreover, evidence shows that the hypothesis that younger starters are better than those who start at older ages may be influenced by other factors such as the modality of tasks. Montrul, Foote, and Perpiñán (2008), for instance, investigated the acquisition of gender agreement in Spanish by L2 learners (n = 72) and heritage speakers (n = 69) in tasks involving different modalities (i.e., oral production, written comprehension, and written recognition), with 22 native Spanish speakers as a control group. They found that the L2 learners performed better in written tasks, whereas heritage speakers did better in the oral task though they had started learning earlier. Thus, the hypothesis that it is better to start learning earlier received only partial support.

The answer for the fourth question is positive. A number of studies found no age effect for the acquisition of L2 grammar. White and Genesee (1996), for example, investigated the judgment of Subjacency and the Empty Category Principle in English. The subjects from mixed L1 backgrounds (i.e. the majority of them were native speakers of French and of Germanic or Romance languages) had near-native (n = 44) and non-native levels (n = 45) of proficiency in English. They were divided into four starting age groups: 0-7 years, 8-11 years, 12-15 years, and 16+ years. The statistical analysis demonstrated that there was an absence of age effect when each grammatical sentence-type was analysed separately. The near-native group showed significantly greater accuracy than the non-native group in judging the ungrammatical sentences, but no significant effect was found due to age or interaction

between age and proficiency group. This result was interpreted as showing that L2 learners could achieve native-like competence under the constraints of universal grammar (UG), even for learners who were first exposed intensively to an L2 after the age of 16. Similar results were found in Yew's (1995, cited in Bialystok 1997) study with 31 Chinese learners of English. The subjects were asked to judge the grammaticality of 160 English sentences presented in oral and written form respectively. It was found that the judgement scores had no correlation with age of arrival, but significantly correlated with length of residence in Canada ($r = .40$). The age effect was present in that the late starters (starting age $\geq$ 15) outperformed the earlier starters (starting age <15) on the written judgement task. This suggested that starting age was not a significant predictor of performance in the grammaticality judgement task. Instead, the modality of the tasks influenced the effect of starting age. The correspondence between first and second language structures was considered the most important factor affecting acquisition.

In sum, studies undertaken in second language settings have produced mixed results. Some studies which detected age effects lent support to the existence of a critical cut-off age, though the cut-off ages differed (e.g., ages 12, 14, 15 or 16). Others provided counter-evidence for the CPH, showing a continued age effect even after puberty. Despite advantages for older children over younger children being found in few studies, it seems that the acquisition of L2 grammatical features favours younger starters more than adult learners. In addition, age effect may also depend on the learners' L1s (e.g., Bialystok & Miller, 1999) and the modality of the task (e.g., Yew, 1995).

3.2.2.2 Research investigating the effect of starting age on L2 proficiency of learners in foreign language contexts

While a large number of studies have investigated age effects in second language contexts, few studies have done so in foreign language settings. This is perhaps because the data in foreign language contexts is neither sufficient nor appropriate to test the existence of the critical period (Long, 1993).

Empirical evidence in foreign language settings generally lends support to the hypothesis that older learners are superior to younger ones and late starters excel over early starters. For example, Muñoz (2003), in her study of bilingual Catalan-Spanish learners of English in

Spanish, found that late starters (starting age = 11) significantly outperformed early starters (starting age = 8) in their oral production scores after both 200 and 416 instructional hours. A stepwise regression analysis showed that L1 proficiency was the strongest predictor at both points of time of testing. Garc á Mayo (2003) found the same results when investigating the acquisition of pro-drop parameter related grammar rules in English with similar participants (i.e., Spanish and Basque bilingual school learners). Older learners (aged 16-17, n = 18) who started learning English at ages 11-12 significantly outperformed younger learners (aged 13-14, n = 26) who started learning English at ages 8-9 at the second point of time of testing (i.e., after 594 hours of learning). In addition, length of exposure had a positive and significant effect on the judgment scores. The results suggested that in a foreign language setting the longer the exposure (i.e. formal study) to the L2, the better the performance. However, an early start did not benefit learning when the hours of instruction were held constant. These findings are in line with those of other studies carried out using different types of tasks such as grammaticality judgment, cloze, dictation, written composition and minimal pair tasks. (Celaya, Torras, & P érez Vidal, 2001; Garc á Mayo, 1999; L ázaro Ibarrola, Garc á Mayo, & Liceras, 2001; Victori & Tragant, 2003).

In summary, when number of years is held constant, an early start does not necessarily result in better performance in the use of grammatical structures in a foreign language setting.

3.2.3 Number of years of study

The number of years of study is another age-related factor relevant to this study. A number of variants of this term have been examined in the literature such as *length of exposure*, *length of residence, length of formal instruction*. Intuitively, it seems likely that the greater the number of years of study, the better the performance. However, empirical studies have yielded mixed results in different contexts.

3.2.3.1 Empirical studies in second language contexts

In second language contexts, some studies have found a significant effect for the number of years of study on the acquisition of L2 grammar. The Heidelberger Forschungsprojekt 'Pidgin-Deutsch' (1978, cited in R. Ellis 2008), for example, found that the length of residence in Germany explained the workers' acquisition of L2 German for the first two years of their stay but was subsequently overridden by other factors such as number of years of formal education and contact with Germans in different settings. Yew (1995) reported a

significant correlation between the judgment scores of 31 Chinese learners of English and their length of residence in the second language setting, Canada ($r = .40$). Geeslin (2003) revealed a clear effect of length of study on the choice of Spanish copular *ser* and *estar* (both meaning 'to be') by English native speakers (n = 28).

The effects of the number of years of study may vary according to the tasks used. R. Ellis (2009b) examined the obligatory use of 17 English grammatical structures in five tests (i.e., an oral narrative test, an oral imitation test, an untimed grammaticality judgement test, a timed grammaticality judgement test, and a metalinguistic knowledge test). The 91 non-native speakers of English had been learning English for an average of 10 years – mostly in a foreign language context – and had lived in a second language context for an average of 1.9 years. It was found that only the judgments of ungrammatical items in the untimed grammaticality judgement test significantly related to the number of years of formal instruction ($r = .27$, p < .05). The mixed results may be due to the fact that the number of years of study included study in both foreign and second language contexts. Clearly, the quality of learning in foreign language contexts is not comparable to that in second language contexts (Johnson & Newport, 1989, 1991).

In contrast, some studies found that the number of years of study had no effect. Patkowski (1980) found that only the amount of informal exposure had a significant effect among the factors tested which included years spent in the U.S., amount of informal exposure to English, and amount of formal instruction. However, this effect was negligible in comparison with the age factor. Johnson and Newport (1989) established that neither the number of years of exposure to English beyond five, nor the amount of classroom instruction was related to the grammaticality judgement scores.

3.2.3.2 Empirical evidence in foreign language contexts

Likewise, little research has been done on the effect of the number of years of study in foreign language contexts. Studies conducted in foreign language settings also have reported mixed findings. García Mayo (2003) found that the length of formal instruction had a positive effect on the judgment of pro-drop parameter related grammar rules in English by bilingual Catalan-Spanish school learners of English. As García Lecumberri and Gallardo (2003) indicated, the combination of high quality and extensive exposure together with early starting

age is a good predictor of native or near-native foreign language acquisition. In contrast, Geeslin and Guijiarro-Fuentes (2005) found that the number of years of study did not correlate with the frequency of choice of Spanish copula *ser* and *estra* (both meaning 'to be') by 27 university students in the U.K. with mixed L1s (i.e. 11 English, 4 French, and 11 German native speakers). One of the possible explanations may be that the choice of Spanish copula is subject to variable rules in Spanish. In addition, the inconsistency may also be due to the subjects, who were school learners in García Mayo's study (2003) but adults in Geeslin and Guijiarro-Fuentes's study (2005).

3.2.4 Measuring language proficiency

All the above studies on starting age and the number of years of study measured L2 proficiency by means of either oral production or metalinguistic tasks (i.e., grammaticality judgement tasks) or both. This raises a question about whether the mixed findings were due to the different types of data. This section will discuss this issue.

3.2.4.1 Oral production

With respect to the effect of starting age, studies conducted in second language settings show mixed results. Some demonstrated that L2 learners' oral proficiency favoured younger learners over their older counterparts (W. Klein & Dittmar, 1979; Montrul, et al., 2008; Oyama, 1978; Patkowski, 1980). Others consistently showed that older children outperformed younger ones (Ekstrand, 1976; Fathman, 1975; Snow & Hoefnagel-Höhle, 1978). The conflicting findings may be attributed to the differences in the number of years that learners had been learning an L2, and whether learners started learning before or after puberty. As Krashen *et al.* (1979) concluded, acquirers who begin natural exposure to second languages during childhood generally achieve higher second language proficiency than those beginning as adults, while older children acquire a second language faster than younger children in the early stages of morphological and syntactic development where time and exposure are held constant.

Findings in foreign language contexts seem to support a reverse hypothesis, namely, that older learners are better than younger ones and late starters are better than early starters (Muñoz, 2003). However, it should be noted that the so-called older learners (aged between 16 and 17 years) in these studies all started receiving formal instruction in English before the age of 13. Thus, this finding for foreign language contexts is consistent with that for second

language contexts in that older children outperformed younger ones. This lends support to Kreshen *et al's* (1979) argument (i.e., older children display a faster rate of learning than younger children in the early stages) because the study in foreign language settings is usually at an initial stage of learning.

Many studies that have examined the effect of the number of years of study have found no significant relationship between oral production performance and the number of years of study in second language contexts (Oyama, 1978; Patkowski, 1980). To my knowledge, no studies have investigated this relationship in a foreign language context.

3.2.4.2 Metalinguistic knowledge

A number of studies have examined the effect of age on L2 grammar using metalinguistic tasks, particularly grammaticality or accessibility judgement tasks. The findings have been mixed. Some lend support to the hypothesis that 'younger = better' and the earlier the better (Bialystok & Miller, 1999; DeKeyser, 2000; Johnson, 1992; Johnson & Newport, 1989, 1991; McDonald, 2000, 2006). Others cast doubt on the hypothesis by showing that some L2 learners who started learning an L2 as adults performed similarly to native speakers in the grammaticality judgement tasks (Birdsong, 1992; Birdsong & Molis, 2001; L. White & Genesee, 1996). Moreover, there are studies showing that starting age is not a significant predictor of the accuracy of judgements (Yew, 1995). The mixed results raise questions as to the reliability and validity of grammatical judgement tasks as a measure of overall language proficiency (R. Ellis, 1991).

Studies in foreign language settings have found that older learners perform better than younger ones in oral production tasks and late starters perform better than early starters in metalinguistic judgment tasks.

Research findings concerning the effect of the number of years of study in second language contexts are mixed. Length of residence in a target language environment significantly predicted the accuracy of grammaticality judgments in some studies (e.g., Yew, 1995), but not in others (e.g., Johnson & Newport, 1989). R. Ellis (2009b) found that the number of years of study only predicted the judgments of ungrammatical sentences. The contradiction may reflect the difficulty in distinguishing or separating the actual period of learning from other related factors such as length of exposure, length of residence, and length of learning

prior to arrival in a second language setting. This situation is much simpler for foreign language learners who have never been exposed to a second language setting. Thus, research findings in foreign language settings are relatively straight forward, showing that length of formal instruction has a positive and significant effect on grammaticality judgement abilities (e.g., García Mayo, 2003).

3.2.4.3 Differences

A comparison of the above findings shows that oral production tasks favour younger learners and early starters over older ones and late starters, but this is not always the case for metalinguistic knowledge tasks. This generally confirms the argument that older L2 learners have superior analytical and cognitive skills (Cummins, 1981). In foreign language settings, it is possible for so-called older learners to outperform younger ones due to their better L1 proficiency, experience in formal learning, and motivation to learn no matter whether oral production or grammaticality judgements are considered. Setting appears to be a significant variable influencing whether starting age has an effect on learners' L2 proficiency. This effect is evident no matter whether L2 proficiency is measured by means of an oral production or metalinguistic judgments.

Studies in second language contexts have found mixed effects for the number of years of study on performance in both oral production and metalinguistic judgement tasks. These inconclusive findings may be due to the confounding of the number of years of study and settings (i.e., whether the study occurs in foreign or second language contexts and whether learning is subject to an instructed or a naturalistic setting). Studies in foreign language contexts have attested to a positive effect of number of years of study on the development of grammatical knowledge. To my knowledge, however, no study has examined this in terms of oral production skills. This constitutes a gap in the current literature.

3.3 Setting

The previous section demonstrated that the effects of starting age and the number of years of study are influenced by setting or learning context to a certain extent. This section addresses the effects of setting on the development of L2 oral and metalinguistic abilities. Although the current study does not examine the effects of study abroad, the studies that have compared the effects of study abroad (SA) and at home (AH) will be reviewed for two reasons. First,

these comparative studies have employed methodologically reliable and valid measures, and included a control group of home-based learners. Second, the distinction between the second and foreign language setting group in my study resemble those between the study abroad group and the at home group in other studies. Of the 90 learners in China, the second language setting, 74 (82%) of them had previously learned Chinese in a foreign language setting. In contrast, of the 20 learners in New Zealand, the foreign language setting, 13 (65%) had never been to China and 4 (20%) students had only had a very short study abroad experience (i.e., five to six weeks).

3.3.1 Theoretical perspectives

L2 learning occurs in the following contrastive contexts: second versus foreign language settings (i.e., depending on whether the target language serves as the native language or a foreign language for the majority of people) (R. Ellis, 1994), naturalistic versus classroom settings (e.g., N. Ellis & Laporte, 1997; Pica, 1983), and majority, official, international or minority language settings (R. Ellis, 2008c; Judd, 1978; Siegel, 2003). Indeed, actual L2 learning settings are relatively complex. For instance, the three types of L2 learning contexts that Collentine and Freed (2004) identified (i.e., classroom learning in the home country, instructional or naturalistic learning in study abroad contexts, and classroom learning in instructed immersion foreign language situations) combine both instructional and social contextual components. The current study concentrates on the relative effects of second and foreign language instructional settings.

3.3.2 Empirical evidence

Research into the influence of learning context on L2 acquisition began with Carroll's (1967) survey of 2,782 U.S. college seniors who majored in French, German, Italian, Russian, or Spanish at 203 institutions. It was found that the duration of study abroad was a more important predictor of L2 proficiency (measured in terms of metalinguistic knowledge) than language learning aptitude or years of language study. Since then, a growing number of studies have tested the efficacy of study abroad (Brecht, Davidson, & Ginsberg, 1995; Churchill, 2006; Coleman, 1998; DuFon & Churchill, 2006; Dyson, 1988; Golonka, 2006; Gore, 2005; Kinginger, 2007; Lafford & Collentine, 2006; Meara, 1994; Murphy-Lejeune, 2002; Opper, Teichler, & Carlson, 1990; Pellegrino Aveni, 2005; Tanaka & Ellis, 2003; Willis, Doble, Sankarayya, & Smithers, 1977). A few studies have investigated the effects of

learning context by comparing the performances of learners in a second language setting with those of a control group in a foreign language setting (Collentine, 2004; DeKeyser, 1986, 1990, 1991; D áz-Campos, 2004, 2006; Hisama, 1995; Isabelli & Nishida, 2005; Iwakiri, 1993; Lafford, 1995, 2004; Rodriguez, 2001; Segalowitz & Freed, 2004; Stevens, 2001; Torres, 2003). Among these studies, even fewer have examined the gains in oral production and metalinguistic knowledge (cf. Collentine & Freed, 2004; Freed, 1995a). A summary of these studies is shown in Table 8.

Table 8: Studies That Have Examined the Effects of Setting on the Development of L2 Grammatical Competencies

Study	Participants	Target feature	Duration	Methods	Main results
DeKeyser (1986, 1991)	American university students of second-year Spanish in Spain (SA = 7) in the US (AH = 5)	Spanish *ser, estar*	16 weeks	Grammar test; oral interview; picture description; recall	SA = AH in grammar and oral proficiency
Huebner (1995)	University students of L2 Japanese at beginning-level in Japan (SA = 12) in the US (AH = 12)	Zero anaphora, full noun phrase in Japanese	9 weeks	Japanese Proficiency Test; Oral Proficiency Interview (OPI); a narrative telling	SA = AH in the use of zero anaphora as opposed to anaphoric pronouns, though the SA group produced more text and showed greater variation than the AH group in the retelling data
Howard (2001)	18 Irish advanced learners of French in 3 groups. Group 1 & Group 2 were about to participate in or just returned from a study abroad program. Group 3 received instruction at home as controls (the number of learners in each group was unavailable).	Past time morphology in French	NA	Oral data elicited through a sociolinguistic interview	SA > AH in accuracy in their use of past time morphology across a more expansive range of aspectual contexts. After controlling for the effect of a number of linguistic factors, the learners' contextual use of past time morphology appears to be relatively similar
Torres (2003)	AH (n = 5) SA (n = 10)	Spanish *clitics (e.g., direct and indirect object pronouns, reflexives)*	16 weeks	OPI	SA = AH in the use of *clitics*
Collentine (2004)	American university students in the US (AH = 20) and in Spain (SA = 26)	Morphosyntactic and lexical development in Spanish	16 weeks	OPI	SA < AH in the accuracy of discrete grammatical features (copula, present-tense-verb, indicative, subordinate-conjunction, and subordinate-clause); SA>AH in narrative abilities

Segalowitz & Freed (2004)	American university students of the third year Spanish (AH = 18) (SA = 22)	Spanish oral fluency	9 months	OPI; various cognitive measures	SA > AH in fluency and proficiency level
Isabelli & Nishida (2005)	Intermediate (fifth or sixth semester) learners (AH = 32) Advanced (third year) learners (SA = 29)	Spanish subjunctive use in subordinate clauses	9 months	Simulated OPI at three times	SA > AH in the use of subjunctive

Note. SA= the study abroad group; AT = the at home group; OPI = Oral Proficiency Interview

All but two studies (Howard, 2001; Huebner, 1995) investigated the acquisition of grammatical features in Spanish by native speakers of English. All studies used oral interview data, but presented conflicting results. On the one hand, some studies have shown that the study abroad setting is beneficial in aiding the development of grammatical structures. Segalowitz and Freed (2004), for example, observed significant improvements in global oral proficiency (measured by the OPI) in the SA group but not in the AH group. Isabelli and Nishida (2005) demonstrated that the SA group performed far better than the AH group with respect to the Spanish subjunctive in oral production over a period of nine months. It should be noted, however, that the superior performance of the SA group may be due to the fact that the SA learners were at an advanced level of proficiency, but the AH learners were at an intermediate level. Howard (2001) found that the SA group achieved higher accuracy in the use of aspect marking (i.e., passé composé vs. imparfait) in French across a wider range of contexts than the two AH groups. On the other hand, no evidence of such a positive effect for the SA setting was found in the studies of the development of Spanish copular *ser* and *estar* (DeKeyser, 1986, 1990, 1991), Japanese zero anaphora as opposed to anaphoric pronouns (Huebner, 1995), and Spanish *clitics* (Torres, 2003). Furthermore, the SA learners did not make as much progress as AH learners in accurate use of Spanish verbs and subordinate conjunctions (Collentine, 2004).

The modality of the data seems to not be an influential factor where setting is concerned. This is because, as shown in Table 3.1, the studies that adopted a grammaticality judgment test (DeKeyser, 1986, 1991) yielded results no different from those that used oral production interview data (Torres, 2003). However, the effects of setting may depend upon the choice of

dependent variables such as the aspects of linguistics. As Collentine and Freed (2004) summarised, studying abroad evidently enhances fluency, lexical abilities, and sociolinguistic awareness, but fails to develop lexical and grammatical abilities quickly. Empirical studies with no control group also lend support to this, showing that SA learners gain in fluency or oral skills (Dyson, 1988; Willis, et al., 1977) but not in morpho-syntactic abilities (Möhle & Raupach, 1983).

Moreover, the effects of setting on the acquisition of L2 grammar also depend upon learners' proficiency level. Freed (1995b) suggested that there might be a proficiency threshold at where learners most benefit from study abroad (Brecht, Davidson, & Ginsberg, 1991; Brecht, et al., 1995; Collentine, 2009; Regan, 1995). As some studies showed, the second language setting appears to be more beneficial for elementary learners than advanced learners (e.g., Collentine & Freed, 2004). Differential performance was observed in learners at different levels of L2 proficiency. Torres (2003), for instance, found that the lower proficiency learners of Spanish rarely used *clitics*, whereas the intermediate learners had begun to use *clitics,* though in a restricted way compared to native speakers' use. The effects of L2 proficiency on the variation in learners' performance will be addressed in Section 3.5.

The mixed findings may also due to the length of the study abroad. The studies shown in Table 8 investigated a large range of study abroad situations, ranging from 9 weeks to 9 months. The studies involving a shorter period of study abroad found no difference between the SA group and the AH group (DeKeyser, 1986, 1991; Huebner, 1995; Torres, 2003) and, in one study (Collentine, 2004) the SA group was even worse than the AH group in the development of grammatical abilities. In contrast, the studies involving a longer period of study abroad (e.g., 9 months) found that the SA group outperformed the AH group, particularly in oral proficiency (Isabelli & Nishida, 2005; Segalowitz & Freed, 2004).

3.3.3 The effects of setting on the acquisition of grammatical structures which are subject to variable use

Another way to probe the effects of setting is to examine the correlation between learners' performance and their exposure to second language settings. This method is often used to investigate the effects of setting on the development of sociolinguistic competence. Thomas (2004), for example, compared omission rates of French negation *ne* by 48 Anglophone

students who spent a year in France and 39 learners who studied in Canada, finding that omission rates of *ne* in oral expression increased slightly from 21.3 per cent to 27.3 per cent for the France (SA) group, and dropped from 32 per cent to 19.7 per cent for the Canada (AH) group. However, some studies found study abroad had little effect. Geeslin and Guijiarro-Fuentes (2005), for example, investigated Spanish copula choice (*ser* and *estar*) by 26 learners from three L1 backgrounds, English (n = 11), French (n = 4), and German (n = 11). The chi-square tests showed that the frequency of the choice of the copular verbs in a written contextualised preference task was not related to their study abroad experience.

The mixed results of these studies may be due to the discrepancies in the data collection methods (i.e., either judgement or preference tasks or oral production tasks), target features (i.e., whether they were categorical or variable) and the measure of learning employed (i.e., either fluency or accuracy, or omission rate).

3.3.4 Summary

The study abroad setting generally boosts the development of oral competence (Collentine, 2004), oral fluency (Freed, 1995a, 1998; Freed, Segalowitz, & Dewey, 2004; Segalowitz & Freed, 2004), lexical breadth (Milton & Meara, 1995), narrative and pragmatic ability (Segalowitz & Freed, 2004), sociolinguistic competence (Marriott, 1995), and communicative skills (Lafford, 1995, 2004). In contrast, the foreign language setting facilitates greater lexico-grammatical growth (Bardovi-Harlig & Dörnyei, 1998; Collentine & Freed, 2004).

However, the differences within each setting regarding both the quality and quantity of learning opportunities are likely to exceed the differences between them (see, R. Ellis, 2008c). Therefore, it is important to consider other factors such as educational and contextual factors (e.g., duration of stay in the second language setting, amount and types of formal instruction, learning opportunities, living conditions, and opportunity for contact with native speakers), methodological components (e.g., treatment designs, sample types and size, testing instruments, pre-experimental proficiency levels) (Lafford & Collentine, 2006), and individual factors (e.g., personality, learning styles, and cognitive abilities) (Segalowitz & Freed, 2004). Lafford (2006) hypothesises that the discrepancy in learners' performances in the two settings can be accounted for by 'individual learner perceptions of specific characteristics of the contexts (*setting, participants [status and roles], end/purpose, norms of*

62

interaction and interpretation)' and 'the interplay of cognitive factors (controlled vs. automatic processing, working memory) and socio-contextual attributes of individual learners'(p. 18), rather than by the context of learning alone.

3.4 L1 effects

L1 transfer is regarded as one of the central issues in SLA (Cenoz, Hufeisen, & Jessner, 2001; Gass & Selinker, 1983; Kellerman & Sharwood Smith, 1986; Odlin, 1989). Researchers' interest in the influence of the L1 on the acquisition of L2 dates from Contrastive Analysis (see Section 2.3.1). Although differences exist regarding how L1 transfer is conceptualised and labelled (N. Brooks, 1960; R. Ellis, 1994, 2008c; Odlin, 1989, 2003, 2005; Sharwood Smith & Kellerman, 1986), it is widely acknowledged that the learners' L1 has a powerful influence on the development of their L2 (DeKeyser & Juffs, 2005; Gass & Selinker, 1992; Gass & Selinker, 1983) as early acquisition studies have attested to (Hyltenstam, 1977).

3.4.1 Theoretical perspectives of L1 transfer

There are approximately five ways to examine the L1 transfer or cross-linguistic influence (R. Ellis, 2008c). The current study is only concerned with the third type, that is, 'comparisons of the use of a particular feature in the interlanguage of learners from two or more different L1 backgrounds' (pp.353-4). Empirical studies that have examined L1 effects on the development of oral and metalinguistic competence are considered.

L1 transfer predicts that it will be easier for learners whose L1s are typologically close to the target language to acquire a target language than for those whose L1s are more distant from the target language. However, empirical studies have provided mixed results, particularly in the acquisition of morphology and syntax (Jarvis & Pavlenko, 2008) (See Tables 9 and 10 for a detailed review).

Table 9: Studies That Have Shown Evidence for L1 Transfer

Study	Subject	Target feature	Tasks	Main findings
Schachter (1990)	18 Dutch EFL learners, 21 Indonesian, 20 Chinese, and 20 Korean ESLs, as well as 19 English native-speaker controls. Their first exposure to the language was after age 12.	Subjacency in L2 English	A grammaticality judgement task	The Dutch speakers, whose native language shows the same range of subjacency phenomena as English, performed at the same level as the native speaker control group, whereas Korean speakers whose native language shows no evidence of subjacency, performed at the level of chance.
Sorace (1993)	24 English and 20 French speakers of Italian who were aged 23-46 and started learning after age 15	Different types of unaccusative verbs (e.g., modal verb constructions, and clitic climbing construction) in Italian	A grammaticality judgement task involving 48 sentences	French speakers were more sensitive to auxiliary choice with accusative verbs than English speakers, because French (in some respects) is similar to Italian in relation to auxiliary selection with accusatives.
Bialystok and Miller (1999)	33 Chinese and 28 Spanish speakers of English who began learning English before or after age 15; 38 native-speaker controls	Five grammar structures in English (presented in both an oral and written form)	A grammaticality judgement test (both accuracy of judgement and time taken to respond were measured)	The younger learners performed differently from the older learners in the grammaticality judgement task in the case of the Spanish learners but not in the case of the Chinese group. A correlation of -.63 between L1 and L2 proficiency was found for Chinese but not for Spanish, and only for oral grammaticality judgment test.
Bruhn de Garavito (1999)	10 English and 10 French near-native speakers and 10 English advanced learners of Spanish who began learning Spanish in a formal setting after puberty; native speakers of Spanish as a control group	Structures related to the clitics *se* (in impersonal constructions and in unaccusative construction) and *le* (in dative constructions) in Spanish	1) An oral interview task 2) Several grammaticality judgment tasks	Although L2 grammars are similar to Spanish native speaker grammars, English native speakers do not perform as well as French native speakers due to L1 effect.
McDonald (2000)	Native Spanish early acquirers (5 or younger) (n = 14) and late (post-	12 grammar rules in English (see Johnson &	A grammaticality judgement test	Native languages appeared to make a difference for early acquirers, whereas a later age of acquisition caused a more

	puberty) acquirers (n=14); Vietnamese early acquirers (n = 14) (starting before age 5) and child acquirers (n = 10) (starting between age 6-10 years); 14 English native-speaker controls	Newport 1989)	consisting of 105 sentence pairs	general problem. Native Spanish early acquirers were not distinguishable from native English speakers, whereas native Spanish late acquirers had difficulty with all aspects of the grammar tested except word order. Native Vietnamese early acquirers had difficulty with those aspects of English that differ markedly from Vietnamese. Native Vietnamese child acquirers had more generalized problems, similar to those of native Spanish late acquirers. Early learners may have trouble mastering the L2 if its grammar is markedly different from their L1, and late learners may be able to master the L2 if its grammar is highly similar to their L1.
Montrul (2001)	1. 18 adult Turkish lower-intermediate and 29 Spanish adolescent intermediate EFL learners,19 English native-speaker controls; 2. 19 Turkish and 31 English intermediate SFL learners, 20 Spanish native-speaker controls; 3. 18 English, 24 Spanish and 9 Japanese TSL learners, 20 Turkish native-speaker controls	Two classes of causative verbs: physical change of state verbs with agentive subjects and psychological change of state verbs with experiencer objects in L2 English, Spanish and Turkish	1) A cloze test 2) A vocabulary translation task 3) A picture judgement task	Morphological errors in the three languages are constrained by the morphological patterns of the learners' first language (L1s) (i.e., English has predominantly zero-morphology, Spanish has anticausative morphology, and Turkish and Japanese both have causative and anticausative morphology). The results confirmed the prediction of the Full Transfer/Full Access Hypothesis of Schwartz and Sprouse (1994, 1996) that the initial state of L2 acquisition is the full computational system of the L1 grammar, including all the abstract features but excluding the morpho-phonological matrices of lexical and functional items.
Sabourin, Stowe, & de Haan (2006)	25 German, 21 Romance, and 24 English native speaking learners of Dutch at an advanced level of proficiency	Grammatical gender system in Dutch	1) A simple gender assignment task, 2) Agreement between the noun and the relative pronoun	Performance on grammatical gender was affected by the L1s of the participants but not by their levels of general syntactic proficiency. A distinct performance hierarchy was found with the German group performing the best (though significantly worse than native speakers), the Romance group performing well above chance, and the English group performing at chance.

				This suggests that L2 acquisition of grammatical gender is affected more by the morphological similarity of gender marking in the L1 and L2 than by the presence of abstract syntactic gender features in the L2.
Ionin and Montrul (2010)	Study 1: 24 Spanish and 29 Korean EFL learners, and 19 native English controls. Study 2: 11 Spanish and 9 Korean ESL learners	Plural noun phrases with articles in English	1) A picture-matching task, 2) A cloze task, 3) A truth-value judgment task, 4) An acceptability judgment task	Study 1. Spanish learners over-accepted the generic interpretation of English definite plurals to a greater extent than proficiency-matched speakers of Korean, an articleless language. Study 2. The results of a follow-up study showed that with advanced proficiency and increased immersion in the target language, Spanish-speaking learners were as target-like as Korean-speaking learners of English on the interpretation of definite plurals. It was suggested that recovery from first language transfer is possible.
B. Yuan (2010)	107 English native speaking learners in the UK and 111 Japanese learners of Chinese in Japan and in China; 20 native-speaker controls. The learners were divided into five proficiency groups according to their performance in a Chinese cloze test	Chinese *wh*-words used as existential polarity words (EPW)	An acceptability judgement test	Japanese learners performed better than English learners in acquiring certain EPW licensers. This may be due to the advantage that Japanese learners have over English learners that *wh*-words can be used as EPWs in Japanese but not in English. It was suggested that that success or failure in establishing interface relations in L2 grammars is likely to depend on a number of variables, including the categorical nature of individual elements involved in the interface relationship, the status of these elements in the target language speaker's grammar, the input that learners are exposed to, and cross-linguistic influence.

Note. EPW = existential polarity words in Chinese

Table 10: Studies That Have Shown No Evidence of L1 Transfer

Study	Subject	Target feature	Tasks	Main findings
Schachter (1989)	20 Chinese, 21 Korean and 20 Indonesian ESL learners, 19 native-speaker controls	Subjacency in L2 English	A written grammaticality judgment task which involved 48 sentences (including ungrammatical and grammatical)	No difference was found between the three different language groups. This suggests that adult learners of a second language do not have full access to UG, specifically to the subjacency principle.
Bley-Vroman & Chaudron (1990)	Spanish, Japanese, and Chinese learners of English at three proficiency levels (*Low,* Mid, High)	Subordinate clauses and anaphora in English	Several elicited imitation tasks as measures of production; A sentence act-out task (with geometric figures) as a measure of comprehension	There is no evidence for a processing difference based upon L1 grammatical contrasts. It was concluded that the L1-based differences reported by Flynn (1984, 1986, 1987a, 1987b) were evidently artifacts of the experimental method and inappropriate analysis, especially of the incorrect use of analysis of covariance in an attempt to correct for a mismatch between experimental groups.
Guijarro-Fuentes, & Geeslin (2003)	19 native speakers from Spain, 10 native Spanish speakers from 9 different origins in the US, and two groups of 22 near-native speakers of Spanish (i.e., 11 Portuguese SSL and 11 English SFL)	Copula choice in L2 Spanish *ser* vs. *estar*	1) A background questionnaire 2) A contextualized grammaticality preference task	Portuguese native speakers do not have any advantage over their English counterparts although Portuguese shares some properties related to copula choice with the Spanish language and English does not. The study suggests that the variation between native and non-native speakers cannot be explained by age factors (i.e., age of arrival, number of years of exposure and chronological age the linguistic variation to) alone.
B. Yuan (2004)	48 French, 51 German and 67 English speaking CFL learners at	The syntactic structure of clausal negation in	1) An oral-production task; 2) A judgement task	Little variation was found between the L2 groups in the acquisition of Chinese negation and their behaviours were all

	mixed levels of proficiency;10 Chinese native-speaker controls	Chinese		native-like. This finding suggests that L2 grammars can have fully and appropriately specified features of functional categories from the initial stage of L2 acquisition even though these features may have different values in learners' L1s.
White, Valenzuela, Kozlowska– Macgregor, & Leung (2004)	48 French and 68 English adult SFL learners at three levels of proficiency	Gender and number agreement in Spanish	1) Oral production data 2) An interpretation task	The results from both tasks show that there were no significant effects for L1 (i.e. there was no difference between French and English native speakers in their use of gender and number agreement) or for prior exposure to another second language with gender, but significant effects for proficiency (i.e., low proficiency groups differed significantly from native speakers, but advanced and intermediate groups did not).
X. Hu & C. Liu (2007)	41 English speaking CFL and 47 Korean speaking CSL learners at mixed levels of proficiency; 15 Chinese natives as a control group	Restrictive relative clauses in Chinese	A written grammaticality judgement test	The prediction that the L1/L2 difference hinders the acquisition of L2 RRCs for English learners but facilitates it for Korean learners was not confirmed. In contrast, English learners distinguished between target-like RRCs and non-target-like RRCs earlier than the Korean learners. It is argued that restructuring of less salient features encoded in functional categories takes longer and may be persistently problematic. It was suggestedthat the fact that Korean is more similar to Chinese (perhaps superficially, same head direction) leads learners not to restructure quickly, while the surface dissimilarity of English and Chinese gives rise to rapid restructuring in L2 grammars of learners.

Note. CFL= Chinese as a foreign language learners; CSL= Chinese as a second language learners; SSL= Spanish as a second language learners; SFL = Spanish as a foreign language learners; ESL= English as a second language learners; RRC= Restrictive relative clauses

3.4.2 Empirical evidence for L1 transfer

The studies in Tables 9 have revealed significant L1 effects. The influence of L1/L2 distance on the L2 acquisition appears to be independent of the typology of L2 (i.e., whether the target language is English, Chinese, Spanish, Italian, Turkish, or Dutch), task (i.e. whether data was collected from oral tasks or grammaticality judgment tasks), and setting (i.e., whether the subjects were recruited in a foreign or a second language learning context). However, advanced proficiency and increased immersion in the target language is likely to overcome the influence of L1 transfer. For instance, Ionin and Montrul (2010) found that Spanish speakers (n = 24) over-accepted the generic interpretation of English definite plurals to a greater extent than proficiency-matched speakers of Korean (n = 29), an articleless language. However, with advanced proficiency and increased immersion in the target language, the Spanish-speaking learners (n = 11) became as target-like as Korean-speaking learners (n = 9) on the interpretation of definite plurals. This suggests that learners can overcome the effects of L1 transfer.

By contrast, some studies (in Table 10) found no effect for L1 transfer regardless of the L2, task, and setting. The conflicting results of these two groups of studies may be due to the different target features and the learners' L2 proficiency levels.

3.4.3 Sociolinguistic studies of L1 influence

The studies that have focused on variability in interlanguage report mixed results regarding L1 effects. Whereas little variation in learners' performance was found in terms of their L1 in some studies (Geeslin & Guijarro-Fuentes, 2005; Xiaoshi Li, 2010; Regan, 1996; Tarone, 1985), the distance between structures in the L1 and L2 served as the best predictor of success of acquisition in other studies (Bialystok, 1997). Godfrey (1980), for example, reported that past tense marking by Spanish and Japanese learners of English varied according to their L1, apart from the difficulty of linguistic context, discourse context, the learners' proficiency level, and the specific tokens investigated. Shannon (1995) found that learners of Arabic (n = 10), which has an article system, marked semantically definite nouns in English with more frequency than learners of Japanese (n = 10), which does not have an article system. There are some studies that revealed a partial effect for the distance between L1/L2 on the variation in L2 learners' performance. Young (1993), for example, in the

examination of the functional constraints on the variable use of L2 English morpheme plural marking, observed that Czechoslovakian learners (n = 12) generally outperformed the Chinese learners (n = 12) as their L1 and English were less distant.

3.4.4 Summary

There are many reasons for the mixed and often contradictory findings of transfer studies. One of them is the lack of a common definition of L1 transfer (R. Ellis, 2008c). Another is the frequency of use of the target feature which is investigated in the target language. If a grammar structure is seldom used even by native speakers, the influence of the L1 is likely to be seriously underestimated (Odlin, 2003). The methods to collect data may be another source of variation. Kellerman (2001) proposed that narratives may serve as a particularly fruitful type of data because they constitute a context for examining both linguistic and conceptual aspects of transfer. As Odlin (2003) puts it, 'the most convincing evidence will come from multiple sources; spoken and written performances as well as responses to measures of perception, comprehension, or intuition' (p. 452).

3.5 Self-rated proficiency

As illustrated above, both the effects of starting age and setting on the acquisition of grammatical structures are mediated by learners' L2 proficiency. However, proficiency has been measured in different ways. In some studies, proficiency data was obtained by means of a standard proficiency test (Huebner, 1995) or a pre-test such as a Chinese cloze test (B. Yuan, 2010), or oral interviews which were rated by trained native speakers (Oyama, 1976; Patkowski, 1980). Proficiency was also estimated on the basis of the stages of the language course that learners were enrolled in at the time of testing. It seems that a valid and reliable or accurate measure of L2 proficiency is not always achievable. The current study used self-rated proficiency instead of an objective measure of proficiency. The following sections review the studies that have examined the relationship between self-rated proficiency and the acquisition of grammatical structures.

3.5.1 Validity and reliability of self-rated proficiency

Self-reported proficiency has proved to be a reliable and valid predictor of measured proficiency in bilingual studies (Hukuta & D'Andrea, 1992; Kominski, 1989; Marian, Blumenfeld, & Kaushanskaya, 2007; McArthur, 1993; McArthur & Siegel, 1983) and in L1

studies (Shameem, 1998). Marian *et al.* (2007) in a bilingual study developed a reliable and valid Language Experience and Proficiency Questionnaire (LEAP-Q) to predict the relationships between self-reported and behavioural measures of proficiency. The statistical analyses suggested that self-reports were reliable indicators of language performance and self-reported speaking proficiency was a relatively accurate predictor of L2 performance.

3.5.2 Empirical evidence for the effects of self-rated proficiency

Only a few studies have employed a self-reported measure of proficiency in SLA. Bialystok and Hakuta's (1999) examination of the critical period hypothesis is one of the few available. Based on U.S. census data, they found a similar linear trend for the relationship between self-rated proficiency scores and age of arrival for both Chinese (n = 24,903) and Spanish (n = 38,787) immigrants, with the Spanish immigrants manifesting slightly higher self-rated proficiency than the Chinese.

To my knowledge, no study has tested the relationship between self-rated proficiency and both L2 metalinguistic knowledge and oral production ability. The current study will partially fill this gap.

3.6 Gender

The last potential source of variation in interlanguage that will be dealt with in this study is gender. The term 'sex' and 'gender' have been used interchangeably to indicate the distinction between male and female in SLA. However, gender places emphasis on the social context, while sex constitutes a biological distinction (see, R. Ellis, 2008c). Sex used in early research in SLA was regarded particularly 'as a static, bipolar opposite in relation to language use and learning' (R. Ellis, 2008c, p. 313). The current study views gender as a 'complex system of social relations and discursive practices differentially constructed in local contexts' (Norton & Pavlenko, 2004)**.**

The study of the influence of gender on L2 acquisition has been based on three paradigms: the socio-cultural paradigm, the sociolinguistic paradigm, and the psycholinguistic paradigm. In socio-cultural approaches, the interest in the relationships between language and gender has been prompted by feminist concerns about the connections between sex, power and language since 1970s (Lakoff, 1975; Pavlenko & Piller, 2001; Thorne & Henley, 1975).

Sociolinguistic accounts of gender effects are motivated by Labov's seminal works on the role of sex in linguistic change. For instance, Labov (1991, 1994, 2001) found that men and women played different roles in linguistic change. In stable sociolinguistic stratification, men use non-standard forms more frequently than women, whereas in the majority of linguistic changes, women use the incoming form more frequently than men Finally, psycholinguistic accounts pay particular attention to the influence of the sex of interlocutors on interaction. In summary, the effect of gender on L2 acquisition tends to be considered more as a social construct than as an individual factor.

3.6.1 Empirical evidence for the effect of gender

Little work has been conducted on the effect of gender on the accurate use of L2 grammatical structures in oral production and on grammatical judgments (Plough & Gass, 1993; Yule & Macdonald, 1990). One relevant study showed that gender had no effect on the English proficiency of 168 third-year university students in China when this was measured by an institutional version of TOEFL (Nisbet, Tindall, & Arroyo, 2005).

A few studies have examined gender effects on the variable use of grammatical structures within the framework of the Labovian sociolinguistic paradigm. Their results are conflicting. On the one hand, gender effects were found in the variable use of some linguistic features such as the English *ing*/*in* by Vietnamese or Cambodian immigrants (Adamson & Regan, 1991), the French *seulemen*t or *juste* (both meaning 'only') (Mougeon & Rehner, 2001) and pronouns *nous* and *on* (both meaning 'we') (Rehner, Mougeon, & Nadasdi, 2003), and the Chinese morphosyntactic marker *de* (Xiaoshi Li, 2010). On the other hand, no significant gender differences have been found in the presence/absence of third person plural marking for French verbs in spontaneous oral discourse in an immersion secondary school context (Nadasdi, 2010), in the omission rates of L2 French pre-verbal particle *ne* by 64 university students of French in the UK (Dewaele, 2004a), in the choice of the French address pronoun *tu* (meaning 'you') by 101 learners in spoken data (Dewaele, 2004b), or in the use of French lexical forms such as *comme/like* with English discursive equivalents (Rehner, 2002).

3.6.2 Summary

In general, gender appears to have little effect on the acquisition of grammatical structures, although a clear gender effect has been observed in phonology (Adamson & Regan, 1991;

Major, 2004), speech norms (Regan, 1998), conversational interactions (Gass & Varonis, 1986, 1994; Pavlenko & Piller, 2001), age and rate of learning (Slavoff & Johnson, 1995), the acquisition of sociolinguistic competence (Mougeon & Rehner, 2001; Rehner, et al., 2003), learning gain during study abroad (Brecht, et al., 1995; Pica, Halliday, Lewis, & Morgenthaler, 1989; Polanyi, 1995), and willingness to communicate (Donovan & MacIntyre, 2004).

Within the studies that have found gender effects, the relative performance of males and females is mixed. Females showed superior acumen to males in certain aspects such as learning L2 French in a primary school setting (Burstall, 1975), overall English proficiency of Chinese university students in Hong Kong (Boyle, 1987), memorising German vocabulary (Nyikos, 1990), discriminating American English accents (Eisenstein, 1982) and achieving better fluency in using L2 French (Dewaele, 1998). The reverse was true, however, for listening vocabulary at university level (Boyle, 1987). In summary, some sociolinguistic variants that are sensitive to gender were preferred by one sex or the other, whereas the obligatory use of grammatical structures seems to resist the influence of gender.

3.7 Conclusion

This chapter has reviewed six potential sources of inter-learner variation: starting age, the number of years of study, setting, L1, self-rated proficiency, and gender. These six factors are summarised below.

1. Starting age has mixed effects on the grammatical development of oral and metalinguistic abilities. In a second language setting starting early is of clear benefit to the development of oral and metalinguistic competence, while in a foreign language setting late starters performed no worse than early starters. Clearly, the effects of starting age vary depending on setting.

2. The number of years of study shows a positive effect in a foreign language setting but a mixed effect in a second language setting. However, the influence of setting may also depend upon the choice of data collection methods and target features.

3. The effects of setting on the acquisition of grammatical structures depend on the duration of stay in the second language setting and on the aspect of language investigated. Study abroad groups gain more than the at home groups in oral and

metalinguistic abilities but only after a relatively long period of study abroad (e.g., 9
months).

4. L1 transfer shows conflicting effects regardless of the target language, task type (i.e.,
   either oral or metalinguistic tasks), and setting (i.e., either foreign or second language
   setting). However, the L1 effect is likely to be overcome with advanced proficiency
   and increased immersion in the target language setting.

5. The effect of self-rated proficiency is still under-researched. In particular, no study
   has examined its relationship with accuracy in the actual use of grammatical
   structures.

6. In general, gender has no effect on the acquisition of grammatical structures.

Although the effects of these potential sources of inter-learner variation have been examined
on the acquisition of L2 English, Spanish, and French, little research has been done on the
acquisition of L2 Chinese, which is typologically different from Indo-European languages.
The only exception is research investigating L1 transfer. The current study intends to address
this gap by focusing on the acquisition of one grammatical structure in Chinese – the ba
construction (*BC*).

Chapter 4 will illustrate the linguistic characteristics of *BC*, discuss its pedagogical
significance, and review relevant acquisition studies.

CHAPTER 4: THE CHINESE BA CONSTRUCTION

4.1 Introduction

This chapter addresses the target structures of this study: two types of the Chinese *BA* constructions (*BC*). The chapter is organised as follows. Section 4.2 provides a brief introduction to general characteristics of *BC*, including its definition, syntactic and functional constraints, and conditions of use. Section 4.3 elaborates the criteria and procedures for selecting the target features of this study, *BC1* – a [locative] nominal *BC* whose complement contains a NP – and *BC2* – a directional verbal *BC* whose complement contains a single or compound directional verb but no NP. Section 4.4 illustrates the linguistic characteristics of these two types of *BC*. Finally, Section 4.5 summarises the chapter.

4.2 Brief introduction to *BC*

4.2.1 Definition of *BC*

BC is a non-canonical word order structure in Standard Chinese or Mandarin. It was called '*ba zi ju*' [3] (lit. a sentence which contains a Chinese word of BA) first by L. Wang (1943) on account of the existence of the preposition [4] BA in the sentential structure. *BC* is also one of the few syntactic structures which have received intense debate in Modern Chinese linguistics (Ho, 1993; Jing-Schmidt, 2005; D. Shi, 1999 ; Sun, 1995, 1996, 2006; D. Xu, 2006; Zou, 1995). From the perspective of syntactic structure *BC* is straightforward (see C. N. Li & Thompson, 1981; Yuehua Liu, Pan, & Gu, 2001). However, there are some demanding issues in relation to *BC*, including a range of constraints on its functions and conditions of use. The complexity of functions of *BC* is partly reflected in the multiple labels for *BC*, such as

---

[3] There are a small number of alternatives for BA, but these alternatives are regional variants (e.g., Ka in Taiwan Southern Min) or belong to the formal register (e.g., *Jiang* as a formal variant of *ba*) (cf. Jing-Schmidt & Tao, 2009).

[4] There are three positions for the word class of *BA* in Modern Chinese. First of all, Chao (1968) labeled *BA* as a 'pre-verb' because *BA* functions like a verb by having an object and always positions preceding the main verb in a sentence. Second, Li & Thompson (1981) termed *BA* a 'coverb' because *BA* sometime can function as a verb but sometime as a preposition in Modern Chinese. Its function as a verb is very weak. In a third position, many researchers (e.g. H. Wang 1981,1985; Zhu, 1981; Lu and Ma, 1985) considered *BA* a preposition because *BA* in the *BC* serves as a function word and has lost the lexical meaning when it functions as a verb (i.e. '*grasp*' or '*take*'). The current research adopts the third position.

disposal construction (Shen, 2002; Song, 1979; H. Wang, 1981; L. Wang, 1943), executive construction (Hashimoto, 1971), accusative construction (e.g. Teng, 1975), and causative construction (W. Hu, 2004, 2005; Ye, 2004; Zhou, 2006). A full discussion of the terminologies of *BC* is beyond the scope of this research, so only the syntactic characteristics and basic functions of *BC* will be introduced below.

4.2.2 Syntactic characteristics and basic functions of *BC*

In syntactic structures, *BC* consists of four obligatory components which can be shown in a template, 'S + BA + NP + VP'. S is the subject which is usually a nominal phrase (NP) or clause; BA is a preposition; the NP following BA (hereafter *ba*-NP) is called the object of BA and refers to a participant in the event denoted by the verb which follows; VP comprises a verb (V) and a post-verbal component, residual[5] (R). V is the main verb which can be any valency (i.e., transitive, ditransitive or intransitive[6]). R can be any type of the major constituents in Chinese from a single aspect marker (e.g. LE) to a resultative DE clause. On some rare occasions, VP can be a bare disyllabic verb[7] or the combination of a preverbal adverbial and a single syllabic verb (e.g. *ba men 'yi tui'* [lit. BA door once push]/ 'quickly push the door').

No variation in the order of the constituents of *BC* is possible. However, the subject in *BC* is usually omitted when it is retrievable from context since Chinese is a pro-drop language (C. N. Li & Thompson, 1976). Negative adverbs – *bu*, *mei* (both meaning 'no') – and modal auxiliary verbs always precede BA (see examples 4.1 and 4.2 respectively), while an aspect marker is usually placed at the final position of the structure (e.g. perspective marker LE in example 4.3). It should be noted that the current study only focuses on the basic structure of *BC*, and other adjuncts in *BC* are not considered.

---

[5] The term *residual* was orally suggested by Dr. Helen Charters.

[6] The intransitive verb is not a typical form to be used here.

[7] For the sake of prosodic rhetoric, the main verb in *BC* can be a bare single-syllabic character in a poem (see, Chao, 1968; Feng, 1996, 2001), but this usage of *BC* is beyond the scope of the current research. In Modern Chinese, some disyllabic verbs (e.g., ti-gao /raise-high/ improve) indicating a telic and perfective event can be used as a main verb in *BC* without a complement.

4.1[8] 他　没　把　饭　吃　完

　　Tā　méi　bǎ　fàn　chī　wán

　　He　NEG　BA　meal　eat　complete

　　'He didn't finish his meal.'

4.2 他　能　把　饭　吃　完

　　Tā　néng　bǎ　fàn　chī　wán

　　He　can　BA　meal　eat　finish

　　'He can finish his meal.'

4.3 他　把　饭　吃　完　了

　　Tā　bǎ　fàn　chī　wán　le

　　He　BA　meal　eat　finish　PFV

　　'He finished all his/the meal.'

BC has aroused much debate among Chinese linguists since the middle of the twentieth century (e.g., Cui, 1995; W. Hu, 2004, 2005, 2010; S. Lü, 1948; Shen, 2002; C. Shi, 2010; Sun, 1995, 1996, 2006; Sun & Givón, 1985; Tai, 1985; H. Wang, 1981, 1985; L. Wang, 1943; Ye, 2004; B. Zhang, 2000; Zhou, 2006). The reasons for this are at least two-fold. First, the existence of *BC* is related to a central concern in Chinese linguistics, namely whether the canonical word order[9] in Chinese language is SVO or SOV. Whereas some researchers believe that the canonical sentence order in Chinese is SVO (see Sun & Givón, 1985), others

---

[8] Each Mandarin example is illustrated as two written forms. The first line is written system of Chinese, Chinese characters. The second line is Romanised pronunciation symbol, *Pinyin* (literally 'spell sound'), the official Romanisation system of the People's Republic of China, which is also the most widely used system in the media and scholarly writings on Chinese in the West. Chinese characters were used because many distinct words have the same phonetic form and characters disambiguate these. Two lines of English are translated below the Mandarin example. The line immediately below the Mandarin example glosses each Mandarin element with the clearest and the most literal English equivalent possible. The second line offers a translation of the whole utterance into idiomatic English, attempting to preserve the 'flavour' of the Mandarin utterance. Following convention, * refers to an utterance that is either structurally or semantically unacceptable to native speakers, and ? refers to an utterance that is odd but not necessarily unacceptable.

[9] According to Greenberg (1966), all languages can be categorised into different types according to the order of the three constituents, S, V, O. The basic word orders include: VSO, SVO, SOV, VOS, OVS, and OSV.

advocate that the canonical sentence order in Chinese has changed from SVO to SOV (e.g., C. N. Li & Thompson, 1976, 1981; Tai, 1985). This latter group of scholars point to the emergence of *BC* as evidence of this[10]. Second, despite considerable scholarly debate, a consensus on the functional constraints and conditions of use of *BC* has yet to be achieved and they continue to be a source of fertile academic discussion. Studies using various approaches from different perspectives in Chinese linguistics have all contributed to this debate (see, Chao, 1968; Cui, 1995; Feng, 1999, 2001, 2002; Hsueh, 1987, 1989, 1994; W. Hu, 2004; L. Jin, 1997; C. N. Li & Thompson, 1981; J. Li, 1924; F.-H. Liu, 1997; S. Lü, 1948; Sybesma, 1999; Tsao, 1987; H. Wang, 1981, 1985, 1995; L. Wang, 1943; Yang, 1998a, 1998b; Ye, 2004; B. Zhang, 2000; W. Zhang, 2001).

4.2.3 *BC* in the Chinese word order system

In Chinese, there exist three word orders – SVO, S BA OV (i.e., *BC*)[11] and OSV. SVO has been increasingly accepted as the canonical word order in Chinese (see Sun & Givón, 1985), while *BC* is regarded as a non-canonical word order. Syntactically, these three word orders can be used interchangeably to convey similar meanings under certain circumstances because the grammatical relationships between their basic constituents remain the same (S-V-O). In discourse, however, the Chinese language relies heavily on word order to convey information about topicality, temporal sequence, and thematic role relations (e.g. agentivity) (see Chao, 1968; Ho, 1993; C. N. Li & Thompson, 1981). Semantic, pragmatic and discourse functions therefore play a more important role than syntactic constraints in determining the selection of word order.

The three word orders (SVO, *BC*, and OSV) can be used to describe the same set of circumstances except for slight distinctions in focal information and implied expressive intention. Thus, multiple devices are available to convey the same meaning without considering discourse constraints. Examples of SVO, *BC*, and OSV are shown in examples 4.4, 4.5 and 4.6, respectively.

---

[10] Tai (1985) argues that the canonical word order in Chinese has changed from SVO into SOV by taking *BC* as one of evidence. The other evidence is that while the locative prepositional phrase (i.e. '*zai+location*' '*at+pl.*' always appears at a post-verbal position in classical Chinese, it is generally at pre-verbal position in Modern Chinese.

[11] See note 3 (p.75)

4.4 他 卖 了 书

    Tā mài le   shū

    He sell PFV book

    'He sold a book/books'

4.5 他 把 书 卖 了

    Tā bǎ shū mài le

    He BA book sell PFV.

    'He sold the/his book'

4.6 书 他 卖 了

    Shū tā   mài le

    book he   sell PFV

    'He sold   the book'

The potential referent of O (where O= ba-NP in the *BC*) differs among the three word orders. In the SVO structure shown in example 4.4, the referent of the noun *shu* 'book' could be any book, books in general, or a specific book which is known only by the speaker (indefinite) or by both the speaker and the hearer (definite). An accurate interpretation requires contextual information. The key point is that the noun *shu* 'book' cannot be topical at the sentential level (i.e., what this sentence is about) no matter whether it is definite or indefinite, known or unknown to the hearer. Rather it must be either new to the hearer or the focus of the sentential expression according to information structure[12]. Therefore, in example 4.4 the O can be any referential value including non-specific, indefinite-specific, definite-specific, and generic, but cannot be topical. Compared to other corresponding word order structures, only the O in the SVO structure can be used to refer to new referents which are first mentioned and non-specific (i.e., cannot even be identified by the speaker). The SVO structure usually emphasises the action and the event that the verb describes. According to a topic-comment

---

[12] Following Steedman *et al.* (1991), information structure can be construed broadly as comprising structural and semantic properties of utterances relating to the discourse status of their content, the actual and attributed attentional states of the discourse participants, and the participants' prior and changing attitudes (knowledge, beliefs, intentions, expectations, etc.). This broad view of information structure involves notions like focus, presupposition, given vs. new, theme vs. rheme and the various dichotomies such as topic vs. comment or focus, ground or background vs. focus etc. define these terms in a simple way for the reader.

view of the Chinese sentence, the O is part of the comment in the SVO in terms of information structure.

In example 4.5 (*BC*), the speaker assumes that the referent of the noun *shu* 'book' (i.e., *ba*-NP) is known to the hearer. Thus, the O (i.e. *ba*-NP), the noun *shu* 'book', in *BC* (example 4.4) can have any referential value apart from non-specific. That means that the *ba*-NP must be specific (i.e., at least known to the speaker). A bare *ba*-NP refers to a generic or definite referent[13]. The O is topical or presupposed because it provides background for the important new information, VP. In the VP consisting of V and R, only the main verb (V) provides new information for the S, but both the main verb (V) and the residual constituent (R) convey new information for the O. The O is given prominence as a subtopic or presupposition of the sentence following the main topic, *ta* 'he', according to the topic-comment approach (Hsueh, 1987, 1989, 1994; C. N. Li & Thompson, 1976, 1981; Tsao, 1987, 1990). Since Chinese sentences are generally subject to pragmatic constraints (Charles N. Li & Thompson, 1981) and follow a temporal order (Tai, 1985), the constituents referring to known information generally precede (i.e. are to the left of) those referring to unknown information. Thus, *ba*-NP in *BC* generally indicates a known entity which exists before the action denoted by the main verb has an effect on it.

In the OSV structure shown in example 4.6, the referent of the O, the noun *shu* 'book', must be old information known to both the speaker and the hearer, serving as the topic of the sentence (i.e., what the sentence is about). Thus, the O modified by indefinite determiner '*yi* "one" +classifier' is unacceptable or ungrammatical in the OSV structure.

In summary, among the three related word order structures, the degree of definiteness of the referent of the O becomes fixed as the O 'moves' towards the front of the sentence. The referent of O can be either definite or indefinite in SVO, but can only be definite in OSV. The O (*ba*-NP) in *BC* is generally definite or generic but can be indefinite-specific when modified by the indefinite determiner '*yi* "one" +classifier'.

---

[13] The *ba*-NP can be modified by an indefinite determiner (i.e. *yi* 'one' + classifier) to refer to indefinite-specific referent (i.e., a referent which is known to the speaker but not identifiable by the hearer) which exists before the action denoted by the main verb has an effect on it. (cf., P. Chen, 2003) (e.g., *Ta ba yi ben shu fang zai zhuozi shang, jiu zou le*. 'He put a book on the desk and then left').

From a pragmatic perspective, the topics in the three word orders also differ. As a topic-prominent language, the left-most position in Chinese is the topic of a sentence (C. N. Li & Thompson, 1981). Whereas S and O is the topic of SVO and OSV respectively, S is the topic and O is the sub-topic which needs to be commented upon in *BC*. The O (*ba*-NP) is generally known or specific information. Thus, the *BC* is understood as a possible topicalisational structure in a situation where the discourse topic in a topic chain is not changed.

4.2.4 Functions of *BC*

BC and parts of *BC* serve different functions. Like any sentence, *BC* expresses relationships between a predicate (the verb) and its subject/object etc. In addition, the choice of the *BC* encodes specific pragmatic information about the subject and object, and aspectual information about the event. The key function of *BC* is characterized as telic (i.e., having an end point) and perfective (i.e., having/ will be finished at an understood reference time) (Huang & Yang, 2004) or bounded events with a combination of the bounded verb and a specific ba-NP (F.-H. Liu, 1997). The specificity of the *ba*-NP is a referential feature or attribute related to information structure not a relationship with a predicate, so this is a kind of pragmatic (context dependent) information rather than what is generally called semantic. The nominal phrases in *BC* reflect semantic (thematic) *roles* such as agent, recipient, goal, and instrument. The features of the classes of the main verb are related to punctual/ non-punctual aspectual (semantic) types of events. This is not only because the main verb in *BC* is subject to certain constraints (i.e., the verbs describing mental activities such as *xihuan* 'like' cannot be used as the main verb in *BC*), but also because the use of *BC* creates a specific aspectual interpretation for the main verb (see, Huang & Yang, 2004; F.-H. Liu, 1997; Yang, 1998a; Yang, 1998b). The functions of *BC* seem to relate to multiple functions, so a simple description of the functions of *BC* is difficult.

From the perspective of pragmatic discourse function, *BC* is considered a kind of topicalisational structure (e.g., Hsueh, 1987, 1989, 1994; Tsao, 1987, 1990). Under the condition where the original subject stays at the topical position, the ba-NP is given prominence at the preverbal position and serves as a subtopic. In a discourse, the topics of the clauses remain the same, but *BC* can give prominence to the object (i.e. ba-NP) and make it a subtopic.

With regard to syntactic constraints, one of the functions of *BC* is to allow all arguments of certain ditransitive verbs to be overtly expressed. This relates to a generalisation that Chinese grammar does not allow both the object and the other relevant argument to co-appear at the post-verbal position except for the indirect object construction (C. N. Li, 1975). This poses a problem for the use of ditransitive verbs which describe spatial movement and change in Chinese, since these verbs select three obligatory arguments: a subject, an object and a phrase indicating final location[14]. In the following examples, 4.7a is grammatically correct, but 4.7b is not.

4.7a 我　把　书　　放　在　　桌子　　上　了
　　　Wǒ　bǎ　shū　fàng　zài　　zhuōzi　shàng le
　　　I　　BA　book　put　at　　　table　　top　PFV
　　　'I put the book on the table'

4.7b* 我　　放　书　在　　桌子　　上　　了
　　　Wǒ　fàng　shū　zài　　zhuōzi　shàng　le
　　　I　　put　book　at　　　table　　top　　PFV

In this case, *BC* still cannot be considered an obligatory structure to use since the OSV structure is an alternative grammatical expression (see examples 4.7c). However, using OSV structure will make the topic change from S to O. Therefore, although *BC* is not obligatory for the specific situation it is the only grammatical choice to express that meaning under the situation where the doer of the action is the topic, S. Because of this pragmatic constraint, Chinese linguists consider this type of *BC* an obligatory type (see Yuehua Liu, et al., 2001).

4.7c 书　我　放　　在　桌子　　上　　了
　　　Shū wǒ　fàng zài zhuōzi shàng le
　　　Book I　put　at　table　top　PFV
　　　'The book, I put it on the table'

---

[14] The nominal phrase at the final position may express a range of other semantic relations (e.g., temporal). In this study I am considering only the locative meanings expressed with this structure.

Aspectually, the function of the typical *BC* is to signal a complete process of change which has an end point (telic), has been finished or will be finished at an understood reference time (perfective) (see Yang, 1998a, 1998b) or is a bounded event (see F.-h. Liu, 1997). More specifically, the template of *BC*, 'S + BA + NP + VP' can be interpreted in semantic terms. The subject of *BC* can have several thematic roles, such as agent, causer, and affecter, because it functions as a dynamic source which executes the action V affecting the O. The function of *BC* can be depicted as that the doer (S) executes an action (denoted by V) which makes the referent of the O (*ba*-NP) undergo a kind of change or impact and achieve a final state expressed by R (see, Sun, 2006)**.** The *ba*-NP refers to a specific entity which is affected.

In a discourse context, S in *BC* can be considered as a discourse topic which serves as the subject or doer of serial actions in successive clauses**,** while the *ba*-NP (O) is the sentence-topic, that is, the referent about which something is being said in the current clause. In the example of a topic chain[15] (example 4.8), S (i.e., *ni* 'you') in the first clause controls multiple successive clauses, so it can be omitted in the second clause as a phenomenon of anaphora. In contrast, the *ba*-NP (i.e. O), *shui* 'water' in the second clause, *BC*, cannot be omitted but rather must be considered as the actual topic of the second clause.

4.8 你　先　　拿　一　个　杯子，然后，把　水　　倒　在　杯子　里

　　Nǐ　xiān　ná　yí gè　bēizi　ránhòu bǎ　shuǐ　dào　zài bēizi　lǐ

　　You first　get a CL　cup,　then　BA　water pour　at　cup　inside

　　'You first get a cup, and then pour some water into the cup'

4.2.5 Conditions of use of *BC*

Researchers' interest in *BC* may also be motivated by the difficulty in finding obligatory conditions of use of *BC*.

The selection of word order structures usually depends upon multiple factors. Kruijff *et al.* (2001) identified an inventory of factors that determine the expression of a specific word order structure, including information structure, syntactic structure, intonation, rhythm, and style. *BC* serves as a good example of these word order structures.

---

[15] According to Tsao (1990), the topic chain is a stretch of discourse headed by one or more topics, which are followed by one or more comment clauses. A discourse unit is equivalent to the English surface sentence.

According to L. Jin (1997), the selection of *BC* is determined by a battery of factors which form a hierarchy in order from obligatory to optional constraints. The linguistic factors corresponding to Kruijff *et al.*'s terms are shown in the parentheses below:

1. *yuyi shang de biaoda yaoqiu* , 'expressive requirements in meaning' (semantic structure).
2. *jufa de qiangzhi xing (bixu peihe pianzhang shang de xuanze)* , 'syntactic constraints (choice limited by obligatory collocations) ' (syntactic constraints).
3. *pianzhang shang de xuanze*, 'discourse constraint**s**' (contexts of use).
4. *shuohua ren de yuyi zhongxin* 'the speaker's focus in meaning' (information structure).
5. *shuohua ren de fengge he aihao* , 'the speaker's personal style and preference' (intonation, rhythm, and style).

From this list, it can be seen that the selection of *BC* is determined by more factors (e.g., semantic structure, contexts of use) than Kruijff *et al.* identified. Accordingly, the determination of the obligatory contexts of using *BC* is beyond merely the linguistic domain. Indeed, the speaker's focus of information or expression exists prior to the generation of sentences. Thus, it would be difficult to judge whether or not a word order is selected adequately without relevant information because the same linguistic context may allow different word order structures to convey similar basic meanings. This may have contributed to the difficulty in creating the obligatory contexts for using *BC*, particularly for the types of *BC* which can be transformed into other word order structures.

After comparing the use of three word orders (the canonical post-verbal form, the *ba* form, and the topicalized preposed form) by native speakers of Chinese in both spoken and written sources, F.–h., Liu (2007) argues that the choice of the *ba* form depends on multiple factors, including information status, weight, and topicality. The *ba* form is more likely to be used under two conditions: (a) when the *ba*-NP carries old information but is less topical than the subject, (b) when the object is new and not light. She also raises doubts on the *ba* -NP's role as a topic in discourse. It should be noted that in F.-h. Liu's study whether information is old or new depends on whether a linguistic form denoted by *ba*-NP has been mentioned in the discourse context rather than whether the speaker assumes the hearer knows the entity or not, which is the position assumed by this study.

4.3 Selection of target types of *BC*

The previous section introduced *BC* and discussed the reasons why *BC* has attracted the attention of Chinese linguists. Linguists' concerns only constitute part of the reasons for selecting *BC* as the target feature of this research. The following sections discuss additional reasons, elaborate the rationales and procedures for selecting the target types of *BC*, and illustrate the linguistic characteristics of the selected *BC*.

4.3.1 Reasons

In addition to the linguistic interest illustrated above, there are two more reasons for selecting *BC* as the target feature of this study. First, *BC* has been identified as one of the most difficult grammatical features for L2 learners of Chinese to acquire (see D. Li, 1996; W. Lü, 1994; J. Shi, 2006). Thus, researching the interlanguage of *BC* is expected to have pedagogical implications. Second, variation in the interlanguage of *BC* is still under-researched although the L2 acquisitional order of *BC* has received considerable attention (e.g., Cheng, 2006; Xiaoping Gao, 1999, 2008; H. G. Jin, 1993; Y. Li & Deng, 2005). This research intends to fill this gap.

However, due to limited space and time, this research focuses on only two of more than 30 types of *BC* identified in previous studies (see Appendix 1 for the classifications of *BC* in previous studies). This decision was made for three reasons. First, the classification of *BC* in previous studies varied depending on researchers' perspectives and research paradigms. For instance, distinguishing criteria inevitably differ between semantic typologies (W. Lü, 1994) and syntactic ones (Zhao, S. Liu, & X. Hu, 1997). Even on the basis of the same theoretical basis (e.g., semantic criteria), typologies vary between studies (see L. Chen, 2005; W. Lü, 1994; Xiong, 1996). Second, in accordance with the general curriculum of L2 Chinese, *BC* is normally first taught at an upper elementary level. Some types of *BC* are presented quite late and some never appear at all. Thus, the presentation order in teaching may have an effect on learners' acquisition order of various types of *BC*. Third, some types of *BC* are seldom used even by native speakers in spontaneous production, let alone by L2 learners. To elicit sufficient data, it is necessary to consider the types of *BC* that are taught early and used frequently.

4.3.2 Criteria and procedure

The two types of *BC* were chosen as target structures in this study on the basis of five criteria: frequency of use, high productivity, prototypicality, presentation order, and early acquisition by L2 learners.

4.3.2.1 Frequency of use

The criterion of the frequency of use indicates that only the types of *BC* used most frequently by native speakers would be targeted. To get the necessary information on the frequency of use, four studies (i.e., L. Chen, 2005; Cui, 1995; W. Lü, 1994; Zhao, et al., 1997) were consulted. Among them, the classification in W. Lü (1994) and Chen (2005)'s studies were based on semantic grounds, whereas that in Cui (1995) and Zhao *et al.*'s (1997) studies followed syntactic grounds. Consequently, these studies found that the types of *BC* used most frequently differed. These studies and the most frequent types of *BC* they found are briefly reviewed below.

Cui (1995) calculated the frequency of use of syntactic types of *BC* based on 1265 tokens collected from a modern novel, *Honglou Meng*, 'A Dream of Red Mansions (DRM) ' , and 3361 tokens obtained from a contemporary novel, *Nanren de Yiban Shi Nüren*, 'Woman – Half of a Man (WH)'. In Cui's study, *BC* was divided into 2 general categories and 9 subcategories (see Appendix 1). In both literary works the most frequently used *BC* were *BC* containing a prepositional phrase complement (*BCPP*), *BC* containing a directional complement (*BCDC*) and *BC* containing a resultative complement (*BCRC*), although the order of their frequency differed, with *BCDC* (31.86%) > *BCRC* (26.8%) > *BCPP* (19.84%) in the modern novel (*DRM*) and *BCPP* (42%) > *BCRC* (24.38%) > *BCDC* (23%) in the contemporary novel (*WM*)('>' indicates a higher frequency). This shows a dramatic increase of the use of *BCPP* in contemporary Chinese despite a decline in the use of *jiang* – a formal variant of *ba*. It should be noted that Cui distinguished these *BC* types by listing examples. His *BCPP* type includes an example of *BC* containing a phrase consisting of a directional verb and a noun phrase, for example, *jin* 'in/into' *qingqing de ba ta la jin wode huaili* 'lit. gently BA her pull into my arms' besides those containing prepositions *zai* 'at', *dao* 'to', *xiang* 'toward', and *gei* 'to' (animate recipient).

Zhao *et al.* (1997) calculated the frequency of use of 32 syntactic types of *BC* based on 702 *BC* tokens collected from the *Sentence Pattern Corpus of Primary School Textbooks* in Mainland China. The three types of *BC* that appeared most frequently are shown below with the frequencies in parentheses.

I. S + BA + NP1 + V + *zai* 'at' / *dao* 'to' / *gei* 'to' (animate recipient) /*cheng* 'as'/ 'into' + NP2 (31.34%)

II. S + BA + NP1 + V + Compound directional complement (13.82%)

III. S + BA + NP1 +V+ V$_{di}$ + locative NP2 (10.83 %)

In Type II, the compound directional complement consists of one of 7 single directional verbs and a directional verb, *lai* 'come' (indicating 'towards the speaker') or *qu* 'go' (indicating 'away from the speaker') but not *qiqu*. The frequency of the sentence structure, S + BA + NP1 +V+ V$_{di}$ + NP2 + *lai/qu* 'come/go' (2.28 %) was calculated separately and was not included in the above three types.

Both W. Lü (1994) and Chen (2005) examined semantic types of *BC*. Nevertheless, their semantic classifications were not exactly identical. W. Lü (1994) identified 6 general semantic types and 19 subtypes of *BC* by considering common semantic meanings of 1094 *BC* tokens extracted from a database of 530,000 words[16]. The most frequent subtypes of *BC* were *BC* containing a prepositional phrase (*PP*) which starts with *zai* 'at', *dao* 'to', *gei* 'to' (animate recipient), *xiang* 'toward', *ru* 'into' (27.8%), *BC* containing a resultative complement (*RC*) (23.3%), and *BC* containing a directional complement (*DC*) (19.9%) . According to W. Lü's definition, *BCPP* belonging to W. Lü's first semantic type refers to something specific changing its location or its relationships due to the effect of action. *BCRC* and *BCDC* belonging to her second semantic type refers to something specific changing or producing some results due to the effect of action. In contrast to Cui's classification, W. Lü considered the *BC* containing a phrase of a directional verb and NP (e.g., *yingai ba ta gan hui yangjuan li qu* 'lit. [You] should BA it lead back sheepfold in go ' '[You should] lead the sheep back into the sheepfold' as *BCDC* rather than *BCPP* because a directional verb was involved.

---

[16] W. Lü (1994) did not provide detailed information on the database she used.

Chen (2005) developed a semantic typology of *BC* through analysing 7710 *BC* tokens obtained from the *Academic Sinica Balanced Corpus of Mandarin Chinese* and the *Corpus of United Knowledge*, the databases collecting the written data used by Mandarin speakers in Taiwan. Chen identified 9 general semantic categories (see Appendix 1 for the frequency of each subcategory of *BC*). The three types of *BC* used most frequently were Resultative *BC* (35.85%) > Target one (22.95%) > Directional *BC* (16%). Her type of Target one actually is the *BC* containing a prepositional phrase complement (PP) which starts with *zai* 'at', *dao* 'to', and *gei* 'to' (animate recipient). Her subtype of Directional *BC* does not contain the phrase of a directional verb and NP.

In summary, based on the data from the four studies discussed above, these three types of *BC* - *BCPP* containing a prepositional phrase, *BCRC* containing a resultative complement, and *BCDC* containing a directional complement - were the most frequently used by native Mandarin speakers. However, the three types of *BC* covered different subtypes of *BC* tokens in the four studies. For example, some of the studies counted *BC* containing *xiang* 'toward' as *BCPP* (e.g., Cui, 1995, W. Lü, 1994), while others did not (e.g. Zhao et al., 1997; Chen, 2005). Also, a *BC* whose complement contains a directional verb, *jin* 'in/into', was considered as *BCPP* in Cui's study (1995) but a *BCDC* in others (e.g., W. Lü, 1994). Table 11 shows a comparison of the prepositions in *BCPP* across the four studies. Because the classifications in the four studies are not identical, a direct frequency comparison of the use of the three types of *BC* is not possible. For this reason, and because this study will focus only on *BCPP*s using *zai* 'at' and *dao* 'to', it is necessary to reclassify *BC* for the purposes of this study.

Table 11: Prepositions in *BCPP* in the Four Studies

Studies	Prepositions in PP	
	Common prepositions	Different prepositions
Cui (1995)		*xiang* 'toward', *jin* 'in/into'
Zhao *et al.* (1997)	*zai* 'at', *dao* 'to', and *gei* 'to'	*cheng* 'as'/ 'into'
W. Lü (1994)	(animate recipient)	*xiang* 'toward', *ru* 'into' (27.8%)
Chen (2005)		

Note. *BCPP* = *BC* containing the verb with a prepositional phrase complement, *= *BCPP* is labelled as Target One in Chen's study.

4.3.3.1.1 The classification of BC in the current study

It is possible to re-classify *BC* according to syntactic criteria and actual use for the purposes of the current study because in these studies the general semantic types were comprised of basic syntactic subtypes. The classification of the three types of *BC* in the current study was based on the previous studies mentioned above (see Table 12 for syntactic structures and examples).

Table 12: The Classification of *BC* in the Current Study

Type	Syntactic structures	Examples
BC1a	S + BA + NP1 + V + *zai* 'at' / *dao* 'to'+ N2 (+ Locative N)	4.9 把 书　放　到　书包　　里 Bǎ shū　fàng dào　shūbāo　　lǐ BA book put　to　book bag inside 'Put the book into the book bag.'
BC1b	S + BA + NP1 + V + V$_{di}$ + NP2 (+ Locative N) + *lai* 'come'/*qu* 'go'	4.10 把 书　　放 进　书包 (里去) Bǎ shū　fàng jìn　shūbāo (lǐ/qù) BA book　put in　book bag (inside /go) 'Put the book into the book bag'.
BC2a	S +BA + NP1 + V + single V$_{di}$	4.11 把 衣服　　穿　　上 Bǎ yīfú.　　chuān　shàng BA clothes　put　on 'Put the clothes on'.
BC2b	S +BA + NP1 + V + compound V$_{di}$	4.12 把 书　　放　　进去 Bǎ shū　fàng　jìnqù BA book　put　in go 'Put the book inside'
BC3a	S +BA + NP1 + V + V	4.13 把 书　　打开 Bǎ shū　dǎ kāi BA book beat-open 'Open the book'
BC3b	S +BA + NP1 + V + SV	4.14 把 衣服　　洗　干净 Bǎ yīfú.　　xǐ　gānjìng BA clothes wash　clean 'Wash the clothes (effectively).'

While most previous classifications of *BC* were based on double criteria, namely, word classes (e.g., directional verbs, stative verbs) and syntactic structures of the post-verbal constituents (e.g., a prepositional phrase), the current study classifies *BC* depending upon whether there was a nominal argument, a noun phrase (NP), present at the post-verbal position, although this study only considers NP indicating locations. Therefore, the *BC* containing a phrase of a directional verb and a NP is considered one subtype of *BC1*.

90

The primary difference between *BC1* and *BC2* is that the post-verbal constituent contains a NP in *BC1* but not in *BC2*. The rationale for the current classification is that native speakers produced *BC1a* (example 4.9) and *BC1b* (example 4.10) at the same discourse context. That is, these two structures contain same pragmatic and expressive functions. Due to the interchangability between *BC1b* (example 4.10) and *BC2b* (example 4.12) in discourse contexts, they are considered as variants under the condition that the noun phrase (*NP2*) (e.g., *shubao li*/'lit. book bag inside') in the complement of *BC1* in example 4.10 has been clarified in the context and does not need to be mentioned again.

BC1 indicates that S executes an action which makes the O move to a new location. *BC1* consists of two subtypes: *BC1a* (example 4.9) and *BC1b* (example 4.10). In *BC1a* the post-verbal constituents are comprised of *zai* 'at' / *dao* 'to' and a NP indicating a location, while in *BC1b* the post-verbal constituent consists of a directional verb (*Vdi*) and a noun phrase, as well as a locative noun (LN)[17] or another directional verb *lai* 'come'/*qu* 'go' optionally appearing at the final position.

BC2 conveys the semantic meaning that S executes an action which makes the O (i.e., *ba-NP*) undergo a change in direction of movement. *BC2* is also comprised of two subtypes: *BC2a* (example 4.11) and *BC2b* (example 4.12). The directional complement at the post-verbal position consists of one of 9 single directional verbs in *BC2a*[18] and one of 13 compound directional verbs in *BC2b*. The compound directional verb is comprised of one of 7 single directional verbs[19] and *lai* 'come' or *qu* 'go' but not *qiqu* 'rise-go'.

BC3 describes the process by which the action of S, denoted by the main verb V, changes the referent of ba-NP into a final state denoted by a resultative complement. *BC*3 also includes

---

[17] Locative nouns in Chinese are a subcategory of nouns which refer to directions and positions. The locative nouns are a closed word class, generally including the following single locative nouns (i.e., *dong* 'east' *xi* 'west' *nan*, 'south '*bei*' north' *shang* 'top' *xia* 'bottom', *zuo* 'left', *you* 'right', *li* 'inside', *wai* 'utside', *zhong* 'middle',,*pang* 'beside', *nei* , 'inside') and compound locative nouns (which consist of a single locative noun and a prefix (i.e., *zhi-, yi-) or a suffix (i.e., '-bian, -mian, -tou').

[18] The 9 single directional verbs are *shang* 'up', *xia* 'down', *jin* 'in/into', *chu* 'out', *hui* 'back', *guo* 'past', *qi* 'rise', *lai* 'come' and *qu* 'go'.

[19] The 7 single directional verbs are *shang* 'up', *xia* 'down', *jin* 'in/into', *chu* 'out', *hui* 'back', *guo* 'past', *qi* 'rise'.

two subtypes: *BC3a* (example 4.13) and *BC3*b (example 4.14). The resultative complement is comprised of an action verb (see Appendix 2 for examples) in *BC3a* but a stative verb (example 4.14) in *BC3*b (see Appendix 3 for examples). It should be noted that *BC3* containing a resultative complement here is only one type of *BC* containing a resultative verbal complement (RVC) defined in F.-h. Liu's (1997) study, which includes *BC* containing a directional complement (e.g. *qilai* 'rise-up').

4.3.3.1.2 Frequency of use of the three types of BC in my classification

Because previous studies calculated the frequency of use of various types of *BC* based on different classifications and databases, it is impossible to get an accurate frequency of use for the targeted types of *BC* in my classification based on their findings. Thus, I calculated the frequency of use of the three types of target *BC* based on 2136 *BC* tokens collected from written texts corresponding to spoken style in the *Lancaster Corpus of Mandarin Chinese* (LCMC). This corpus was used for three reasons. First, there is a lack of free speech corpora of Mandarin Chinese that include sufficient *BC* tokens produced by native speakers[20]. Second, the LCMC corpus is constructed by written texts of Mandarin Chinese published in Mainland China, where the native speakers in the current study originate from. Third, the texts in the LCMC written corpus are categorised according to literary genres. This makes it possible to identify the written texts corresponding to speech styles (e.g. fiction) and to extract the *BC* tokens.

As Table 13 shows, the frequency of use of the three target types of *BC* follows the order: *BC1* (n = 563*)* > *BC2* (n = 322) > *BC3* (n = 244). According to criterion one, frequency of use by native speakers, *BC1* and *BC2* were selected as the candidates of the target feature of this research.

---

[20] At the time of the writing, little spoken data was available, even though there are more open sources lately.

Table 13: Frequency of Use of the Three Types of *BC* in the LCMC

Type	Definition	Formula (S + BA + NP + V + R)	Frequency
BC1	Locative nominal phrase type (LN)	R= a proposition *zai* 'at' /*dao* 'to' + NP (PP) R = V_{di} +N+ (locative noun /*lai* 'come'/*qu* 'go')	563
BC2	Directional verbal complement type (DC)	R = a single directional verb R= a compound directional verb	322
BC3	Resultative complement type (RC)	R= an action V R = a stative verb	244

4.3.2.2 High productivity

A second criterion takes into account the productivity of *BC*. According to this criterion, the target types of *BC* should be highly productive rather than a formulaic or fixed expression. Productivity has been addressed mainly in morphology according to the degree to which native speakers use a particular grammatical process for the formation of *novel structures* (Katamba, 2004). Some studies calculated the productivity of syntactic structures by using a formula for calculating the productivity of morphological words: spontaneous generation of items not encountered before, regular combination of a comparable base, and a transparent production whose meaning can be identified from the base and the formation (Zeldes, 2009). In the current study, however, the productivity of the three types of *BC* is considered in terms of how many alternatives are possible in the language. At least three factors contribute to this issue, namely, 1) whether the post-verbal constituent (R) contains an open or closed class, 2) how sensitive or selective the structure is to the range of main verbs that are possible, 3) how easily the structure is learned as an item or a rule (Skehan, 1998).

As Table 13 shows, the complement in *BC1* consists of one of two prepositions (only two possible prepositions, *zai* 'at' and *dao* 'to', are of concern in this study) and seven single directional verbs[21], a noun phrase that is an open class, a locative noun as well as an optional *lai* 'come'/*qu* 'go'. Possible main verbs can be selected from a large range of verbs from ditransitive verbs to transitive verbs. The structure of the complement of *BC1* (P/V_{di} + NP_2 + [LN]/ [*lai/qu*]) is so fixed that it is easy to be learned as a rule.

---

[21] See note 17.

The directional complement in *BC2* consists of one of 9 single syllable directional verbs and 13 compound directional verbs[22], which are a closed word class. Although the structure of VR is rather simple, the collocation of the main verb and directional complement seems subject to item-based learning due to the lexical meaning of the directional verbs. Some collocations are subject to conventional word combinations and fixed as idiomatic expressions (e.g., *chuan shang* 'put on'). Thus, the productivity of *BC2* is limited.

The resulative complement in *BC3* consists of one of action verbs and stative verbs which are an open class. However, the collocation between the main verb and the complement seems subject to item-based rules. For example, among the action verbs and stative verbs that serve as a resultative complement (*RC*) of 239 *BC3* tokens found in the LCMC (see Appendix 2 and Appendix 3 respectively), the most frequent action verb is *kai* 'open', appearing in 31 of 141 tokens. The most frequent stative verb is *hao* 'good', occurring in 35 of 107 tokens. Of the 141 action verbs which serve as a *RC*, 21 occurred only once, while of the 107 stative verbs which serve as a *RC*, 31 occurred only once. Thus, *BC3* is considered of low productivity.

I also counted the number of verbs and stative verbs (which is similar to adjective in English in meaning) which can serve as a resultative complement (RC) in the *Dictionary of Collocation of Verb and Resultative Complement* (1987). This dictionary collected a total of 103 action verbs and 119 stative verbs which can serve as a complement in *RC*. Among them, the verbs or stative verbs which could serve as a RC in the *BC* are much fewer. That means that the action verb and stative verb which could serve as a RC in the *BC* are relatively closed classes, respectively. They are also selected by the main verb. Sometimes the collocation of the main verb and RC is so fixed that even native speakers consider it one word, for example, *da- kai* 'beat-open' in *ba shu da kai* 'lit. BA book beat-open' / 'open your book(s)' (example 4.13). Therefore, the collocation of V and RC in *BC3* is predicted to follow exemplar-based rules and item-based learning.

---

[22] The 13 compound directional verbs are constructed by combining one of the 7 single directional verbs, *shang* 'up', *xia* 'down', *jin* 'into', *chu* 'out', *hui* ' back', *guo* 'across', *qi* 'upward' and one of the 2 single directional verbs *lai* 'come' (indicating towards the speaker) or *qu* 'go' (indicating away from the speaker' except for *qiqu** 'up +away'.

According to this criterion, highly productive types will be selected as the target feature. Since *BC3* has lower productivity than *BC2* and *BC1*, this research focuses on only *BC1* and *BC2*.

4.3.2.3 Prototypicality

A third criterion is prototypicality, that is, the prototypes of *BC* (i.e., a representative for a large range of forms of *BCs*) will be targeted. This criterion was considered because prototype schemas (i.e. L2 learners' initial representation of new forms) can be a psycholinguistic source of variation in interlanguage (Adamson & Elliott Jr, 1997). It can be assumed that the prototype might be easier to learn than a non-prototype because learners' production will be based on comparison to the prototype until the correct target form is acquired.

However, criteria for determining the prototype of linguistic structures differ. There are at least three kinds of criteria. First, Bates and MacWhinney (1982, 1987) indicated that all of the critical predications of prototype theory must satisfy three requirements: 1) family resemblance, that is, members of the family should 'best fit' or have 'maximum overlap' with the prototype; 2) heterogeneous membership, that is, a grammatical category will contain members that overlap with the prototype but not with one another; and 3) maximum distance from other categories, that is, prototype should have maximum overlap with the category that is ultimately assigned and minimum overlap with competing categories.

Second, G. Hu (2002) summarised three criteria for identifying the prototypicality of linguistic structures on the basis of review of previous studies: 1) a minimum distance from members of the same category but at a maximum distance from members of a contrasting category (Rosch, 1975, 1978; Taylor, 1995, 1998); 2) the relative frequency of members of a category (Barsalou, 1985); and 3) members that are frequently encountered tend to be viewed as more prototypical of their category (de Villiers, 1980; Smith & Medin, 1981).

Third, Taylor (2003) stated that the prototype of linguistic categories possibly originates from: 1) the inherent properties of human perception, 2) the frequency of encounter as a representative exemplar of the category; 3) the order of learning, and 4) 'a large number of attributes of the category and a smallest number of attributes of other categories' (pp.55-58).

The prototype of *BC* were identified based on the common components of the above three kinds of criteria, which are: 1) representation of a larger number of members in the category, but not overlap one another, 2) maximum distances from competing categories, 3) the frequency of encounter. Two psycholinguistic factors related to learning process in Taylor's criteria (i.e. inherent properties of human perception and the order of learning) were not considered here because prototypicality only focuses on inherent linguistic properties of grammatical structures. The order of learning will be examined separately in Section 4.3.3.5.

According to the first of these criteria, it is hard to determine the prototype of *BC*. This is because the three types of *BC* illustrated above share the same structure, $S + ba + N + V + R$ and the differences between them lie in R. Thus, they do not overlap each other and it is hard to say which covers a larger number of *BC* types. Productivity would play a role here.

According to the second criterion, *BC1* has the maximum distance from its competing word order, SVO, because both *BC2* (including a directional verbal complement) and *BC3* (including a resultative complement) can be transformed into a SVO structure, but *BC1* cannot. Clearly, *BC1* seems most eligible to represent all types of *BC*.

According to the last criterion, *BC1* is encountered most frequently in LCMC (see Table 13), so it appears to be the prototype of *BC*.

4.3.2.4 Presentation order in pedagogy

The fourth criterion is concerned with the influence of presentation order in pedagogy. According to this criterion, the types of *BC* that are the common focus of various elementary Chinese textbooks were targeted. With regards to the influences of instructional orders, there are two camps of opinions. Researchers in one position assert that the input order of the structural syllabus does not affect the actual acquisitional order (Hyltenstam & Pienemann, 1985; Krashen, 1982; Pienemann, 1998), whereas those with the opposing view provide evidence to show that instruction does make a difference (Doughty, 1991).

In consideration of the impact of presentation order, the presentation orders of *BC* in the five elementary L2 Chinese textbooks extensively used by the participants of this study were examined (i.e., J. Xu & Ren, 2004; Kang & Lai, 1999; X. Liu, 2003; Sanders & N. Yao, 2009; T.-c. Yao *et al.,* 2005).

Table 14: Presentation Orders of *BC* in Elementary Textbooks of L2 Chinese

Authors	Textbooks	The lessons introducing *BC*	The types of *BC* presented	Presentation order [1]
J. Xu & Ren (2004)	Boya Chinese II	Lesson 41	A) **S+BA+N1+V+ zai 'at'/dao 'to'**/gei 'to'/cheng 'into' **+N2** B) **S+BA+N1+V**+C (**SV/ V$_{di}$** /frequency/ zhao	BC1=BC2=BC3
Sanders & N. Yao (2009)	Fundamental Chinese Spoken Chinese	Chapter 17	A) **S+BA+N1+V+ zai 'at'/dao 'to'+N2** B) S+BA+N1+V+ cheng 'into'+N2 C) Finish dealing with sth. **S+BA+N1+V**+C (**SV/V$_{di}$**) D) Cause somebody to enter into an extreme emotional or exhausted state (more details in Chapter 20)	BC1=BC2=BC3
T. Yao *et al. (*2005)	Integrated Chinese Level I	Lesson 16	A) **S+BA+N1+V+V$_{di}$** B) **S+BA+N1+V+V$_{di}$+N2+(qu 'go')**	BC2=BC1
Kang & Lai (1999)	Chinese Conversation 301 sentences	1. Lesson 34 2. Lesson 37	1. **S+BA+N1+V+V$_{di}$./V./SV.** 2. A) **S+BA+N1+V+ zai 'at'/dao 'to'**/jin 'in'+**N2+ (locative noun)** B) S+BA+N1+V+dative N2	1.BC2=BC3> 2.BC1
X. Liu (2003)	New Practical Chinese Reader I	1. Lesson 16 2. Lesson 18	1. S+BA+N1+V+LE 2. A) S+BA+O+ V+ dative N B) **S+BA+O+ V+ lai 'come'/qu** 'go'+(LE)	2. BC2

Table 14 shows the presentation orders of the various types of *BC* in the five textbooks. The *BC* structures in bold are related to the *BC* types in the current study. The three types of *BC* classified in the current study were presented earliest in nearly all five textbooks although their presentation orders were not exactly the same. Whereas some textbooks (Sanders & N. Yao, 2009; J. Xu & Ren, 2004) intensively presented multiple structures in one lesson, others displayed *BC2* or *BC3* prior to *BC1* (Kang & Lai, 1999; X. Liu, 2003) following a step-by-step procedure. In short, according to the criterion of the early presentation, all three types of *BC* in the current study could be targeted.

4.3.2.5 Acquisition orders of the three types of *BC*

The last criterion is the acquisitional orders of *BC* by first language (L1) learners and by second language (L2) learners. This criterion was established for two purposes: 1) to take into account psycholinguistic constraints on actual learning difficulty; and 2) to facilitate data collection. Based on my observation in teaching and consensus in the literature (X. Liu, 2003), *BC* is seldom used by L2 learners. Thus, on the basis of the assumption that the earlier a type of *BC* is acquired the more chances L2 learners would use it, it was expected that it would be easier to collect *BC* tokens from learners, particularly those at a lower level of proficiency.

4.3.2.5.1 Acquisition orders of the ba constructions by L1 Children

Four studies have investigated the acquisition orders of *BC* by L1 learners of Chinese (Cheung, 1992; P. Li, 1993; Xiangnong Li, Zhou, & Kong, 1990; Tse, Tang, Shie, & Li, 1991). Xiangnong Li *et al.* (1990) investigated the emergence order of the 9 syntactic structural types of *BC* consisting of 17 subtypes. The subjects were 90 Chinese children (male 43, female 47) in Anhui, China. They were randomly selected from a kindergarten and evenly divided into 9 age groups ranging from 1.5-5 years with an interval of 6 months. Both spontaneous and elicited data were collected from each child in two ways (i.e., observation and interview prompted by topics) for three times with 1.5 hours each time. The analysis of obtained 843 *BC* tokens showed that the following three types of *BC* appeared earliest (i.e., at 2 years of age). It can be seen that all three types of *BC* in the current study emerged earliest in the L1 acquisition order of *BC* according to Xiangnong Li *et al.*'s (1990) study.

4.15 BC+ NP   + V.   + V_{di}.   (BC2)

     把   这个   手套   搞   下来

     Bǎ   zhège   shǒutào   gǎo xiàlái

     BA this CL glove   take down-go

     'take down this glove'

4.16 BC+ N.　　+ V. + V. /Adj.[23]　　　　　　　　(BC3)

把 这 球　 摔 掉

Bǎ zhè qíu shuāi diào

BA this ball throwoff

'throw away this ball'

4.17 BC +NP1 +　　　V. + zai 'at'/dao 'to'+NP2 + locative noun　　(BC1)

把　 手帕　　　 放　 到　　　　 口袋　 里边　　　 (好不好)?

Bǎ　shǒupà　　fàng　dào　　　　kǒudài　lǐbiān　　(hǎo bù hǎo`

BA　handkerchief　put　　to　　　　　　pocket　inside　　(alright or not)

'Put the handkerchief into the pocket, alright?'

把　 我 抱 到 这 里面 去

Bǎ wǒ bào dào zhè lǐmiàn qù

BA me carry to here inside go

'Carry me in there'

Cheung (1992) examined the acquisition of *BC* by focusing on the following questions. 1) How did children treat BA during their grammatical development? 2) What is the categorical status of BA in children's grammar? 3) Did children rely on the objective affectedness linking rule in the acquisition of *BC*? The subjects were 32 Mandarin speaking children (5-6 year olds) in Taiwan. The data were collected mainly from three experimental tasks: 1) picture comprehension; 2) sentence imitation, and 3) sentence production, besides spontaneous speech samples from a previous study. The primary results of the comprehension test suggested that the object affectedness rule alone does not determine the acquisition of *BC*. Findings from the imitation and comprehension tests and analysis of spontaneous speech samples suggested that *BC* is a verb-like construction in children's grammar. Findings based on the free speech and prompted speech data showed that the *BC* including a PP (prepositional phrase) and a RVC (resultative verb complement) were the most frequent *BC* patterns (see Appendix 4 for details). *BC* containing a RVC includes *BC2* and *BC3* clarified in this study.

---

[23] Original article used *xingrongci* 'adjective', which is called stative verb (SV) in this study.

In addition, Cheung (1992) also analysed spontaneous longitudinal data collected for three years at one-month intervals by Tse *et al.* (1991). The subjects in Tse *et al.*'s study were 21 Mandarin-speaking children (6 females and 4 males) at ages between 1 and 6 years in Taipei, Taiwan. Speech samples from 10 children who had attended more than 10 recording sessions were analysed. Cheung found that the *BC* with RVC was the most frequent BA pattern (see Appendix 5) among the 10 BA patterns targeted.

P. Li (1993) compared the acquisition of *ba* and progressive aspect marker *zai* by Chinese children in Beijing. The constraints of ba-NP (definiteness and specificity) and ba-VP (including resultative verbs and perfective aspect marker LE) were investigated. The subjects were 99 children with ages between 3 and 6 years. They were asked to describe 18 situations elicited by interview prompts and toys. The analysis using the CLAN programs (MacWhinney & Leinbach, 1991; MacWhinney & Snow, 1985) showed that 296 of 1107 sentences collected were the *ba* constructions and 90% of them were *BC* with a RVC. It should be noted that the *BC* containing a RVC included *BC2* and *BC3* in the current study because the RVC includes all verbal compounds at the post-verbal position.

These studies on L1 acquisition orders of *BCs* differed in many aspects such as the age groups and locations of the subjects, data collection methods ranging from cross-sectional experimental tests to longitudinal spontaneous speech data, data analysis methods including frequency (Cheung, 1992) and emergence order (Xiangnong Li, et al., 1990). However, these studies yielded almost the same order for the three types of the *BC* that concerns the current study, that is, *BC* containing a RVC (corresponding to *BC2* and *BC3* in the current study) and *BC* containing a PP (corresponding to *BC1a*) appeared earliest and most frequently in L1 learners' speech production.

4.3.2.5.2 Acquisition orders of the ba constructions by L2 learners

The developmental order of all types of *BC* has received considerable attention in TCSOL. The studies that have examined the acquisition of *BCs* by L2 learners number no less than ten (e.g., Cheng, 2006; Du, 2004; Xiaoping Gao, 1999, 2008; Huang & Yang, 2004; H. G. Jin, 1989; Y. Li & Deng, 2005; Z. Lin, 2001; S. Liu & Wang, 2003; Wu, 2001; Xiong, 1996; Yu, 2000; S. Zhang, 2002). However, some of them have concentrated on the acquisition order of constituents within the *BC* structure (Huang & Yang, 2004; H. G. Jin, 1993), influential

factors of the interlanguage of *BC* (Xiong, 1996), and the effectiveness of textbook exercises in eliciting *BC* (S. Liu & Y. Wang, 2003). This study is only concerned with the studies have examined the acquisition orders of the various types of *BC*s irrespective of the classifications of *BC* used.

Five studies that have examined the acquisitional order of *BC* types relevant to the three types of *BC* in the current study are reviewed below (see Table 15).

Table 15: Studies on the Acquisition Order of Different Types of *BC*

Study	Target feature	Subject	Tasks	Measure	Data	Main findings
Xiaoping Gao (1999)	17 subtypes of BC based on Zhao *et al.'* s (1997) structural categories	90 English, Japanese, Korean speaking learners in Beijing, China	1) Contextualised grammaticality judgements in the form of multiple choice 2) Translation	Accuracy Statistical analysis	Written	BC containing a PP > BC containing a DC and a RC
Z. Lin (2001)	15 subtypes of BC based on W. Lü's (1994) semantic categories	120 Korean learners in Beijing, China	1) Grammaticality judgements in the form of multiple choice 2)Translation	Accuracy	Written	BC containing a DC > BC containing a PP
Y. Li & Deng (2005)	15 subtypes of BC based on W. Lü' s (1994) categories	111 learners in mixed L1 backgrounds in Guangzhou, China	Contextualised sentence making	Accuracy	Written	*BC* containing a PP > other types of *BC* (e.g. containing a perfective marker LE) > *BC* containing a verbal complement indicating concrete meaning > *BC* containing a verbal complement indicating abstract meaning
Cheng (2006)	3 types of BC, containing 1) a locative PP, 2) a RC, and 3) a DC	30 Japanese, at elementary level in Wuhan, China	1) Grammaticality judgement tasks in the form of multiple choice 2) A sentence completion task	Accuracy Avoidance	Written	BC containing a locative PP> BC containing a RC and a DC
Xiaoping Gao (2008)	16 types of *BC*	1) 43 English speaking learners in New Zealand 2) 2 English and 1 Korean speaking learners	1) Cross-sectional acceptability judgements in the form of multiple choice 2) Longitudinal l interviews and oral narrative tasks	1) Accuracy 2) Frequency	1) Written 2) Oral	*BC* containing a PP complement > *BC* containing a RC> *BC* containing a DC

Xiaoping Gao (1999) examined the acquisition order of 17 types of *BC* by 90 L2 learners in mixed L1 backgrounds (i.e., Korean, Japanese and English) at three proficiency levels in the target language environment – Beijing, China. The categories of *BC* were based on Zhao *et al.*'s (1997) syntactic classification. The data was collected mainly from a two-part written questionnaire – contextualised grammaticality judgements in the form of multiple-choice and sentence translation. The acquisition order of *BC* was calculated by combining the rank orders of accuracy scores in the multiple choice and translation. The order of the three types of *BC* relevant to the current study was the *BC* whose complement contains a prepositional phrase (PP) (*BC1a* in this study) > the *BC* whose complement contains a directional verb (related to *BC1b* and *BC2* in this study) > *BC* containing a resultative complement (*BC3* in this study).

Z. Lin (2001) examined the acquisition order of 15 types of *BC* by 120 Korean adult learners of Chinese in Beijing. The types of *BC* were identified according to W. Lü's (1994) semantic categories. His data was collected from a multiple choice questionnaire similar to that used by Xiaoping Gao (1999) and sentence translation. He found that the accuracy order for the three types of *BC* related to the current study was *BC* containing a *DC* > *BC* containing a *PP* > *BC* containing a *RC* by combining the results in the two tests (i.e., *BCDC* > *BCRC* > *BCPP* in multiple choice and *BCPP* > *BCDC* > *BCRC* in translation).

Y. Li & Deng (2005) examined the accuracy order of 15 types of *BC* by adopting W. Lü's (1994) semantic categories. The participants were 162 adult learners of Chinese in mixed L1 backgrounds in Guangzhou, China, as well as 19 Chinese speakers as a control group. Their data was collected from a sentence-making test (i.e. making sentences according to linguistic contexts) in class. The analysis of the percentage of accurate use yielded the following order of *BCs*: *BC* whose complement contains a PP (related to *BC1a* in the current study) > other types of *BC* (e.g., containing a perfective marker, a reduplicated verb, and a measure word phrase) > *BC* containing a verbal complement indicating concrete meaning (related to *BC1b*, *BC2* and *BC3* in the current study) > *BC* containing a verbal complement indicating abstract meaning.

Cheng (2006) examined the acquisition order of two general types of *BC* by 14 Japanese learners of Chinese at an elementary level of proficiency. Written data was collected from three tests: 1) a multiple choice regarding the grammaticality of *BCs*; 2) a sentence

completion test, and 3) an error correction test. According to his calculations, the accuracy of *BC* containing a PP (related to *BC1* in the current study) scored higher than that of *BC* whose complement containing a directional verb and a stative verb (relevant to *BC2* and *BC3* in the current study).

Xiaoping Gao (2008) investigated the accuracy order of 16 types of *BC* by 43 learners of Chinese in New Zealand. The cross-sectional data was collected from 40 learners enrolled at Stage Two and Three Chinese courses using grammaticality judgments in a written questionnaire and analysed using obligatory analysis. Among the 16 target types of *BC*, *BC* including a PP received the highest accuracy score, followed by *BC* containing a RC and a *DC*. The longitudinal data was collected from 3 learners enrolled in Stage Two Chinese courses (two native speakers of English and one native speaker of Korean). The data were collected from oral interviews and oral narrative tasks for a period of 8 months at monthly intervals, and analysed using frequency analysis (R. Ellis & Barkhuizen, 2005). The frequency order lent support to the accuracy order found in the cross-sectional data.

The differences between the findings of these studies may be due to the following factors: the participants' L1s and proficiency levels in Chinese, the different settings of the studies, and the methods of data collection and data analysis. More specifically, the participants were speakers in mixed L1s in some studies (Xiaoping Gao, 1999, 2008; Y. Li & Deng, 2005) but native speakers of a single language in the others, such as Japanese in Cheng's study (2006) and Korean in Z. Lin's study (2001). All but one study (Xiaoping Gao 2008) were conducted in the second language setting, China. The participants were at mixed proficiency levels in most of the studies (e.g., Xiaoping Gao, 1999, Z. Lin, 2001, Y. Li & Deng, 2005) but at an elementary level of proficiency in Cheng's study (2006). Thus, whether or not these factors contributed to the variation in the learners' use of *BC* is one of the motivations of the current study.

In addition, the methodological limitations in these studies also motivated the current study. First, nearly all studies collected data from untimed grammaticality judgements (e.g., multiple choices) in written form, making sentences, filling in blanks, and translation. In these written tests learners could monitor their answers or complete the test by guessing. The translation test itself may cause L1 transfer. In addition, most of the studies did not contain

distractors and pilot questions to see whether the questions could effectively elicit target types of *BC*.

Despite the above discrepancies and the different typologies of *BC* adopted, these studies achieved considerable consistency in finding that the *BC* containing a PP was acquired earlier than the *BC* containing a verbal complement and other types of *BC*. More specifically, apart from Lin's study, the accuracy order of *BC* in the other four studies could be shown as *BC1* >*BC2* and *BC3* when using the labels of the current study. According to the criterion of the early acquisition by L2 learners, *BC1* and *BC2* were considered as the target structure of this research.

In sum, *BC1* and *BC2* were targeted through examining the above selection criteria[24].

4.4 Characteristics of the targeted *BC*

This section further illustrates the characteristics of the two target *BC* types in detail. *BC1* is characterised as *BC* containing a noun phrase, NP2, in the complement which indicates the locative destination of the movement of the *ba*-NP. *BC1* is divided into two subtypes: *BC1a* and *BC1b*. A structural difference between *BC1a* (see example 4.18) and *BC1b* (see example 4.19) is that the post-verbal components in *BC1a* consist of a coverb (also called a preposition) (i.e., *zai* 'at' or *dao* 'to') and a NP2 while those in *BC1b* consist of a directional verb and a NP2. The NP2 indicates a location, more specifically, the final position or destination of the movement of NP1 (i.e., ba-NP).

4.18. S BA NP1 **V P   NP2   LN**   **(PFV)**
      他 把   衣服   放 在   桌子   上   (了)
      Tā bǎ   yīfú. fǎng zài zhuōzi shàng   le
      He BA clothes place at table   top   (PFV)
      'He put the clothes on the table'

4.19. S BA NP1 **V Vdi NP2 (LN) (come / go)**   **(PFV)**

　　他　把　衣服　**拿**　**进**　房间　（里）（来/去）   （了）

　　Tā bǎ yīfú. ná jìn fángjiān lǐ lái/qù   le

　　He BA clothes carry enter room (inside)(towards/away from the speaker)   (PFV)

　　'He took/ brought the clothes into the room'

A second distinction between *BC1a* and *BC1b* is whether there is a need to include a locative noun (LN) (e.g. *shang* [*bian /mian /tou*] 'top', *li [bian /mian /tou]* 'inside') following NP2. In *BC1a*, it is necessary to contain a locative noun following NP2 unless NP2 refers to a location (e.g. *tushuguan* 'library'). In contrast, in *BC1b,* it is not necessary to do so.

A third difference between these two subtypes is the transformability into a SVO structure. Although neither *BC1a* nor *BC1b* can be transformed into a grammatical SVO with the same meaning (see example 4.20), *BC1b* can be transformed into a structurally correct serial verb construction (i.e., a string of verbs or verb phrases within a single clause) which contains two VO structures (i.e. SVOVO) (see example 4.21). This may be because *zai* 'at'/ *dao* 'to' in *BC1a* have generally lost their functions as action verbs, whereas the verbal functions of directional verbs remain**.** Due to its non-transformability, *BC1* is called an obligatory type under the condition that the agent of the action is the topic.

4.20. * S　　V_{dy}　　NP1　　**P　NP2　LN**　　　　　**(PFV)**

　　他　放　　衣服　**在　桌子　上**　　　　（了）

　　Tā fàng   yīfú.   **zài zhuōzi shàng**   le

　　He place   clothes **at   table   top**   (PFV)

　　'He put the clothes on the table'

4.21.? S　　V_{dy}　　NP1　　**V_{di} NP2**　　**(LN) (come/go)**　　**(PFV)**

　　他　拿　衣服　**进　房间**　(里)　(来/去)　　　了

　　Tā　ná　yīfú　**jìn fángjiān** (lǐ)　(lái/qù)　　　le

　　He carry clothes **enter room**　**(inside) (come/go)** (PFV)

　　'He took/ brought the clothes into the room'

BC2 is characterised as *BC* whose complement contains a single or compound directional verb indicating the direction of the movement of the ba-NP. *BC2* also consists of two subtypes. *BC2a* contains a single directional verb and *BC2b* a compound directional verb. Both types of *BC2* are shown in example 4.22. There are no major differences between *BC2a* and *BC2b* in their functions. Both can be freely transformed into a grammatical SVO structure (example 4.23) with the original meaning remaining. Therefore, the choice of *BC2* and SVO mainly depends on speakers' intentions and the requirements of the discourse context.

4.22. S   BA   NP   **V**   **(V$_{di}$)** **V$_{di}$**   (PFV)

他　把　衣服　**拿**　**(进)来**　　　（了）

Tā　bǎ　yīfú.　**ná**　**(jìn)lái**　　　le

He　BA clothes　**carry (enter) come**　(PFV)

'He brought the clothes in'

4.23. S   **V**   **(V$_{di}$)** **V$_{di}$**   NP   (PFV)

他　**拿**　**(进)来**　　　衣服　　　（了）

Tā　**ná**　**(jìn)lái**　　　yīfú.　　　le

He　**carry (enter) come**　clothes　　　(PFV)

'He brought the clothes in'

To sum up, the primary difference between *BC1* and *BC2* is that *BC1* can be considered as an obligatory type while *BC2* can be considered as an optional type. That is, *BC1* is subject to syntactic constraints and there is no alternative canonical SVO structure available in the context for *BC1*. In contrast, the choice of *BC2* is mainly determined by the speaker's intention of expression, discourse context and even the speaker's preference. *BC1* and *BC2* are related to each other pragmatically because *BC1b* and *BC2b* can be treated as variants under the condition where the post-verbal NP2 in *BC1b* is implied in the context but not explicitly mentioned.

4.5 Summary

This chapter has addressed the basic characteristics of *BC* and the selection and features of the target types of *BC*. Based on five criteria (i.e., high frequency of use by native speakers, high productivity, prototypicality, early presentation in textbooks, and the acquisitional

orders by L1 and L2 learners), two types of *BC* – *BC1*, a (locative) nominal *BC* whose complement contains a NP and *BC2*, a directional verbal *BC* whose complement consists of a directional verb but no NP – were selected as the target feature. Their characteristics were further clarified and compared.

The next chapter will report on the use of these target types of *BC* by native speakers of Chinese in a pilot study.

CHAPTER 5: PILOT STUDY WITH NATIVE SPEAKERS

5.1 Introduction

The previous chapter elaborated the process of selecting the target types of *BC* and their basic functions. This chapter reports on a pilot study with native speakers of Chinese, providing information on its purpose, participants, instruments, procedures, analysis and results, reliability and validity, and problems and changes to be made.

5.2 Purpose

A pilot study with native speakers (NSs) of Chinese was conducted in Auckland, New Zealand between May and July 2008. The overall objective of the pilot study was to establish the appropriate design, instruments, materials, and procedures for the main study. In addition, there were four specific objectives: 1) to ensure the instructions and materials were sufficiently clear and adequate to elicit the target types of *BC*, 2) to check the procedures and permit problems or difficulties to be detected, 3) to obtain baseline data for the main study, and 4) to examine the reliability of the proposed instruments.

5.3 Participants

5.3.1 Sample

Native speakers of Chinese were recruited in Auckland, New Zealand with reference to three criteria: 1) they could speak Mandarin in daily life when living in China, 2) they had grown up in Northern dialect areas (also called Mandarin dialect areas) in P. R. China (see the map in Appendix 6), and 3) they had no exposure to other language speaking countries before the age of 18.

These criteria were established on the basis of the following rationale. First, a large number of Chinese people speak their native dialects in daily life, particularly in dialect speaking areas. Although the majority of them can also speak Mandarin, their everyday use of their native dialects influences their use of *BC*. According to the findings of my two preliminary studies with native speakers, *BC* was used less often and/or in a different manner by the speakers of some dialects such as Cantonese. That is, under the discourse contexts where

most Mandarin speakers from North China used *BC*, those from South China (e.g., Cantonese speaking areas) seldom used this S BA OV structure. This was also in line with Du's (2004) observation that the use of *BC* by native Mandarin speakers from Mainland China and those from Taiwan differed. To avoid the influence of native dialects, therefore, only NSs who speak Mandarin in daily life in China were recruited. Second, because native Mandarin is based on northern dialects, people from these areas were required for the study. Third, some studies (e.g., Cook, 2003) showed early exposure to second language environments might potentially affect the use of the first language. To avoid the potential influence of L2 on L1, only Mandarin speakers who came to NZ after 18 years of age were recruited.

The Chinese native speakers were recruited by the placement of advertisements on local Chinese websites and by word of mouth in the Auckland Chinese community. Thirty respondents were contacted by email or by phone. Of them, 22 qualified respondents were provided with both English and Chinese versions of a Participation Information Sheet (PIS) (see Appendix 7) as an email attachment. The PIS indicated voluntary participation and anonymous responses, described the study and its potential benefits to second language learners of Chinese, and expressed appreciation by offering an incentive of a $10 gift voucher. After obtaining the participants' confirmation, one-on-one meetings were arranged.

5.3.2 Characteristics of native speakers

All 22 native Mandarin speakers originally came from Northern dialect or Mandarin dialect areas in China. They had lived in New Zealand for between 0.5 and 12 years (M = 6). There were equal numbers of male and female participants (11 of each). Their ages ranged between 20 and 45 years (M = 32). The daily language in Auckland was Mandarin for 12 NSs, both Chinese and English for 8 NSs, and English for 2 NSs. Ten participants (45%) were studying toward degrees or certificates at universities and 12 (55%) were gainfully employed. Their educational qualifications varied from graduates of high school in China to those with a doctorate. Eight held master's degrees or above, 10 bachelor's degrees, and 4 were studying towards Diplomas of English language. Their educational backgrounds were spread across Arts, Commerce, Science, Engineering, and Medical Science.

5.4 Instruments and Procedures

Five instruments[25] were utilised to collect data from native speakers of Chinese: 1) a background questionnaire, 2) an Oral Production Task prompted by Video clips (OPTV), 3) an Oral Imitation Task (OIT), 4) an Untimed Grammaticality Judgment Task (UGJT), and 5) a follow-up interview. The current study employed these instruments to collect data on participants' language background, and their oral and metalinguistic performance in using *BC*.

These instruments were administered in the following order: background questionnaire > the OPTV> the OIT > the UGJT> a follow-up interview. Each of these tasks was conducted in one-on-one interviews between the researcher and participant. The participants were first asked to read the Participation Information Sheet (see both English and Chinese versions in Appendix 7) and to sign the consent form (see both English and Chinese versions in Appendix 8). They were informed that the interview would be recorded and their responses would be kept anonymous. The participants were asked questions about their personal background based on a semi-structured background first questionnaire (See Appendix 9 for both English and Chinese versions). They were then asked to do the tasks in the order specified above. Finally, after completing all the tasks, the participants were asked about their perceptions of the difficulty of the tasks in a follow-up interview. All the oral data produced was recorded on a digital recorder and later transcribed by the researcher.

---

[25] Originally, six tasks were used to collect data. However, three of them —an oral production task prompted by interview topics (OPTI), a metalinguistic knowledge task (MKT), and a stimulated recall — were discarded for a number of reasons. First, in the OPTI participants provided different narrations responding to the interview question\s. This made the obligatory contexts involved in the narrative discourses incomparable. Second, the data in the stimulated recall was collected based on the data in the OPTI. Accordingly, the stimulated recall was eliminated as well. Third, the MKT was discarded because both native speakers and learners had no clear metalinguistic knowledge of *BC*. So it was hard for them to verbalise metalinguistic rules of *BC*. All native speakers claimed that they made judgments by feel, although some of them had made efforts to verbalise the rules based on their implicit knowledge. Most of higher proficiency learners indicated that they made a judgement by feel. The learners who claimed that they judged the sentences by rule just simply said that the rule was '*ba zi ju*'/the *ba* construction. Interestingly, the majority of lower proficiency learners claimed that they made judgement by rule or by guessing, however, their judgments were generally wrong. It seems that metalinguistic knowledge is not related to participants' linguistic competence. Therefore, this task was discarded.

The following describes the purpose of design and the content of each instrument and procedures for administering the instrument.

5.4.1 Background questionnaire

A background questionnaire was used to collect some personal information from the native speakers. The questionnaire (Appendix 9) contained both closed and open-ended questions regarding the participants' age, gender, the dialects and languages they used most frequently in China and in New Zealand, length of residence in New Zealand and other English speaking countries, home language, the languages spoken by the people whom the participants lived with, and their study majors, if any.

At the beginning of the meeting, the participants were asked to fill in a written version of the background questionnaire. The researcher checked their answers and asked further questions to elaborate upon the answers given in the open ended questions, if it was deemed necessary. During this discussion, the participants provided their personal account about what dialects and languages they spoke in China and in New Zealand. This oral conversation was recorded.

5.4.2 Oral production task prompted by video clips (OPTV)

An oral production task prompted by video clips (OPTV) was administered to elicit the actual use of *BC* in oral production. This more controlled oral production task was used because *BC* is a late acquired grammar feature and it is unreasonable to expect low intermediate learners to use the *BC* in free production. In this OPTV, silent video clips served as prompts (see Chapter 6 for a detailed explanation of the reasons for choosing this task). It was expected that participants would focus on the meaning conveyed in the video episodes rather than on linguistic form.

Because it was hard to find existing films or video clips to elicit a sufficient number of the target types of *BC*, my own video clips were produced by: 1) composing *BC* sentences for each type of *BC* using basic words[26], and 2) videoing my performance of the scenarios that each sentence depicted. The digital version of each video clip was approximately 4MB and 10 seconds long.

---

[26] These words have been learned by the learners in New Zealand. The wording of the items was checked by Chinese instructors at the University of Auckland.

Sixteen out of 24 video clips (8 for each type of *BC*) and prompt questions were selected as material for the pilot study after being trialled in two preliminary studies with native speakers. The first preliminary study (n = 10) found that the prompt questions —'*Ni neng shuoshuo fasheng le shenme ma*? / Could you please describe what happened on the screen?' and '*Pingmu shang nei ge ren zuo le shenme*? / What did that person on the screen do?' – failed to elicit a sufficient number of *BCs,* with participants instead using (S) VO or serial verb structures to answer these questions. The second preliminary study[27] (n = 20) found that the reformulated prompt questions – '*Ni zai [pingmu shang] kan dao le shenme dongxi*? / What have you seen on the screen?' and '*Na ge ren dui na ge dongxi zuo le shenme*? / What did the actress do to the stuff?' – more effectively elicited the target *BC* in response to 16 video clips. Therefore, the 16 video clips were selected to elicit *BC* in the OPTV in this pilot study with native speakers.

The video clips were designed to elicit two types of *BC*, *BC1* – whose complement contains a NP, and *BC2* whose complement consists of a directional verb. The video clips showed various actions by a woman as she moved things to new places, such as 'placing a book on the desk', 'pouring water from one cup to the other and pour it back', 'posting a picture on the wall and taking it away', 'taking out a mobile phone from her bag, having a look, and putting it back' and 'moving a chair into a room and moving it out', 'picking up a newspaper, having a look, and throwing it into the rubbish bin' (see Appendix 10 for a detailed list of the actions). The actions that involved a destination for the movement of the objects and were designed to elicit *BC1,* while the actions that involved a direction were designed to elicit *BC2*. Each video clip was designed to elicit one to two *BCs*. The 16 video clips were intended to elicit 8 *BC1* and *BC2*, respectively.

In the OPTV, the participants were invited to watch the video clips shown on the screen of either a digital camera or a computer one by one. The researcher sat in a position where she could not see the screen of a digital camera, laptop, or a computer, and was busy making notes. The participants were asked to start the video clips themselves. After watching each video clip, they were asked to tell the researcher what object they saw on the screen and what

---

[27] The second preliminary study with native speakers intended to elicit five types of *BC* using three tasks: an oral narrative task prompted by interview questions, an OIT displayed in discourse, and an untimed grammaticality judgment test. The native speakers who had participated in preliminary studies were not invited to this pilot study with native speakers.

the actress did to/with it. The participant was allowed to watch the video clips a second time if they could not see clearly. The participants' oral production data was recorded on a digital recorder and transcribed by the researcher afterwards.

5.4.3 Oral imitation task (OIT)

The oral imitation task (OIT) was conducted to elicit the oral production of *BC* under the conditions requiring attention to meaning. While processing the auditory stimuli, participants' primary focus was intended to be on meaning. If there is any chance to focus on form, it was secondary.

The OIT contained 20 items as well as 5 training examples (Appendix 11). The training examples were comprised of SVO structures which are not related to *BC*. Each of the 20 items consisted of a compound sentence containing two clauses. The first clause consisted of 3 to 8 Chinese characters (i.e., syllables), providing the background or context of use of *BC*. The second clause, related to the target *BC*, consisted of 6 to 11 Chinese characters (i.e., syllables). The participants were required to imitate the second clause following the first clause. There were two advantages of this design. The first clause could provide background or context for using the target *BC*, and the repetition of the first clause could prevent participants from rote memorising the targeted second clause. To ensure the items were sufficiently easy to understand, the sentences were constructed with the use of relatively high frequency words.

The 20 stimuli contained 7 *BC*s and 13 (S) VO structures (see Table 16). The *BC*s included 1 grammatical, 1 mis-formed and 1 overused *BC1*, 2 grammatical, 1 mis-formed and 1 overused *BC2*s. The 9 (S) VO structures contained 5 underused *BC1s* (i.e., ungrammatical (S) VO which must be corrected into *BC1s*) and 4 underused *BC2s* (i.e., ungrammatical (S) VO which must be corrected into *BC2s*), and 4 distractors. In short, all 20 stimuli consisted of 7 grammatical and 13 ungrammatical items, while the expected responses were 7 *BC1* and 7 *BC2* and 6 (S) VO structures.

Table 16: The Stimuli and Expected Responses in the OIT*

	Stimuli			Expected responses
7	**13**			**20**
Grammatical items	**Ungrammatical items**			**Grammatical items**
	Mis-formed BC	**Overused BC**	**SVO (Underused BC)**	
1 BC1	1 BC1	1 BC1	5 SVO	7 BC1
2 BC2	1 BC2	1 BC2	4 SVO	7 BC2
4 SVO				6 SVO

Note. Well-formed *BC* = grammatical *BC*, Mis-formed *BC* =ungrammatical *BC*, Underused *BC* indicates the case in which *BC* was not supplied under an obligatory context; Overused *BC* indicates the case in which *BC* was used in the context where *BC* should not be used. *BC2 contains one more grammatical item than *BC1*. This is because one of the SVO structures corresponding to *BC2* was grammatical but that corresponding to *BC1* was ungrammatical.

For uniformity and consistency, all items and training examples were pre-recorded on a cassette tape and a digital recorder using a headset microphone. The audio recordings were made by the researcher at a moderate speed. Care was taken to make the texts sound natural and to ensure that the target structure was not unnecessarily emphasised in any way. The items were presented in a fixed order – SVO and *BC* were evenly distributed – to avoid a cluster of *BCs* that might cause the participants to focus attention on *BC*. Each item was followed by a 10-second pause. The OIT took 8 minutes to complete.

The participants were first required to listen to training examples. When they were completely clear how to do the task, the participants were asked to listen to each item once. They were asked to restate the second clause in good Chinese following the repeated first clause and a beep.

5.4.4 Untimed grammaticality judgment task (UGJT)

The untimed grammaticality judgment task (UGJT) was designed to assess participants' metalinguistic competence to use the target type of *BC*. In this task, participants were asked to judge whether or not the sentences presented in written form were correct. The participants, therefore, were more likely to use their metalinguistic knowledge. Because in this task participants focus on linguistic form and no time limit was imposed for this task, their

judgments and corrections were considered to reflect a 'careful style' (cf., Tarone, 1983, 1985).

As for the OIT, the stimuli in the UGJT were constructed as compound sentences consisting of two clauses. The first clause was grammatical and provided the background or context for using *BC* in the second clause. The second clause was underlined and related to the target feature in question. Unlike in the OIT, the items in the UGJT were presented in written form (i.e., Chinese characters and their Romanised orthography- *Pinyin*).

The distribution of the stimuli in the UGJT is shown in Table 17. The UGJT consisted of 30 items and 4 training examples which were not related to the target feature, *BC* (see Appendix 13). The 30 items consisted of 10 distractors consisting of (S) VO structures and 10 stimuli related to each type of *BC*. The 10 items for each type of *BC* included 2 well-formed, 5 mis-formed, 1 overused *BC*, and 2 SVO structures which needed to be corrected into *BC*.

Table 17: The Stimuli and Expected Responses in the UGJT

10 Grammatical items	Stimuli			Expected responses
	20 Ungrammatical items			30 Grammatical items
Well-formed *BC*	Mis-formed *BC*	Overused *BC*	SVO (Underused *BC*)	
2 *BC1*	5 *BC1*	1 *BC1*	2 SVO	9 *BC1*
2 *BC2*	5 *BC2*	1 *BC2*	2 SVO	9 *BC2*
6 SVO			4 SVO	12 SVO

Note. Well-formed *BC* = grammatical *BC*, Mis-formed *BC* = ungrammatical *BC*, Underused *BC* = other structures supplied in an obligatory occasion for using *BC*; Overused *BC* = *BC* used in an occasion where *BC* should not be used.

The mis-formed *BCs* were composed according to the error types found in previous L2 acquisition studies, that is, incorrect word order, incorrect semantic meaning of the ba-NP (O), missing or incorrect main verb (V), incorrect verb residual (R), and the inappropriate semantic relationship between ba-NP, main verb and verb residual (R) (see, Du, 2004; Xiaoping Gao, 1999; H. G. Jin, 1993; D. Li, 1996; S. Zhang, 2002). Each mis-formed *BC* contained one error. In short, 10 grammatical and 20 ungrammatical stimuli were presented.

The expected responses consisted of 9 grammatical *BCs* for each type of *BC* and 12 (S) VO structures.

There were two versions of the items. The sentences that appeared at the end of version A were presented at the beginning for version B. The order of the items was reversed to reduce the possibility that ordering would affect the results.

When completing the UGJT, the participants were provided with a written form of the questions with the 4 training examples shown at the beginning (see Appendix 12). The researcher showed participants how to do the training examples. After the participants confirmed that they were fully clear about what to do in the task, they were asked to begin the UGJT following the procedures:

- read each sentence aloud;
- orally judge whether or not the underlined parts of the items were grammatical
- express how certain they were about their judgment;
- orally correct the underlined part they judged to be ungrammatical (even if the sentence was in fact grammatical).

The participants were asked to read each sentence aloud before making their judgments. This requirement was based on my observations in my preliminary studies. In my preliminary studies with native speakers, some native speakers had inadvertently misread ungrammatical sentences as grammatical ones. In one of my studies with learners (Xiaoping Gao, 2008), one learner told other students that she made judgments by guessing even without reading the sentences. Therefore, in this pilot study I reminded the participants to re-read the sentence carefully when they misread the stimuli.

Both judgements and correction of this task were performed orally for two reasons. First, writing Chinese characters is time consuming and written responses may be influenced by learners' skills in writing Chinese characters rather than their linguistic knowledge. Second, it is easier to monitor answers in writing than in speaking.

5.4.5 Interviews

A follow-up interview was conducted to gather information on the extent to which the participants were aware of the target features in the various tasks, and their perceptions about task difficulty. The interviews were semi-structured (see Appendix 14 for more information). The questions included what target features the participants thought the tasks had focused upon, whether the items were difficult, which task was the most difficult and easiest and why, and what task procedures influenced their performance. The participants' responses were recorded on a digital recorder.

5.5 Analysis and measures

The native speakers' orally produced data in all the tasks was transcribed into Chinese characters. The native speakers' use of *BC1* and *BC2* was coded without consideration of false starts. For the purpose of the pilot study, the number and percentage of the native speakers who had provided expected responses were calculated for each stimulus in the three tasks.

It should be noted that the participants' performance in the above three tasks (i.e., the OPTV, the OIT, the UGJT) yielded four measures: elicited oral production scores in the OPTV, elicited oral imitation scores in the OIT, and untimed grammaticality judgment scores and correction scores in the UGJT.

In the transcription of the OPTV, the elicited oral production was scored based upon the use of *BC1* and *BC2* in the 16 video clips. Each correct suppliance was scored one point. Ungrammatical or overused *BC* were scored zero. For each video clip, the number of the native speakers who provided the target *BC* was counted, and the percentage was also calculated.

In the transcription of the OIT, the native speakers' imitation was scored based on 16 items related to *BC*. Each expected response was scored one point, and each unexpected response was scored zero.

In the transcription of the UGJT, the participants' judgments were scored based on the 20 items related to *BC*. Each correct judgement was awarded one point, and the incorrect

judgement and 'not sure' responses were scored zero. A total score and separate scores for both grammatical and ungrammatical items were calculated

The participants' correction scores were calculated for the 16 ungrammatical (including 2 overused) stimuli in the UGJT. Complete correction was scored one point, and no correction and partial correction was scored zero.

5.6 Results

5.6.1 The OPTV

All the NSs' production in the OPTV was grammatical. When describing some video clips, (i.e., video 3, 4, 5, 6, 9, 13, 14, 15), some NSs produced other types of the *BC* (OTBC). When describing some video clips, e.g. video 4, 7, 13, 14, some NSs used SVO structures instead of the expected *BC*. Table 18 summarises the number (and percentage) of the NSs who produced *BC* and relevant SVO structures in the OPTV.

Table 18: The Number (%) of NSs Who Produced Four Types of Structures in the OPTV

Video clips	BC1	BC2	Other BC	SVO
1	22 (100%)	3 (14%)	-	-
2	22 (100%)	3 (14%)	-	-
3	1 (5%)	20 (91%)	1 (5%)	-
4		15 (68%)	7 (33%)	1 (5%)
5	4 (19%)	18 (82%)	5 (24%)	-
6	18 (82%)	4 (19%)	2 (10%)	-
7	8 (38%)	17 (77%)	-	2 (10%)
8	22 (100)	-	-	-
9	9 (43%)	11 (48%)	7 (33%)	-
10	21 (95 %)	21 (95%)	-	-
11	21 (95%)	1 (5%)	-	-
12	19 (86%)	5 (24%)	-	-
13	18 (90%)	6 (29%)	2 (10%)	1 (5%)
14	13 (59%)	9 (55%)	1 (5%)	1 (5%)
15	14 (64%)	15 (68%)	1 (5%)	-
16	14 (64%)	21 (95%)	-	-

Table 18 shows that the most successful video clips for eliciting *BC1* were videos 1, 2, and 8. By comparing the number of native speakers who used *BC1* and *BC2*, the video clips where at least 75% of NSs used the target BC was chosen as the threshold for selecting the video clips to be used in the final study. On this basis, video clips 1, 2, 6, 8, 10, 11, 12, 13 were selected to prompt *BC1*, video clips 3, 5, 7, 10, 16 were selected to elicit *BC2* , and video clips 4, 9, 14, and 15 were used as distractors.

5.6.2 The OIT

The native speakers' performance was the most consistent in the OIT. They corrected all the ungrammatical items as expected: 7 *BC1*, 7 *BC2* and 6 SVO structures, and did not make any mistakes when restating the second clauses of the items. Table 19 displays the number (and percentage) of NSs who produced the expected two types of *BC* and SVO and unexpected OTBC and SVO.

Table 19: The Number (%) of NSs Who Produced Expected and Unexpected Responses in the OIT

Item	Stimuli	Expected responses			Unexpected responses
		BC1	BC2	(S)VO	Other BC
1	UG(S)VO		21 (100%)	1 (5%)	
2	GBC1	22 (100%)			
3	OUBC1			20(91 %)	2 (9%)
4	G(S)VO			22 (100%)	
5	UG(S)VO	21(95%)			1 (5%)
6	GBC2		20(91%)	2 (9%)	
7	UG(S)VO	21 (100%)			
8	GBC2		22 (100%)		
9	UG(S)VO	21 (100%)			
10	UG(S)VO		22 (100%)		
11	UGBC1	22 (100%)			
12	UG(S)VO		22 (100%)		
13	UG(S)VO	22 (100%)			
14	UG(S)VO		22(100%)		
15	G(S) VO			22 (100%)	
16	UG(S)VO		22 (100%)		
17	UG(S)VO	22 (100%)			
18	OUBC2		2 (9%)	20 (91%)	
19	G(S)VO			22 (100%)	
20	G(S)VO			22 (100%)	

Note. G= grammatical items, UG= ungrammatical items, OU= overused *BC*.

From Table 19, it can be seen that at least 19 (91%) of NSs produced *BC1*, *BC2* and SVO structures as expected responses. Few NSs supplied unexpected resources. In a similar manner to the OPTV, a threshold of 75% of NSs producing *BC* was used to select the video clips for the main study. On this basis, all the OIT items could be used.

5.6.3 The UGJT

In the UGJT, the NSs quickly judged the items once they finished reading them aloud. The NSs' judgments on 30 items were almost 100 percent correct (Just one NS said ungrammatical item 12 was acceptable). The NSs' corrections showed slight differences on 5 items: items 2, 4, 6, 12, and 27. Table 20 summarises the number and percentage of NSs who made accurate judgments and corrections.

Table 20: The Number (%) of NSs Who Made Expected Judgments and Corrections

Stimuli	Number (%) of NSs who made accurate judgments	Number (%) of NSs who made corrections	
		using expected structures	using other structures
1. G(S)VO	22 (100%)		
2. GBC1	22 (100%)		
3. UGBC2	19 (86%)	19 (86%)	3 (14%)
4. UG(S)VO	22 (100%)	22 (100%)	
5. UGBC1	22 (100%)	22 (100%)	
6. GBC2	22 (100%)		
7. UG(S)VO	22 (100%)	22 (100%)	
8. UG(S)VO	22 (100%)	22(100%)	
9. UGBC2	21 (95%)	18 (82%)	4 (18%))
10. G(S)VO	22 (100%)		
11. OUBC1	22 (100%)	22 (100%)	
12. GBC2	22 (100%)		
13. UG(S)VO	22 (100%)	22 (100%)	
14. UGBC1	22 (100%)	22 (100%)	
15. UGBC2	22 (100%)	20 (91%)	2 (9%)
16. G(S)VO	22 (100%)		
17. UG(S)VO	22 (100%)	22 (100%)	
18. UGSVO	22 (100%)	20 (91%)	2 (10%)
19. UG(S)VO	22 (100%)	21 (100%)	
20. UGBC1	22 (100%)	22 (100%)	
21. OUBC2	22 (100%)	22 (100%)	
22. G(S)VO	21 (100%)		
23. UGBC1	22 (100%)	22 (100%)	
24. UGSVO	22 (100%)	22 (100%)	
25. G(S)VO	22 (100%)		
26. GBC1	22 (100%)		
27. UGBC2	22 (100%)	21 (95%)	1 (5%)
28. G(S)VO	22 (100%)		
29. UGBC1	22 (100%)	22 (100%)	
30. UGBC2	22 (100%)	22(100%)	

Note. G= grammatical items, UG= ungrammatical items, OU= overused *BC*.

From Table 20, it can be concluded that the objectives of the UGJT were achieved because nearly all of the target *BC*s were successfully supplied by the native speakers.

5.7 Reliability and validity

Before conducting the study, the wording of the items was checked by two native speakers of Chinese, one working as a research assistant at the Department of Applied Language Studies

and Linguistics and the other working as a tutor in Chinese at the School of Asian Studies at The University of Auckland.

To ensure interrater reliability, two interraters were trained to check the coding and scoring of the data. Two Chinese PhD students in Applied Linguistics were trained as raters to double code 15% of the data randomly selected. Interrater reliability was calculated using Pearson Product Moment Correlation. The correlation between the scoring of the researcher and that of the raters (Table 21) in the three tasks achieved a high level of agreement after negotiation. Therefore, interrater reliability was established.

Table 21: Pearson Product Moment Correlations between the Scoring of the Researcher and That of the Raters for the Three Tasks

Task	Raters' scoring	The researcher's scoring	
		Pearson correlation	Sig.(2-tailed)
OPTV	R1	.81	$p<.01$
	R2	.78	$p<.05$
OIT	R1	.79	$p<.01$
	R2	.85	$p<.01$
UGJT	R1	.93	$p<.001$
	R2	.89	$p<.001$

Note. R1 = the first rater's scoring, R2 = the second rater's scoring

Validity deals with whether the tasks measured what they were designed to measure. Both content validity and construct validity of the instruments were considered based on the findings of the pilot study.

Content validity refers to 'the representativeness of our measurement regarding the phenomenon about which we want information' (Mackey and Gass 2005, p.107). The study found at least 75% native speakers used the target *BC1* and *BC2* in the OPTV, and at least 90% of native speakers provided expected responses in the OIT and expected judgments and corrections in the UGJT. This suggests that the items in these tasks were effective in eliciting the target *BC1* and *BC2*. The stimuli in the OIT and UGJT include all possible uses of *BC* such as grammatical, mis-formed, overused, and underused items for each type of *BC*. The participants' performance in these tasks could serve as effective measures of their ability in using *BC1* and *BC2*. Therefore, the content validity was established.

Construct validity of the pilot study was established by analysing the data in the follow-up interview. The OPTV and OIT were designed to manipulate the contexts where learners' primary focus was on meaning, while the UGJT was designed to manipulate a context where learners' primary focus was on linguistic form. In the follow-up interview, all the native speakers claimed that they paid attention to meaning while doing the two oral tasks. None of them were aware of what linguistic structure was being targeted. Although these native speakers were aware that the UGJT tested some Chinese grammar features after judging and correcting sentences, few of them pointed out that the target feature was *BC*. Since the participants paid less attention to form in the OPTV and OIT than in the UGJT, their production in the two oral tasks could be considered closer to the 'vernacular style', while their judgments and corrections in the untimed UGJT could be considered as the reflection of their 'careful style' (cf. Tarone 1983). In addition, construct validity can be enhanced when multiple estimates of a construct are used (Mackey and Gass, 2005). This study employed two elicited oral scores in the OPTV and the OIT to manipulate the contexts where participants' primary focus was meaning, and two measures of untimed grammaticality judgment and correction in the UGJT to manipulate the contexts where participants' primary focus was form. To this end, it can be concluded that the construct validity of the study was also established.

5.8 Problems and changes

Generally speaking, the above results demonstrate that the instructions, items, and the procedures of the tasks could successfully elicit the target types of *BC* and provide effective measures of participants' oral and metalinguistic abilities to use *BC*. The only concern was the length of time it took participants to complete all the tasks, and the number of items. Some NSs indicated that there were too many items in the OPTV and UGJT. This caused some participants fatigue, which might have influenced the results. L2 learners, especially, may need even longer to complete the tasks. Some items were therefore removed from the OPTV, and some of the items in the UGJT were simplified. The following strategies were adopted for amending the items.

1. In the OPTV, 12 video clips which successfully elicited the target *BC* were selected for the main study. Seven of them were expected to elicit *BC1* and 5 to elicit *BC2*. The remaining 4 four video clips served as distractors.

2. The procedures of the OIT were modified. Some native speakers asked to listen to a few items again because they did not pay attention initially. The stimuli were therefore recorded at a slower speed with a longer interval (20 seconds) between items. Participants in the main study were permitted to re-listen to the stimuli once again after completing all 20 items if they needed, but only the last attempt was scored.

3. In the UGJT, two of the ungrammatical distractors were changed into grammatical ones. This was expected to reduce the time needed for judging and correcting items in the UGJT.

CHAPTER 6: METHODOLOGY

6.1 Introduction

The previous chapter reported on the pilot study with native speakers. This chapter describes the methodology of the main study. It begins by outlining the research questions that the study addresses, followed by a detailed discussion of its research design, participants, instruments, procedures, task reliability and validity, and methods of analysis.

6.2 Research questions

The general purpose of the present study was to investigate the variation in L2 learners' use of *BC*. The following questions were formulated:

1: Is there intra-learner variation in the learners' use of *BC*?

    1a. Is there intra-learner variation in the learners' use of *BC* in terms of *BC* type?

    1b. Is there intra-learner variation in the learners' use of *BC* in terms of task type?

2: Is there inter-learner variation in the learners' use of *BC*? If so, what are the sources of the inter-learner variation?

6.3 Design

A cross-sectional quantitative study was designed to answer the above research questions. This design involved five instruments: a background questionnaire, three tasks (i.e., OPTV, OIT, and UGJT), and a follow-up interview. The accuracy scores for *BC* in the four measures in the three tasks served as dependent variables. The accuracy scores in the OPTV and in the OIT were intended to measure learners' oral competence in using two types of *BC*, while untimed grammaticality judgment and correction scores in the UGJT were intended to tap learners' metalinguistic knowledge of *BC*. When addressing research question one, *BC* type and the method of measurement were considered as independent variables. It was expected that participants' primary focus was on meaning in the two oral tasks but on form in the two metalinguistic measures. When examining research question two, the background questionnaire was used to gather information about starting age, gender, first language, number of years of study, setting, and the self-rated proficiency. The six potential sources of inter-learner variation served as independent variables. A follow-up interview was

administered to gather qualitative data on learners' perceptions about task difficulty. The qualitative data was used to help interpret the findings of the quantitative analysis.

The study was implemented in two steps. First, a pilot study was undertaken with 22 native speakers of Chinese during March-July 2008 in Auckland, New Zealand (Chapter 5). The purpose of this pilot study was to check whether the instructions and the items of the instruments were clear and sufficient to elicit the target *BC*s, and to obtain baseline NS data. In the second step, the main study was conducted with 20 learners in New Zealand – a foreign language context – and with 90 learners in China – a second language context. The following describes the methodology of the main study in detail.

6.4 Participants

A total of 110 adult learners of L2 Chinese participated in the main study, 20 in New Zealand, and 90 in China. These learners had a mean age of 24 (SD = 5.3), ranging from 18 to 47. The learners at age 19 (8.18%), 20 (11.82%), 21 (17.27%), 22 (14.55 %), 23 (11.82%), and 24 (9.09%) collectively accounted for 72.73% of the sample, with the modal age being 21. The background information of the two groups of learners is described in detail below.

6.4.1 The L2 learners in New Zealand

L2 learners in New Zealand were recruited according to two criteria: 1) having taken Chinese lessons for at least one year, and 2) home language is not Chinese. These criteria were established for these reasons. First, *BC* was taught at the late elementary stage. Learners would not have a chance to learn *BC* until they had studied Chinese for at least a year under formal instruction. Second, two learners were born in a Cantonese speaking family, but their daily language was English. These learners were accepted because *BC* is not part of Cantonese grammar.

New Zealand learners were recruited by advertising at two universities in Auckland after obtaining ethics approval. The participant information was passed to students by posting the Participant Information Sheet (see Appendix 7) on online study resources and through word-of-mouth from their Chinese instructors. The first 20 respondents were recruited and received a $10 gift voucher for their voluntary participation.

127

The 20 New Zealand learners were students enrolled in Chinese language courses at a university in Auckland. Ten were female and 10 male. They were between 17 and 33 years of age (M = 19). 11 were English and 9 Korean native speakers. The 9 Korean speakers had lived in New Zealand for 6 to 13 years (M= 7.8). These learners had mixed language background and proficiency in Chinese. Fourteen were studying at Stage Two Chinese course and 6 at Stage Three. These Chinese language courses for each academic year required an attendance of 120 hours of Chinese lessons (5 hrs/wk $\times$ 24 weeks). All Korean learners could speak English, and 4 of them could speak another foreign language (e.g., Japanese and Spanish). Three English native speakers understood a little bit Cantonese, and 6 had also studied other foreign languages (e.g., Japanese, Korean, and Spanish).

These learners came from diverse language backgrounds. All but five started learning Chinese after age 17 at university. Four (3 English and 1 Korean native speakers) started learning Chinese at ages of 11 to 14 at schools in New Zealand, where they received an average of 600 hours of Chinese lessons. One Korean native speaker had learned Chinese at an international school in China at age 11-12 and finished three months' Chinese language study at a Chinese university at age 20. One Korean and one English native speaker (at Stage Three) had studied Chinese language at a Chinese university for a year in Beijing. Two English speaking learners studied at schools in China for 6 weeks and 4 months, respectively. One English and one Korean speaker (at Stage Two) attended 3 weeks' study abroad program in China. Only three learners (1 English speaker and 2 Korean speakers) who had studied Chinese in China for over one year sat the Standard Chinese Proficiency Test, *Hanyu Shuiping Kaoshi* (HSK)[28]. The English learner achieved band 4 at an elementary level and the two Korean learners achieved band 6 and 7, respectively, at an intermediate level. Table 22 summarises the background information of the learners in New Zealand.

---

[28] The participants in this study attended the old version of the HSK test, which contains four sections (i.e., listening, grammar, reading, and comprehensive expressions) and discriminates 11 bands of proficiency. No oral element is involved until band 9, an advanced level.

Table 22: Background Information of the L2 Learners in New Zealand

Group	English-speaking learners (n = 11)	Korean learners (n = 9)	Total (n = 20)
Age	19-33 (M = 20.7)	19-24 (M=20.9)	19-33 (M= 20.8)
Gender Male	8	2	10
Female	3	7	10
Stage Two	9	5	14
Stage Three	2	4[a]	6
# of learners who experienced exposure to China	3	4	7
Years of residence in NZ	18.5	7.9	14.7
#Other languages speakers	6	4	10

With regard to the purposes for learning Chinese, all English speakers chose personal interest but only 67% of Korean learners did so. Some Korean learners said that they took Chinese language courses to please their parents. Nearly an equal proportion of English (73%) and Korean (78%) learners recognised that they studied Chinese to enhance their employment opportunities.

6.4.2 The L2 learners in China

The learners in China were recruited according to the same criteria except that they had started learning Chinese in China after age 13. This criterion was added after I interviewed a Korean participant who performed like native speakers of Chinese. He had lived in Beijing since 5 years old and actually was a bilingual learner. His data was finally eliminated to ensure that all participants of this study were adult learners.

The learners in China were recruited by advertising at 6 universities and 3 Chinese language schools in Beijing after obtaining permission, on two local English websites, and by word of mouth in Korean and English speakers' communities in Beijing and among Chinese teaching staff. It was stated in the advertisement that participants in this study could receive 40 yuan (equivalent to NZ $10 then) as compensation. Of 96 respondents, three learners (1 English and 2 Korean speakers) were eliminated due to they started learning Chinese in China at an age younger than age 13 (i.e., age 5, age 8 and age 12), and the other three (2 English and 1

Korean speaker) were eliminated because they reported that they had never learned *BC* in the follow-up interviews.

The China group contained 45 English and 45 Korean native speakers. All but one were either working or studying at language schools or universities in Beijing[29]. Forty six were female and 44 male. The English native speakers came from 5 English speaking countries: U.S. (n = 25), New Zealand (n = 7), Canada (n = 6), UK (n = 6) and Australia (n = 1). The backgrounds of these learners were mixed. Sixteen learners (6 Korean and 10 English speakers) started learning Chinese in China, whereas others (n = 74) had prior exposure to Chinese in foreign language contexts. At the time of testing, 37 learners (22 Korean and 15 English speakers) were studying Chinese courses at universities in Beijing as part of a study abroad program.

Overall, the learners' Chinese proficiency was mixed, ranging from elementary to advanced. Some (n = 21) were enrolled in either intermediate or advanced Chinese Language courses in the universities, whereas others were studying towards an undergraduate degree program or a higher degree in various disciplines, such as Economics and Politics. Thirty one of 45 Korean learners sat the standard Chinese proficiency test, *Hanyu Shuiping Kaoshi* (HSK), and 29 had received scores with an overall mean of band 6.73 out of a possible band 11. Only 8 of 9 English speakers had been tested through the HSK, receiving an average of band 5.5.

On average, the China group had been learning Chinese for 2.56 years and resided in China for 1.4 years on average. All 45 Korean learners had studied English as another foreign language, whereas 31 of 45 English speakers had experience in learning other foreign languages such as French and Spanish. The background information of the learners in China is summarised in Table 23.

---

[29] One student was studying an intermediate course at Shanxi University.

Table 23: Background Information of the L2 Learners in China

		CSL learners in China (n = 90)		
		L1-English speakers (n = 45)	L1-Korean speakers (n = 45)	Total (n = 90)
Age	Mean	26.11	23.22	24.67
	Range	20 - 40	18 – 47	18 – 47
	SD	5.81	4.88	5.53
Gender	Female	18	28	46
	Male	27	17	44
Starting age	M	20.78	19.73	20.26
	Range	5 - 32	14 – 41	5 - 41
	SD	5.99	4.41	5.26
	M	2.68	2.53	2.56
	Range	0.5 -10	0.6-10	0.5 -10
Years of study	SD	1.82	1.69	1.75
	M	1.48	1.34	1.41
Years of residence in China	Range	0.17-12	0.17-10	0.17-12
	SD	1.80	1.76	1.77

The purposes for learning Chinese of both the Korean and English learners followed the same pattern: Interest > Employment > Degree > Visit, although more Korean learners selected multiple reasons. The percentages of learners who selected the different purposes are shown in Table 24.

Table 24: The Learners' Purposes for Learning Chinese

L1 group	Purposes			
	Degree	Employment	Interest	Visit
English	11%	40%	49%	7%
Korean	38%	53%	76%	0%

6.5 Instruments

The main study used the five instruments that had been trialled in the pilot study with native speakers. The instruments were administered in one-on-one interviews between the researcher and the learners. After reading the participant information sheet and signing the consent form, the participants were informed that the interview would be recorded and that their responses would be kept anonymous. The instruments were administered in the same order as in the pilot study: a background questionnaire > the OPTV > the OIT > the UGJT > a

follow-up interview. The order of the three elicitation tasks was determined to minimize the possibility that learners could pay conscious attention to *ba* in *BC*. Finally, the follow-up interview was conducted. All oral data was recorded on a digital recorder and transcribed by the researcher later.

The following sections describe the rationale for using each instrument, content of the instrument, and procedures for administering it.

6.5.1 Background questionnaire

A background questionnaire was used to obtain information about learners' personal background. The questionnaire included both closed and open ended questions regarding the learners' age, gender, reasons for learning Chinese, starting age and experience of learning Chinese, self-rated proficiency in four skills (reading, writing, listening and speaking), their scores in the Chinese proficiency test if available, etc. (see Appendix 15 for the English and Chinese versions of the background questionnaire).

There are a few advantages in using questionnaires (Mackey & Gass, 2005). Questionnaires can elicit comparable information from a number of respondents, and gather longitudinal information and personal opinions from learners in a short period of time. In addition, they can be administered in many forms.

In this study, the questionnaire was administered orally by the researcher in one-one-one meetings with the learners. At the beginning of the meetings, the researcher asked the learners questions following the semi-structured background questionnaire in either Chinese or in English. The learners could answer the questions in either of the languages. The researcher made notes on the written questionnaire. Learners were given ample time to amplify their answers. All oral data was recorded on a digital recorder.

Oral answers were used to mitigate the learners' difficulty in writing, and to gather rich qualitative data in a relatively short period of time. English was used whenever the learners had difficulty in understanding and expressing themselves in Chinese. This was possible because English was the other common language of the English and Korean learners.

6.5.2 Oral Production Task prompted by Video clips (OPTV)

The OPTV was designed to elicit learners' oral production by using video clips and question prompts. Silent films or video clips have been used to investigate cross-linguistic features in SLA (e.g., Bardovi-Harlig, 2000; Chafe, 1980; Erbaugh, 1976). For example, Tarone's seminal studies (Tarone, 1985; Tarone & Parrish, 1988) examined variability by using silent films to elicit oral narration. Tarone argued that oral narrative data represents a vernacular style of speech because less attention to form was possible. H. G. Jin (1989) adopted silent films to elicit five types of *BC* which contain a nominal complement to investigate acquisition order. Du (2004) used self-recorded video scenes to elicit the *BC* containing the perfective marker LE and the resultative verbal complement (RVC) (See Chapter 4 for definition of RVC).

There are a number of advantages in using the OPTV. First, *BC* only occurs in a few specific contexts, so it is hard to collect naturalistic data by observation. Second, the same prompts in the OPTV could predict the maximum potential occasions to use *BC*, while free oral production elicited by interview topics usually contains different numbers of obligatory occasions for using *BC*. Third, in contrast to static pictures (see, P. Li, 1990), dynamic video clips matches the essential function of *BC* – describing a complete dynamic process of movement or change. In addition, when producing speech data in the OPTV, participants are more likely to focus on meaning than on linguistic form. So learners' production was considered closer to the 'vernacular style'.

However, no existing film or video clips were available for eliciting the target BC types in this study. Although Du (2006) used self-recorded video clips to prompt learners' oral production of *BC*, the 'pairs of video clips' were not very successful. In Du's design, the first video scenes were used to elicit SVO structures, while the second ones were used to elicit *BCs.* The assumption was that SVO structures would be used when the entities denoted by the definite be-NPs appeared for the first time, and *BCs* would be used when the entities appeared again. As Du recognised, some native speakers used *BCs* to describe the first video clips, while L2 learners generally used SVO structures to describe both video clips. Indeed, using SVO structures to describe both video clips is grammatical and acceptable if the prompt question is 'Could you please describe the video scenes?'. Inspired by Du's (2006) study, I have realised that to elicit *BC* it is essential to make the speaker realise that the entity denoted

by *ba*-NP is known to the hearer. To do so, two prompt questions were employed for each video prompt (see p.112 for detailed information about the choice of questions).

In this study, these video clips were produced to involve reference to two or more people/objects and thus created contexts for the use of the target *BC*. Half the participants received the items in one order and the other in the reversed order in order to control for potential ordering effects of the items.

To guard against the possibility that the video clips posed differing levels of difficulty for the learners the video clips were counterbalanced in each group at each testing time.

The 16 video clips which had been trialled on native speakers in the pilot study were used in the main study. Twelve of them which successfully elicited the target types of *BC* from native speakers were analysed. The other 4 video clips were treated as distractors. The maximum number of obligatory occasions was 7 for using *BC1* and 5 for using *BC2*.

The OPTV was administered to the participants individually. Participants were asked to answer questions – 'what things did you see on the screen?' and 'what did the actress do to/with them' – in complete Chinese sentences after watching each video clip. They were allowed to watch the video clips more than once if they requested. Chinese nouns were even provided to learners when they asked for help.

6.5.3 Oral Imitation Task (OIT)

The OIT that had been tested by the pilot study with native speakers was also used to collect L2 learners' oral data in the main study. In this task, participants' primary focus is on meaning.

The OIT is often used as a means to determine the nature of learners' grammatical systems. As Mackey and Gass (2005) noted, the basic assumption underlying OIT is that a sentence will be relatively easy to repeat if it is part of a person's grammar otherwise it will be difficult. It is assumed that sentences are 'filtered' through the learners' grammatical systems. Although the process of OIT is not well understood (Bley-Vroman & Chaudron, 1994), it is considered that the OIT is reconstructive in nature and goes beyond rote memory and repetition. The OIT has been identified as an effective way of measuring learners' implicit

linguistic knowledge (R. Ellis, 2005a; Erlam, 2006). The OIT was considered to elicit a speech style falling between a 'vernacular style' and a 'careful style' on the continuum of speech styles (Tarone, 1985; Tarone & Parrish, 1988).

In the typical OIT, the sentences are designed to incorporate specific grammatical structures. Munnich, Flynn and Martohardjono (1994)pointed out that the OIT can incorporate both grammatical and ungrammatical sentences. L2 learners' ability to repeat them accurately is viewed as a reflection of their internal grammatical systems.

This study increased the difficulty of the OIT by composing most of items as compound ungrammatical sentences. There were two reasons for this. First, most Chinese words are single-syllable. A sentence which contains 7 words or syllables is relatively easy to rote memorize. To avoid rote memorization and attract participants' attention to meaning, ungrammatical sentences were devised to make it impossible for the learners to imitate correctly without understanding the meaning of the sentences. Successful imitation of sentences in the OIT requires the ability to understand the meanings that the phonological forms convey, identifying the occasions to use *BC*, and producing correct sentences based on the meaning of the original sentences.

The OIT that had been tested in the pilot study with native speakers was also used to collect L2 learners' oral data in the main study. The OIT consisted of 20 items. Seven of them involved *BC* structures and 13 (including 4 distractors) were presented in the form of (S)VO. The expected responses were 7 *BC1* and 7 *BC2* as well as 6 SVO structures (see Table 16 for more detail). There were 8 maximum obligatory occasions for using *BC1* and *BC2*, respectively.

Although the OIT successfully elicited the target *BC* from the native speakers, the pre-recorded stimuli were modified by slowing down the recording speed and extending intervals between stimuli from 10 seconds to 20 seconds. In addition, learners were allowed to listen to all stimuli a second time if they needed. Only their final responses were scored. This was done because a large number of the participants were unable to understand the sentences on the first hearing. Therefore, in order to elicit the data from learners it was necessary to allow them to listen to the stimuli a second time. The other reason for doing so is that all tasks in this study were untimed in order to reduce the influence of time pressure on learners.

At the beginning of the OIT, participants were asked to listen to the pre-recorded instructions and five training examples (i.e. three grammatical and two ungrammatical simple sentences). They were allowed to re-listen to the instructions and examples until they fully understood what was required. The pre-recorded sentences were played to participants one by one. Participants were instructed to restate the second clause in good Chinese when presented with the first clause following a beep as soon as they had heard the whole sentence. Their responses were recorded on a digital recorder and analysed later. The collected data were scored following the same scoring scheme as used in the pilot study.

6.5.4 Untimed Grammaticality Judgment Task (UGJT)

The untimed grammaticality judgment task (UGJT) was used to collect data reflecting the learners' metalinguistic knowledge of the *BC*.

The GJT has been widely used as a means of data collection in SLA research (Birdsong, 1989; Chaudron, 1983, 2003; R. Cowan & Hatasa, 1994; Davies & Kaplan, 1998; R. Ellis, 1991; Gass, 1994; Mandell, 1999; Sorace, 1996) . However, the validity and reliability of this task are uncertain. First, it is not always clear if participants make judgements on the basis of their intuitions (implicit knowledge), memory (L. R. Brooks & Vokey, 1991; Vokey & Brooks, 1992), context, certain parsing strategies, or their metalinguistic knowledge (Y. Han, 2000; Y. Han & Ellis, 1998; Sorace, 1996). R. Ellis (1991) showed that learners who were uncertain about their responses changed their responses in subsequent tests. R. Ellis (2005a) further demonstrated that time pressure in the GJT might determine what the GJT measures. That is, L2 learners were more likely to access their implicit knowledge in a timed GJT but explicit knowledge in an untimed GJT. The reliability of the GJT is further threatened by the instability of NS judgments. This increases the difficulty of scoring grammaticality judgements. On the other hand, some researchers (e.g., Gass, 1994; Mandell, 1999) have showed that the reliability of the GJTs was reasonably high and that GJT data are reliable and valid measures of linguistic knowledge.

Despite the uncertainty about GJTs, the advantages of using them to collect data are clear. The GJT can easily collect data from a large population and involve target features which are not frequently encountered in the natural setting (R. Ellis & Barkhuizen, 2005). This study employed the UGJT also because *BC* is not frequently encountered in natural settings.

In this study, the UGJT was devised to require untimed grammaticality corrections, which are also widely used in SLA research (R. Ellis & Barkhuizen, 2005; E. C. Klein, 1995; Salaberry, 1998; Schachter, 1989; Trahey & White, 1993). Untimed grammaticality corrections are often used to verify the basis of participants' judgments.

The UGJT included 30 items and 4 training examples (Appendix 13). All but two items that had been tested in the pilot study with native speakers were used in the main study (see Table 17 for the distribution of the items). Two distractors were reformed into grammatical sentences to save time and mitigate learners' fatigue. There were 9 obligatory occasions for each type of *BC*.

After being shown four training examples, participants were asked to read each sentence aloud and then to judge whether the underlined part was categorically grammatical or ungrammatical or they were not sure about the grammaticality. They also had to orally correct the parts which they considered ungrammatical. The entire process was conducted without any time pressure.

6.5.5 Interviews

A follow-up interview was conducted after the participants had completed all tasks. The purpose of the follow-up interview was to establish whether the learners had recognised that the focus of the tasks was on *BC*, and to obtain their personal perceptions about task difficulty and explanations of their performance in the tasks. The interview guideline can be seen in Appendix 14.

The follow-up interview was administered orally following the semi-structured guidelines. The interviews were conducted between the researcher and the learners using either Chinese or English whichever the learners preferred.

6.6 Analysis and measures

This section describes the criteria for coding the transcribed oral data from the OPTV and the OIT, lists the scoring schemes for the four measures of *BC* in the three tasks, and presents the method of analysis.

6.6.1 Coding criteria

Data collected from four tasks were transcribed into normal orthography (both Chinese characters and English) by the researcher and cross-checked by a second transcriber. The transcription of the oral data in the OPTV and the OIT was coded using the following criteria.

1. The target *BC* without a subject was considered grammatical. This is because that Chinese is a pro-drop language. If a sentence shares the same subject with the previous sentence in discourse, its subject is normally omitted.

2. The accurate use of *jiang* was considered to reflect the accurate use of *BC* because *jiang* is the formal form of *ba*.

3. Other types of *BC*, for example, *BC* containing a resultative complement (RC) was not considered.

4. The use of other functions and meanings of *ba* was not considered as the use of *BC*.

 For example,

  "她把了手机"

 ' *Tā bǎ le shǒujī,---*'

 Lit. She ba PFV marker mobile phone ---

In this sentence, *ba* is used to express the meaning of 'take or grasp'. However, the lexical meaning intended here is not related to the grammatical meaning of *ba*. Therefore, the use of *ba* in this context was not scored.

5. False starts were not considered

 For example

  "她把 (1.2), 她拿了一个杯子"

 '*Tā bǎ (1.2), tā ná le yi gè bēizi, ---*'

 * 'She *ba* (1.2), she took a cup---'

In this sentence, *ba* produced in this sentence was a false start, so was not scored.

6. In the case of self repetitions or corrections, the last version was scored.

For example, when coding the following sentence, the non-ba structure, '*xiang fang yi zhang kapian zai qiang*', was scored.

∗ "她把一张卡片(.6) 想放一张卡片在墙, 可是她又拿走了"

Tā bǎ yi zhāng kǎpiàn (.6) xiǎng fang yi zhāng kǎpiàn zài qiáng , kěshì tā yòu ná zǒu le.

Lit. She ba one CL card (.6), intend put one CL card at wall, but she again take away PFV.

∗ She *ba* a card, intended to stick a card on the wall, but she took it away.

7. When a few VPs follow the ba-NP, only the first one was counted as the VP of *BC*. The others were interpreted as examples of pro-drop (omission of a topical argument).

For example, when coding the following sentence, only the first clause following the ba-NP (i.e., '*Ta ba shouji fang dao shubao li*') was counted as a *BC*.

"她 把手机放到书包里，又拿出来，看了看"

'*Tā bǎ shǔjī fàng dào shūbāo lǐ, yòu ná chūlái, kàn le kàn.*'

Lit. 'She ba mobile phone put into (her) bag, again take out, have a look. '

She put her mobile phone into her bag, and then took it out again to have a look.

6.6.2 Measures and scoring schemes

The three tasks yielded four measures of the learners' use of *BC*: elicited oral production, oral imitation, untimed grammaticality judgments and error correction. The same scoring scheme as used in the pilot study was adopted (see Section 5.4.3.5). That is, each correct suppliance, imitation, judgement or correction related to the target *BC* was awarded one point. Each inaccurate suppliance, imitation, judgement and correction was scored zero, as was the 'not sure' response in the UGJT.

Examples that reflect the scoring of the learners' responses in the OPTV are shown below.

a) 她把那个包放在桌子上 = 1 point

 Tā bǎ nà ge bāo fàng zài zhuōzi shàng

Lit. She *ba* that bag put at table top.

 'She put the bag on the table.'

b) *她把那个包放在桌子 = 0 point

　　Tā bǎ bāo fàng zài zhuōzi

Lit. She ba bag put at table

c) *她放那个包在桌子(上) = 0 point

　　Tā fàng shūbāo zài zhuōzi shàng

Lit. She put bag at zhuozi top

d) *她把了书包= 0 point

　　Tā bǎ le shūbāo

Lit. She *ba* LE bag

Only sentence a) demonstrates the correct use of *BC1,* scoring one point. The other three examples are misused *BC* and were scored zero. Sentence b) is an ill-formed *BC1*, in which the necessary locative noun at the final position, *shang*, 'top', was omitted. Sentence c) is a case of under-used *BC1*, in which *BC1* was not supplied in an obligatory context. Sentence d) is an example of overused *BC1*, in which the preposition – *ba* – was misused as a main verb.

6.6.3 Methods of Analysis

For the purpose of this study, accuracy was used to examine variation in the learners' use of *BC1* and *BC2*. In the oral production tasks, accuracy refers to how well the target language is produced in relation to the rule system of the target language (Skehan, 1996). In the UGJT, accuracy was defined as percentage of accurate judgments and correction.

In the OPTV, an obligatory occasion analysis was carried out to establish the percentage of correct suppliance of each target *BC*. The accuracy of use was calculated using Pica's (1983) formula for target-like analysis (R. Ellis & Barkhuizen, 2005) to take into account overused *BC*. Accuracy was calculated by dividing the total number of accurate suppliance by the total number of obligatory occasions and the number of overused *BC* (see the following formula).

correct use in contexts　　　× 100 = per cent accuracy

obligatory contexts + # overuse

The obligatory contexts for using target type of *BC* were established based on the native speakers' baseline data. An occasion where 75% of the native speakers used the targeted type of *BC* was considered an obligatory context. The native speakers' baseline data showed 7 obligatory contexts for using *BC1* and 5 for using *BC2*. Total scores were calculated by averaging the percentage scores for each type of *BC*. The maximum possible accuracy score for each type of *BC* was therefore 1 (or 100%).

In the OIT, the participants' failure to imitate the target clause at all or to reproduce it in such a form that they did not create an obligatory context for the target *BC,* was coded as avoidance. Each imitated sentence was scored 1 when the target *BC* was correctly supplied or 0 when the target *BC* was either incorrectly supplied or avoided. In addition, occasion where a learner overused *BC* were included in the analysis to calculate target-like use. . Scores were expressed as percentage correct.

When scoring grammaticality judgements in the UGJT, each item was scored dichotomously as correct vs. incorrect, with 'not sure' responses scored as incorrect. A percentage accuracy score was calculated.

When scoring error correction in the UGJT, only the 8 ungrammatical stimuli were considered. Each accurate correction supplied scored 1 and each inaccurate or non-accurate correction scored 0. A percentage accuracy score was calculated.

 In consideration of learners' limited vocabulary, misuse of lexical nouns (e.g., NP2) was not considered an error when scoring, but misuse of the locative nouns (LN) in *BC1a* (S+*ba*+ NP1+ V+P+NP2+ LN) was. This is because the LN is one of the determinants of a correct *BC*.

6.7 Reliability and validity

To ensure the validity and reliability of the instruments, 15 percent of transcribed written data randomly selected from each task was double coded by two raters. Both of them were native speakers of Chinese and doing PhDs in Applied linguistics or in Chinese. They were

provided with the coding scheme (see Section 6.7.1) and the formulae for calculating target-like accuracy in terms of the four measures (i.e., oral production, oral imitation, untimed grammaticality judgement, and error correction). Inter-rater reliability for all measures was examined using the Pearson Product Moment Correlation Coefficients. The correlation coefficients between the researchers' and the raters' scoring were moderate. In particular, the correlation for the OPTV (r = 0.35 for the first rater, r =.42 for the second rater) were quite low. The discrepancy was due to the fact that the raters were not sufficiently clear about the classification of *BC1* and *BC2*. After being briefed on the classification of *BC1* and *BC2* and basic Chinese grammar, the raters re-coded the data and calculated accuracy scores. The recalculated Pearson Product Moment Correlation Coefficients between the scores of the researcher and those of the two raters are shown in Table 25. The correlation coefficients indicate that the researchers' and raters' scoring correlated strongly. It can, therefore, be claimed that interrater reliability was established.

Table 25: Pearson Product Moment Correlations between the Scoring of the Researcher and That of the Raters for the Four Measures

Task	Raters' scoring	The researcher's scoring	
		Pearson correlation	**Sig.(2-tailed)**
OPTV	R1	.81	*p*<.01
	R2	.78	*p*<.05
OIT	R1	.83	*p*<.01
	R2	.86	*p*<.01
UGJT-GJ	R1	.95	*p*<.001
	R2	.80	*p*<.001
UGJT-EC	R1	.79	*p*<.05
	R2	.82	*p*<.05

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT-GJ = grammaticality judgements in the UGJT, UGJT- EC= error correction in the UGJT; R1 = the first rater's scoring, R2 = the second rater's scoring

In addition to the internal validity (i.e., content validity and construct validity) established in the pilot study with native speakers, the main study also examined external validity. External validity is concerned with to what extent the research findings can be generalized in a large population beyond the research the confines of the research setting and participants (Mackey & Gass, 2005).

In this study, external validity was established through random sampling and achieving representativeness of the samples. First, convenience sampling was adopted to select early respondents from a large population through advertising in New Zealand and in China. The 110 learners were recruited from 6 universities and 2 language schools in Beijing, China, where there were well-established Chinese as a second language programs and two universities in Auckland, New Zealand that were the only universities in Auckland offering Chinese language courses. Second, the learner sample, particularly the 56 English speaking learners, is highly representative. This was because the overall population of English speaking learners of Chinese was limited and the English speaking learners came from five English speaking countries: the U.S., Canada, the U.K., Australia, and New Zealand. Therefore, this study can reasonably claim the external validity.

6.8 Statistical analyses

A number of statistical methods were employed to investigate the research questions. The first research question regarding intra-learner variation was examined using statistical methods to compare means. The effect of *BC* type was tested by means of either paired-samples *t*-tests or Wilcoxon Signed Rank Tests according to whether or not the assumption of normality was met. The effect of task type was examined using either a one - way analysis of variance (ANOVA) or the Friedman Test depending on the normality of the data.

The second research question regarding inter-learner variation was investigated using three types of multivariate statistical techniques. Oral and metalinguistic scores for *BC1* and *BC2* served as dependent variables, while six potential sources of inter-learner variation (i.e., starting age, gender, L1, setting, number of years of study, and self-rated proficiency) served as independent variables. First, an independent samples *t*-test and its non-parametric counterpart, the Mann-Whitney *U* test, were performed to compare group means of male/female, English/Korean speaking learners, and learners in New Zealand/ China in order to examine the effects of gender, L1, and setting. Second, Pearson Product-Moment Correlation Coefficients were used to probe the relationships between each of the factors (i.e., starting age, years of study, and self-rated proficiency) and the oral/metalinguistic scores for the two types of *BC* in the four measures. Finally, four factors, L1, setting, starting age, and years of study, which figured in statistically significant effects, were adopted as the

independent variables in a stepwise multiple regression analysis. The oral/metalinguistic scores for *BC1* and *BC2* were used as the dependent variables.   144

Significance levels (p<. 05, p<. 01) were calculated and reported for each statistical analysis. However, significance levels are sometimes influenced by sample size. The data with a sufficiently large sample size often yield significant results, while the data having a small sample size generally produce insignificant results. Accordingly, the statistical significance is not always meaningful (cf. J. Cohen, 1997a, 1997b; Norris & Ortega, 2000). To include some index of effect size or strength of relationship, effect sizes were calculated using the following methods corresponding to statistical tests used (Table 26). The data were interpreted according to corresponding Cohen's guidelines. Different methods were used to calculate effect size because using one method (e.g., Cohen's *d*) was not appropriate for both parametric and non-parametric tests.

Table 26: Statistical Methods Used for Calculating Effect Size

Statistical tests	Effect size	Formula	Interpreting guidelines
Paired samples *t*-test	Eta squared	$\eta^2 = \dfrac{t^2}{t^2 + N - 1}$	Cohen (1988, pp. 284-7): .01 = small effect, .06 = moderate effect, .14 = large effect (cited from Pallant, 2007, p.236)
Independent samples *t*-tests	Eta squared	$\eta^2 = \dfrac{t^2}{t^2 + (N1 + N2 - 2)}$	Cohen (Cohen, 1988, pp. 284-7): .01 = small effect, .06 = moderate effect, .14 = large effect (cited from Pallant, 2007, p.236)
One-way repeated measures ANOVA	Partial Eta Squared	$h_p{}^2 = SS_{between} / SS_{total} + SS_{error}$	Cohen (1988, pp. 284-7): .01 = small, .06 = moderate, .14 = large effect
The Mann-Whitney U tests	r	R = z / square root of N, where N = total number of cases.	Cohen's (1988) criteria .1 = small effect, .3 = medium effect, .5 = large effect.
Wilcoxon Signed Rank Test	r	r = z/squared root of N (where N = two times of cases),	Cohen's (1988) criteria: .1 = small effect, .3 = medium effect, .5 = large effect (cited from Pallant, 2007, p.225)

145

CHAPTER 7: INTRA-LEARNER VARIATION IN THE INTERLANGUAGE OF *BC*

7.1 Introduction

The previous chapter described the methods used in this study. This chapter will investigate the first research question:

1. Is there intra-learner variation in the learners' use of *BC*

    a. in terms of *BC* type?

    b. in terms of task type?

This question investigates to what extent there is intra-learner variation in the learners' use of *BC* in terms of *BC* type and task type. The two target types of *BC* were: *BC1* – a (locative) nominal *BC* whose complement contains a NP and *BC2* – a directional verbal *BC* whose complement contains a directional verb but no NP. The learners' use of two types of *BC* was elicited from three tasks: an Oral Production prompted by Video clips (OPTV), an Oral Imitation Task (OIT), and an Untimed Grammaticality Judgement Task (UGJT). The learners' performance in the UGJT was assessed by two measures: grammaticality judgments (GJ) and error correction (EC). Thus, the effect of task type was examined in terms of four measures: oral production in the OPTV, oral imitation in the OIT, and grammaticality judgments and error correction in the UGJT.

The learners' accuracy scores for using *BC* in the OPTV were calculated using obligatory occasion analysis (Brown, 1973; R. Ellis & Barkhuizen, 2005). In consideration of the overuse of *BC* (i.e., use in non-obligatory contexts), Pica's (1983) formula for target-like analysis was adopted (see Section 6.7.3). The accuracy scores in the OIT were also calculated using the formula for target-like analysis. For grammaticality judgements, each correct judgment scored 1, while each incorrect judgment or 'not sure' response scored 0. Each accurate correction was awarded 1 point, while each inaccurate or no correction scored 0. A percentage accuracy score for judgments and correction was calculated. The maximum possible score for each type of *BC* was 1 (or 100%) for each measure.

7.2 Results – Intra-learner variation

7.2.1 The effect of *BC* type

The first part of research question one investigates the effect of *BC* type on the learners' accuracy of use of *BC*. The hypothesis that there will be a difference between the 110 learners' accuracy of use of *BC1* and *BC2* was tested by means of either paired-samples *t*-tests or Wilcoxon Signed Rank Tests according to whether or not the assumption of normality was met.

7.2.1.1 Descriptive statistics

Table 27 shows descriptive statistics for the 110 learners' accuracy scores for *BC1* and *BC2* across the three tasks and four measures. Skewness and kurtosis are included to test whether the distribution of the data is normal.

Table 27: Descriptive Statistics – Accuracy for *BC1* and *BC2* across the Three Tasks and Four Measures (n = 110)

Type of BC	Task	M	SD	Range	Skewness	Kurtosis
BC1	OPTV	40.23	34.43	0-100	.22	-1.43
	OIT	27.82	23.44	0-90	.69	-.39
	UGJT- GJ	73.73	19.60	10-100	-1.01	1.07
	UGJT- EC	55.45	24.50	0-100	-.23	-.49
BC2	OPTV	16.73	28.09	0-100	1.81	2.40
	OIT	19.77	18.96	0-75	1.03	.62
	UGJT- GJ	62.95	21.60	0-100	-.48	-.18
	UGJT- EC	48.64	22.20	0-90	-.35	-.49

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT-GJ = grammaticality judgements in the UGJT, UGJT- EC= error correction in the UGJT; *BC1*= a (locative) nominal *BC* whose complement contains a NP; *BC2* = a directional verbal *BC* whose complement contains a directional verb but no NP.

Table 27 shows large ranges for the accuracy scores for *BC1* and *BC2* in each of the three tasks and four measures. The largest range, 0-100%, was found in four sets of accuracy scores (i.e., accuracy of use of *BC1* in the OPTV and the UGJT-EC, accuracy of use of *BC2* in the OPTV and the UGJT- GJ). The range was 0-90% for the accuracy of use of *BC1* in the OIT and that of *BC2* in the UGJT-EC, and 10%-100% for that of *BC1* in the UGJT- GJ. The smallest range, 0-75%, was found in the accuracy of use of *BC2* in the OIT. These large

ranges suggest that the learners exhibited dramatic variation in their accurate use of *BC* in terms of both *BC* type and task measures. The largest ranges occurred in oral production and the smallest in oral imitation.

The skewness and kurtosis values, especially those above or below 1, indicate some degree of violation of the normal distribution. The negative skewness value for *BC1* in grammaticality judgments in the UGJT indicates a clustering of high values, whereas the positive skewness values for *BC2* on the two oral tasks (i.e., OPTV and OIT) indicate a clustering of low values. Positive kurtosis values for *BC1* in the UGJT-GJ and for *BC2* in oral production in the OPTV indicate a peaked distribution with most scores clustered in the centre, while negative kurtosis values for *BC1* in OPTV indicate a rather flat distribution with many scores in the extremes. This was because a large number of learners (59%) avoided using *BC2* in the OPTV. According to Tabachnick and Fidell (2007), skewness will not 'make a substantive difference in the analyses' with a reasonably large sample (p.80). However, kurtosis will result in an underestimate of the variance and an increased chance of type II errors (i.e., failing to reject a null-hypothesis when it should have been rejected). This was taken into account when analysing the results.

The violation of the assumption of normal distribution was dealt with in two ways. The first was to ignore the violation by running paired-samples *t*-tests. Kinnear and Gray (2000) suggest that a *t*-test can still safely proceed even with moderate violations of these assumptions provided that the samples are not too small, do not contain outliers (atypical scores), and are of equal or nearly equal size. These conditions were met in the present study. For consistency, paired-samples *t*-tests were performed to compare the learners' mean accuracy scores for *BC1* and *BC2* in the four measures. The second solution was to run a Wilcoxon Signed Ranks Test, the non-parametric equivalent of the paired samples *t*-test, to test the three pairs of scores that were not distributed normally.

7.2.1.2 Paired-samples *t*-tests

To investigate the effect of *BC* type on the learners' mean accuracy in using *BC*, paired-samples *t*-tests were performed separately for each of the four measures.

Table 28: Paired-Samples *T*- Tests – Accuracy for *BC1* and *BC2* on the Four Measures (n = 110)

Pairs	Paired Differences						df	Sig.(2-tailed)	Effect size
	M	SD	Std. ER M	95%Confidence Interval oft the Difference					
				Lower	Upper				
OPTV BC1-BC2	.24	.31	2.94	.18	.29	7.99	109	.000	0.37
OIT BC1-BC2	.08	.15	1.41	.05	.11	5.71	109	.000	0.23
UGJT- GJ BC1-BC2	.11	.18	1.76	.07	.14	6.14	109	.000	0.26
UGJT- EC BC1-BC2	.06	.16	1.48	.04	.10	4.61	109	.000	0.16

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT = an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; UGJT- EC = error correction in the UGJT; *BC1*= a (locative) nominal *BC* whose complement contains a NP; *BC2* = a directional verbal *BC* whose complement contains a directional verb but no NP.

.

The results (Table 28) show that the learners' mean accuracy scores were significantly higher for *BC1* than for *BC2* on all four task measures, p < .0005 (two-tailed), with 95% confidence intervals of the difference.

Because four *t*-tests were involved, the Bonferroni adjustment was applied following the formula, convention *p*-value .05/n, to avoid type I error (i.e. an overestimate of the differences between groups leading to a rejection of the null hypothesis when it is true). Using the new significance level of .01 (.05/4 = .01) as a criterion, the differences between accuracy scores for *BC1* and *BC2* were still significant since all paired-samples *t*-tests reached a .0005 significance level.

Of the four measures, the largest difference between the mean accuracy for *BC1* and *BC2* occurred in the OPTV: MD = 24%, *t* (109) = 7.99, p<.0005 (two-tailed). According to Pallant (2007), the effect size (i.e. the magnitude of the effect) for the paired samples *t*-test is assessed by eta squared (one of the most commonly used effect size statistics) using the guidelines proposed by Cohen (Cohen, 1988, pp. 284-7): .01 = small effect, .06 = moderate effect, .14 = large effect. The eta squared is calculated using the formula: $\eta^2 = \dfrac{t^2}{t^2 + N - 1}$,

where N = number of the sample. The eta squared statistic ($\eta^2 = 0.37$) indicated a large effect size for the paired samples *t*-tests in the OPTV. The smallest difference occurred in the error correction in the UGJT: MD = 6.82%, *t* (109) = 4.61, *p* <.0005 (two-tailed). The eta squared ($\eta^2 = 0.16$) still indicated a large effect size.

7.2.1.3 Wilcoxon Signed Rank Tests

Wilcoxon Signed Rank Tests were carried out to compare the mean accuracy scores for *BC1* and *BC2* in the three measures: oral production in the OPTV, oral imitation in the OIT and grammaticality judgment in the UGJT, where the data was not normally distributed.

Table 29: Wilcoxon Signed Rank Tests – Accuracy in the OPTV, OIT, and UGJT-GJ (n = 110)

			Ranks			Task statistics		
Task	Pair		N	Mean Rank%	Sum of Ranks%	Z	Exact Sig. (2-tailed)	Effect size (r)
OPTV	BC2-BC1	Negative Ranks	69	46	2686	-6.60	.000	.44
		Positive Ranks	11	21	395			
		Ties	30					
		Total	110					
OIT	BC2-BC1	Negative Ranks	64	52	3396	-5.32	.000	.36
		Positive Ranks	26	28	700			
		Ties	20					
		Total	110					
UGJT-GJ	BC2-BC1	Negative Ranks	71	52	3672	-5.45	.000	.37
		Positive Ranks	23	35	794			
		Ties	16					
		Total	110					

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP.

The Wilcoxon Signed Rank Tests (Table 29) yielded results consistent with the paired-samples *t*-tests, indicating that the differences between the median of the accuracy scores for

BC1 and *BC2* on all three measures (z = -6.6 in the OPTV, z = -5.32 in the OIT, and z = -5.45 in the UGJT-GJ) were significant, *p*<.0005. The effect size for this test was calculated using the formula: *r* = z/squared root of N, where N = two times of cases, and assessed using Cohen's (1988) criteria: .1 = small effect, .3 = medium effect, .5 = large effect (cited from Pallant, 2007, p. 255). Thus, the *r* values (*r* = .44 in the OPTV, *r* = .36 in the OIT, and *r* =.37 in the UGJT-GJ) suggest medium effect sizes for the differences between the accuracy scores for *BC1* and *BC2* on all the three task measures.

To summarise, the learners achieved significantly greater accuracy in using *BC1* than in using *BC2* in terms of four measures, suggesting that the *BC* type had a clear effect on the learners' accuracy of use.

7.2.2 The effect of task type

The second part of research question one is concerned with the effect of task type on the learners' accuracy of use of *BC*. As Table 27 shows, the skewness and kurtosis values for *BC1* were largely normally distributed (see the next section for a detailed discussion) on each of the four measures, but those for *BC2* were not. Therefore, a one-way analysis of variance (ANOVA) was performed to compare the mean accuracy scores for *BC1* across the four measures, while the Friedman Test, a nonparametric test, was applied to compare the median accuracy scores for *BC2*.

7.2.2.1 One-way repeated measures ANOVA

Prior to running the one-way repeated measures ANOVA, the assumptions of normality and homogeneity of covariance of the data were tested.

As suggested above, the assumption of normality was largely met. Although the skewness and kurtosis values in UGJT-GJ and kurtosis value in OPTV (Table 27) showed a little bit bias, they are sufficiently close to 1 that the biases could be ignored in accordance with Coakes, Steed, and Price (2008). Further, according to Pallant's (2007) claim, most statistical techniques are robust or tolerant of violations of the normality, particularly with large sample sizes (e.g. 100+). Therefore, the biases were not considered likely to cause a major problem in this study (n = 110).

151

Table 30: Mauchly's Test of Sphericity – Task Effect

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
task	.47	81.68	5	.000	.67	.69	.33

The assumption of homogeneity of covariance was assessed by Mauchly's Test of Sphericity. The result is shown in Table 30. The Chi-square value (81.68) is significant, so the assumption of homogeneity of covariance was violated. Therefore, a Bonferroni adjustment was performed to ensure a trustworthy solution (Kinnear & Gray, 2009).

Table 31: Multivariate Tests – Task effect

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Task	Wilks' Lambda	.18	165.30	3	107.00	.000	.82

Table 32: Bonferroni Post-hoc Tests – Task Effect

(I) Task	(J) Task	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
OPTV	OIT	.13*	.03	.000	.06	.19
	UGJT-GJ	-.33*	.03	.000	-.42	-.25
	UGJT-EC	-.15*	.03	.000	-.22	-.08
OIT	UGJT-GJ	-.46*	.02	.000	-.51	-.40
	UGJT-EC	-.28*	.02	.000	-.33	-.23
UGJT-GJ	UGJT-EC	.18*	.01	.000	.14	.22

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; UGJT- EC= error correction in the UGJT; *. The mean difference is significant at the .05 level.

The results of the one-way repeated measures ANOVA (Table 31) show that task had a significant effect on the learners' accuracy scores for *BC1*, Wilks' Lambda =.18, $F (3,107) =$ 165.30, p<.0005. The effect size of this test was assessed by Partial Eta Squared. The result is shown in Table 31 and interpreted by Cohen's guidelines (1988, pp.284 -7): .01 = small, .06 = moderate, .14 = large effect (cited from Pallant, 2007, p.255). The partial eta squared (.82) suggests a very large effect size for the one-way ANOVA.

The Bonferroni *post-hoc* tests (Table 32) suggested that the learners' accuracy scores for *BC1* significantly differed between the four task measures, p<.0005, with a 95% confidence

interval. The accuracy scores followed the order: OIT < OPTV < UGJT-EC < UGJT-GJ, with highest mean accuracy score in the UGJT-GJ and the lowest in the OIT. Figure 1 depicts the profile plot of the mean accuracy scores of *BC1* in the four measures.
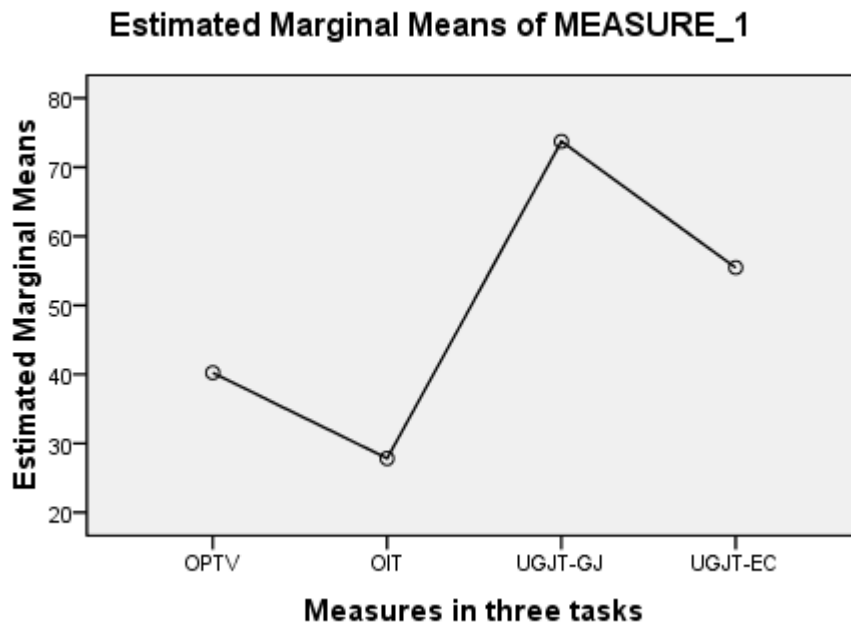
Figure 6: Profile plot of the learners' accuracy scores for *BC1* on the four measures

To summarise, both the one-way repeated measures ANOVA and its *post-hoc* test suggest that task type has a definite effect on the learners' accuracy scores for *BC1*.

7.2.2.2 Friedman Test

Because the data for *BC2* violated the assumption of normality (see Table 27), the hypothesis that task type will have an effect on the learners' accuracy scores for *BC2* was examined by means of the Friedman Test.

Table 33: Friedman Test – Accuracy Scores for *BC2*

Task	Mean Rank	Median	N	Chi-Square	df	Asymp. Sig.
OPTV	1.60	.00	110	214.36	3	.000
OIT	1.73	12.50				
UGJT-GJ	3.72	62.50				
UGJT-EC	2.95	50.00				

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; UGJT- EC= error correction in the UGJT

The Friedman Test (Table 33) indicated that there was a statistically significant difference in the accuracy scores for *BC2* across the four measures in the three tasks, χ^2 (3, n = 110) = 214.36, p < .005. The median accuracy scores for *BC2* follow the order: OPTV (0) < OIT (12.5%) < UGJT-EC (50%) < UGJT-GJ (62.5%).

The *post-hoc* testing involved six individual Wilcoxon Signed Rank Tests. A Bonferonni adjusted alpha value, .05/6 = .01, was used to control for type I error. Effect size statistics were calculated for each specific comparison using the Wilcoxon Signed Rank Tests.

Table 34: Post-hoc Tests and Effect Size – Accuracy Scores for *BC2*

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)	Effect size
OIT - OPTV	Negative Ranks	29	48.50	1406.50	-2.154[a]	.031	-
	Positive Ranks	58	41.75	2421.50			
	Ties	23					
	Total	110					
UGJT-GJ - OPTV	Negative Ranks	9	21.94	197.50	-8.426[a]	.000	0.57
	Positive Ranks	99	57.46	5688.50			
	Ties	2					
	Total	110					
UGJT-EC - OPTV	Negative Ranks	13	33.88	440.50	-7.430[a]	.000	0.50
	Positive Ranks	91	55.16	5019.50			
	Ties	6					
	Total	110					
UGJT- GJ -OIT	Negative Ranks	1	1.00	1.00	-9.062[a]	.000	0.61
	Positive Ranks	108	55.50	5994.00			
	Ties	1					
	Total	110					
UGJT-EC - OIT	Negative Ranks	8	7.94	63.50	-8.829[a]	.000	0.60
	Positive Ranks	100	58.23	5822.50			
	Ties	2					
	Total	110					
UGJT-EC - UGJT-GJ	Negative Ranks	83	52.25	4337.00	-7.659[b]	.000	0.52
	Positive Ranks	12	18.58	223.00			
	Ties	15					
	Total	110					

Note. OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; UGJT- EC= error correction in the UGJT

The *post-hoc* tests (Table 34) showed significant differences between all four task measures, at the *.05* level or higher. However, according to the Bonferonni adjusted alpha value (.01), the difference between the scores in the OIT and the OPTV was non-significant (p >.01). The effect size for the Wilcoxon Signed Rank Test was calculated using the formula: $r = z$/squared root of N (where N = two times of cases). The results were interpreted using Cohen's (1988) criteria: .1 = small effect, .3 = medium effect, .5 = large effect (cited from Pallant, 2007, p.225). Given the *r* values for assessing effect sizes, all significant differences between the pairs of task measures showed large effect sizes (i.e., *r* = .57 for the difference between UGJT- GJ and OPTV, *r* = .50 for the difference between UGJT- EC and OPTV, *r* = .61 for the difference between UGJT- GJ and OIT, *r* =.60 for the difference between the UGJT- EC and OIT, and *r* = .52 for the difference between the UGJT-EC and UGJT-GJ). It can be concluded that learners' accuracy of use of *BC2* significantly differed in three of the measures.

7.3 Discussion

Research question one asked whether there is intra-learner variation in the learners' use of *BC*. The results have established that there is significant variation in the learners' accuracy scores in terms of *BC* type and task type. Firstly, both the Paired-samples *t*-tests and the Wilcoxon Signed Rank Tests (Section 2.2.1) suggested that the learners exhibited consistently greater accuracy in using *BC1* than *BC2* across the four measurements in the three tasks. Secondly, the *post-hoc* Bonforroni test for the One-way repeated measures ANOVA showed that the learners' accuracy scores for *BC1* significantly differed between the four measures in the three tasks. For *BC2*, the Friedman Test and its *post-hoc* tests showed a similar divergence except for a non-significant difference between the OPTV and the OIT.

To explain these results, the following questions will be discussed:
1. What are the factors that can explain why *BC1* was used more accurately than *BC2*?
2. What explains the task variation for *BC1*?
3. What explains the variation in the scores on the different measures for *BC1* and *BC2*?

7.3.1 Explaining learners' greater accuracy on *BC1*

The Paired-samples *t*-tests and the Wilcoxon Signed Rank tests (Tables 28 and 29) showed that the learners manifested significantly greater accuracy in using *BC1* than *BC2* in the four

measures. As the native speakers' performance was used as the baseline for calculating the learners' accuracy scores, a comparison of the native speakers' and learners' use of *BC1* and *BC2* is shown in Table 35.

Table 35: Native Speakers' and L2 Learners' Performance on the Four Measures in the Three Tasks

Measures in tasks	BC type	#OC	NS (n = 22)					NNS (n = 110)				
	BC	#OC	Correct Use (%)	Incorrect Use (%)	Avoided (%)	#OU	Accuracy	Correct Use (%)	Incorrect Use (%)	Avoided (%)	#OU	Accuracy
OPTV	BC1	8	165 (94%)	0	9 (6%)	0	94%	358 (41%)	59 (7%)	458 (52%)	0	41%
	BC2	5	99 (90%)	0	11 (10%)	0	90%	93 (17%)	8 (1%)	451 (82%)	0	17%
OIT	BC1	7	152 (99%)	0	1 (.0%)	3	98%	223 (29%)	85 (11%)	462 (60%)	19	28%
	BC2	7	140 (91%)	0	14 (9%)	2	91%	168 (22%)	39 (5%)	562 (73%)	68	20%
UGJT-GJ	BC1	9	196 (99%)	0	-	0	99%	733 (74%)	238 (24%)	-	-	74%
	BC2	9	187 (94%)	0	-	0	94%	624(63%)	307 (31%)	-	-	63%
UGJT-EC	BC1	8	154 (100%)	0	-	-	100%	484 (55%)	396 (45%)	-	-	55%
	BC2	8	106.5 (97%)	0	-	-	97%	431 (49%)	449 (51%)	-	-	49%

Note. **#OC**= number of obligatory occasions; #OU= number of overused occasions; OPTV= an oral production task prompted by video clips; OIT= an oral imitation task; UGJT= an untimed grammaticality judgement task; UGJT- GJ = untimed grammaticality judgements in the UGJT; UGJT- EC= error correction in the UGJT; *BC1*= a (locative) nominal *BC* whose complement contains a NP; *BC2* = a directional verbal *BC* whose complement contains a directional verb but no NP.

As Table 35 shows, there is a slight difference between the native speakers' use of *BC1* and *BC2*. The difference mainly reflected their choice of whether to use *BC* or another SVO structure. However, these differences were not statistically significant according to the results of the Wilcoxon Signed Rank Test (Table 36), when a Bonferroni adjusted alpha value, .05/4 = .01, was applied.

Table 36: Wilcoxon Signed Rank Test – *BC* Effect - Native speakers (n = 22)

L1		OPTVBC2 - OPTVBC1	OITBC2 – OITBC1	UGJT-GJBC2 – UGJT-GJBC1	UGJT-ECBC2 – UGJT-ECBC1
NS	Z	-2.19	-2.08	-1.98	-1.81
	Asymp. Sig. (2-tailed)	.03	.04	.05	.07

The learners' use of *BC1* and *BC2* was both less proficient and more variable compared to the native speakers' performance. Learners scored more highly for *BC1* than for *BC2*. This difference will be explained in terms of the following factors:

- discourse context (i.e., the external determinants of the use of the target feature)
- functional value (i.e., the form-function mapping of a grammatical feature)
- saliency (i.e., the ease with which formal features attract attention, e.g., phonological property, structural position)
- regularity (i.e., regularity of the grammatical rule)
- potential for L1 transfer (i.e., the potential L1 transfer caused by the similarities or differences between the equivalent structures in L1 and L2)
- ease of elicitation (i.e., the ease with which the target structure can be elicited in the tasks)
- Prototypicality (i.e., whether the target feature is prototypical or peripheral)

Arguably, accuracy scores encompass two components: the decision to choose the target feature, and the accuracy of its use. The former is related to the discourse function determined by the discourse context outside the target feature, while the latter is related to syntactic and semantic functions determined by the relationships of the components within the target structure. The two components will be separately considered below.

- **Discourse context**

The discourse context is related to the discourse function of the target feature, which is determined by the background information or components outside the target feature. The learners' ability to use *BC* in appropriate discourse contexts could be detected relatively clearly in the two oral tasks. This is because the learners relied on their own linguistic resources to produce responses.

While all native speakers used *BC1* and *BC2*, the number of the learners who attempted *BC* varied from task to task. The difference between the number of the learners who attempted *BC1* and *BC2* was the largest in OPTV (75% and 41% respectively). As Table 7.9 shows, the lower accuracy scores for *BC2* in the two oral tasks were more likely due to greater avoidance (i.e., 52% for *BC1* and 82% for *BC2* in the OPTV; 60% for *BC1* and 73% for *BC2* in the OIT) than incorrect use (i.e., 7% for *BC1* and 1% for *BC2* in the OPTV, and 11% for *BC1* and 5% for *BC2* in the OIT). In fact, the percentage of the incorrect use was smaller for *BC2* than for *BC1*. Generally, the learners' overuse of *BC* in the OPTV was far less than in the OIT and UGJT (see Table 35). This may be because the OIT and UGJT contained stimuli with overused errors. This finding suggests that the lower accuracy scores for *BC2* were mainly caused by the learners' difficulty in identifying the discourse context for using it, rather than in composing the structurally correct form. To further compare the differences between *BC1* and *BC2* in their detailed functions and their functional value is discussed below.

- **Functional value**

Functional value mainly concerns the form-function mapping of a linguistic feature. As indicated in Chapter 3, both *BC1* and *BC2* have a complex one-to-many form-function mapping. More specifically, both of them follow the same syntactic structure (S+ ba +NP + V+ R), realise the same semantic structural meaning (i.e. the causer's motion makes causee undergo a change in direction or location and stay in the final state), and contain the same four semantic elements (i.e. causer, causee [or theme], cause [or motion], and effect) (W. Hu, 2004; Sun, 1996; Ye, 2004; B. Zhang, 2000). In addition, both of them are subject to other functional constraints: for instance, discourse constraints (e.g., ba-NP [O] is a sub-topic) (Tsao, 1990), and semantic or pragmatic constraints (e.g., definite or specificity related to ba-NP [O]) (H. Wang, 1985). However, due to the structural difference between the complement

of *BC1* and *BC2*, *BC1* must be used when the effect contains a locative nominal phrase, whereas *BC2* can be used interchangeably with SV(C) O. Therefore, the selection of *BC1* is primarily determined by syntactically obligatory constraints, while the selection of *BC2* is mainly determined by implicit semantic, pragmatic and discourse functions.

The results of this study suggest that a syntactic structure, which is subject to obligatory syntactic functional constraints (e.g. *BC1*), might be easier to learn than a structure which is only subject to semantic or discourse functional constraints (e.g. *BC2*). More specifically, a grammar feature which is only subject to internal constraints might be easier to acquire than one which is subject to external constraints.

Besides the external factors indicated above, two internal factors, saliency and regularity, also explain the results.

- **Saliency**

Saliency is a cognitive construct relating to L2 learners' language processing procedures. The saliency of *BC1* and *BC2* is distinguished here according to two factors that affect how easy a feature is to notice in the input: perceptual or more specifically phonological salience (i.e. how easy it is to hear or perceive a given structure) and syntactic category (i.e. the lexical or functional property of a given structure (cf., R. Ellis, 2006; Goldschneider & DeKeyser, 2001).

Formally, the differences between *BC1b* (e.g. *Ni ba zhuozi na dao/jin fangjian li* [lit. You BA table move to room inside]) and *BC2b* (e.g. *Ni ba zhuozi na jinlai* [lit. You BA table move in (towards the speaker who is in the room)]) lie mainly in the constituents of the complement. That is, the complement of *BC1,* the locative NP, consists of 'a preposition/directional verb + noun + (a locative noun or a directional verb *lai/qu* which indicates towards/away the speaker) (e.g. *dao/jin fangjian li* [lit. to/in room inside]). The directional verbal complement in *BC2* consists of a single or compound directional verb (e.g. *jin* 'in', *jinlai* [lit. *in-come* indicating towards the speaker]. The perceptual salience of the complement of *BC1* and *BC2* is contrasted in terms of three sub-factors: number of phones – one of the factors that Goldschneider and Dekeyser (2001) identified–, number of syllables, and stress. The complement of *BC1* generally contains more phones and syllables (i.e., at least six phones

and three syllables) than that of *BC2* (i.e. at most four phones and two syllables). The stress, by means of "lengthening the duration" or "expanding the tonal range"(Chao, 1968; H. Lin, 2001) on the noun in the complement of *BC1*, lengthens the sound and makes it more salient, whereas the directional complement in *BC2* is normally not stressed[30] (Chao, 1968; Lamarre, 2008) . Hence, *BC1* is more salient than *BC2* based on the assumption that the more phones or syllables, or stressed components, the more likely the structure is to draw learners' attention.

With regards to syntactic category, the complement of *BC1* encompasses both a functional category (i.e., a preposition/ directional verb) and a lexical category (i.e., a noun phrase), but the directional complement in *BC2* mainly executes a functional category (cf., Lamarre, 2008). According to Goldschneider and Dekeyser's assumption that a lexical category is more salient than a functional category, the complement of *BC1* is more salient than that of *BC2*. Thus, it can be concluded that the higher accuracy score for *BC1* might have been due to the fact that the *BC1* is more salient phonologically and syntactically than *BC2*.

- Regularity

The regularity of *BC1* and *BC2* will be distinguished according to two factors defined by Hulstijin and de Graaf (1994) and Ellis (2006, p.436): *scope* (i.e., "number of the cases covered by a particular rule") and *regularity* (i.e., "the extent to which a rule holds true"). Here, the rules relating to the use of *BC* comprise three levels: the rules that govern the entire *BC*, the construction of the complement, and the collocation of V-Complement. First, *BC1* which is subject to an obligatory rule is more regular than *BC2* which is subject to an optional rule. In turn, the obligatory use of *BC1* may make the scope of *BC1* (case of use) broader than that of *BC2*. Second, although the rules for constructing the complement are relatively regular for both *BC1* and *BC2*, the scope of the complement of *BC1* is relatively broad. That is, the locative NP in *BC1* consists of one of two prepositions or seven directional verbs[31] in this study, a noun which can be selected principally from an unlimited number of nouns, and in most cases a final component which can be selected from one of locative nouns[32] and two

---

[30] According to Chao (1968), the directional complement is pronounced as a neutral tone unless it functions as a potential complement.

[31] See note 17 (p.90).

[32] See note 16 (p.90).

directional verbs *lai* 'come'/*qu* 'go'. The directional complement in *BC2* is composed of either one of 9 single or 13 compound andative directional verbs. Third, the rule for constructing VR is more regular for *BC1* than for *BC2*. For example, the prepositions, *zai* 'at' and *dao* 'to' in the complement of *BC1* are relatively productive because they can be collocated with a large number of action verbs serving as the main verb and nouns serving as its objects. In contrast, single directional verbs in *BC2* are more sensitive to the main verb and often used in idiomatic expressions which are subject to a fixed collocation (e.g. *chuan shang* 'put on [a coat]'). In this sense, the VR structure in *BC2* seems subject to item-based learning. In short, *BC1* presents more regular characteristics than *BC2* in terms of scope and regularity of rules.

- **Potential for L1 transfer**

BC is a unique word order structure in Chinese. Undoubtedly, there is no completely equivalent structure performing the same functions as *BC* in the learners' L1s (i.e., English or Korean). Since the major difference between *BC1* and *BC2* lies in their complements, I will be mainly concerned with the relationships between the complements of *BC1* and *BC2* and their equivalent structures in English and Korean (see Table 37).

It is assumed that the simpler the corresponding relationship, the easier to learn the target feature. More specifically, the greater difficulty in using *BC2* may be attributable to the complexity of corresponding relationships between the complements of *BC1* and *BC2* and their equivalent structures in the L1.

Table 37: The Equivalent Structures for the Complements of *BC1* and *BC2* in Chinese, English, and Korean

	Chinese	English	Korean
BC1	**BC1a: S+ ba +O + V + <u>preposition + N. + LN.</u>** E.g. 她 把 书　放 在 桌子　上 　*Tā bǎ shū　fàng <u>zài zhuōzi shàng</u>* She BA book put at table　top. 'She put the book on the table' **BC1b: S+ ba +O + V +<u>directional verb + N.+</u>** **<u>（LN）+ (lai/qu).</u>** E.g. 她 把　桌子　拿 进 房间　（里）（去） 　*Tā bǎ zhuōzi ná　<u>jìn fángjiān (lǐ) (qù)</u>* She BA table　move in　room　(inside)(go) 'She moved the table into the room'	**a. S + V + O + <u>preposition + N.</u>** E.g., She put the book <u>on the desk</u> **b. S + V + O + <u>preposition + N.</u>** E.g., She moved　the table <u>into the room</u>	**(1) S+ (2) O + (3) V** a.그　여자 는　책 을　<u>책상 위</u> 에 놓았습니다 　the woman particle　book particle <u>desk</u> <u>on</u> particle put-past 　'The women put the book on the table' b. 그　여자 는　탁자 를　방　<u>으로</u> 옮겨 놓았습니다. 　The　woman particle table particle <u>room</u> <u>into</u> move-past . 'The women moved　the table <u>into the room</u>: (The equivalent structure of the locative NP in the complement of *BC1* is consisted of a locative NP and a post-positioned particle in Korean. This structure can appear in three positions in the structure as shown by the parentheses.)
BC2	**BC2a: S + ba +O + V+ <u>a single directional</u>** **<u>verb</u>** E.g. 你 把　书 拿 来 　*Nǐ bǎ　shū <u>ná</u> <u>lái</u>.* 　You BA book <u>carry</u> <u>come</u> 　'(Please) <u>bring</u> (your) book' **BC2b: S + ba +O + V+ <u>a compound</u>** **　<u>directional verb</u>** E.g. 你 把　书 拿　出来 　*Nǐ bǎ shū <u>ná</u> <u>chūlái</u>.* 　You BA book take out 　'(Please) take <u>out</u> your book'	The meaning of the directional complement in *BC2* is encompassed in the lexical meaning of a single action verb. **a. S + <u>V</u> +O (e.g. v.= bring/ take)** 　E.g. 'take'= *dai qu/zou* ' carry + go/away' 　'bring' =*dai lai*, 'carry + come' 　　E.g. You <u>bring</u> (your) book **b. S +V +O + <u>adv.</u>** The meaning of the direction is expressed by a preposition or adverb (e.g. in/out) E.g. You take <u>out</u> (your) book	The meaning of the directional complement in *BC2* is encompassed in the lexical meaning of a single verb **a.　S+O+<u>V.</u> (e.g. put on (clothes))** E.g. 책　을　<u>가지고 오세요</u> 　Book　particle　bring **b. S+ O + <u>adv.+ V</u>** The meaning of the directional complement in *BC2* is expressed by a pre-positioned adverb 책　을　<u>꺼내세요</u> Book　particle take out　　('take' and 'out' cannot be separated)

As Table 37 shows, the difference between the locative nominal complement in *BC1* and its equivalent structures in English and Korean mainly lies in the differences in structural construction and positions of the corresponding structures in a sentence. In contrast, the correspondence between the directional complement in *BC2* and its equivalent structures in English and Korean involve both structural and lexical categories, and so are relatively complex. Particularly, lexical categories are involved in conceptual formulation (Levelt, 1989). The direction of an action or movement is sometimes encompassed in the meaning of a single verb in English and Korean (see the examples shown in Table 37), so learners might have not perceived the need to use a directional complement, nor have considered using *BC2*. The relatively complex mapping of the directional verbal complement in *BC2* and its equivalents in English and Korean may have increased the difficulty in using *BC2*. The potential conceptual L1 transfer might have contributed to the lower accuracy scores for *BC2*.

- **Ease of elicitation**

The difference between accuracy scores for *BC1* and *BC2* was much larger in the OPTV than in the other three task measures. This may be because the complement of *BC1* contains a nominal component which is relatively easy to elicit by means of visual images in the video clips. The noun indicating a location is generally explicitly represented by concrete entities in the real world, so it is more likely to attract the learners' attention than the abstract directions involved in *BC2*. For instance, a video clip showed that a lady moved a chair into a room and then she moved it out. When answering the question about what the lady did to the chair, the native speakers generally used *BC2* in the second clause by saying "*Ta ba nei ba yizi ban jin fangjian, ranhou, you ba ta ban chuqu* [lit. she BA that CLASSIFIER chair move in (a) room, then again BA it move out- away from the speaker] She moved a chair into a room, then, she moved it out again". However, some learners produced *BC1* instead of *BC2* by saying, "T*a ba yizi ban dao fangjian li(mian), ta ba yizi ban dao fangjian wai (mian)* [lit. She BA chair move to room inside, she BA chair move to room outside]. She moved a chair into a room. She moved the chair out of the room". This might be because the "room" is more concrete and able to attract the learners' attention than the abstract direction, "out- away from the speaker". Moreover, for native speakers, using *BC1* to repeat the "room" in the second clause seems redundant. Using *BC2* to emphasise the direction of the movement was a better choice. For L2 learners, it may be hard to perceive this functional requirement or to express abstract idiomatic meaning by selecting appropriate directional verbs. The nominal destination related

to the "room" might have served as an obligatory constraint for the learners' selection of *BC1*. Thus, the visual images in the OPTV might have favoured the use of *BC1* over *BC2* as *BC1* contains a locative nominal phrasal complement. This could be one of the possible explanations for why more learners used *BC1* than *BC2* in the OPTV.

- **Prototypicality**

As addressed in Chapter 4, *BC1* can be considered as a prototypical form and *BC2* a peripheral form of *BC*. The finding that the learners' accuracy of use of *BC1* is higher than that of *BC2*, therefore, lends support to the argument that a prototypical form is easier to acquire than a peripheral form (cf., G. Hu, 2002).

7.3.2 Causes of task variation for *BC1*

The second part of research question one addressed the intra-learner variation derived from tasks. The effect of task type on the accuracy scores for *BC1* and *BC2* was examined separately by means of different statistical methods. This was because the learners' use of *BC1* and *BC2* was significantly different. The accuracy scores for *BC1* were normally distributed, but those for *BC2* were not. The sources of task variation in using *BC1* will be discussed first.

The learners' accuracy of use of *BC1* on the four measures follows the order: UGJT- GJ (74%)> UGJT- EC (55%)> OPTV (40%)> OIT (28%) (see Table 27).

The UGJT proved the easiest task. Judging the grammaticality of *BC1* proved significantly easier than correcting errors. The learners' judgement and correction scores on the UGJT were significantly higher than those on the two oral tasks. The OIT proved the most difficult task. The one-way repeated measures ANOVA suggested that task types had significant effects on the learners' accuracy scores for *BC1*.

The task complexity of three pairs of task scores (i.e., oral vs. metalinguistic tasks, the OIT vs. the OPTV, and the UGJT-GJ vs. UGJT- EC) will be explained according to the following dimensions: nature of the learner's response, attention, resources, time pressure, and processing pressure.

7.3.2.1 Oral versus metalinguistic tasks

The differences between the two oral tasks and the metalinguistic task, UGJT, were explained in terms of three factors: nature of the learner's response, instant vs. delayed response, and focus on form.

First, the nature of the response required by a task might have potentially influenced the difficulty which the learners experienced. Arguably, a task that simply requires a learner to judge the grammaticality of sentences (i.e. GJ in the UGJT) will be easier than a task that requires learners to produce sentences. The mean score for the untimed grammaticality judgements was notably higher than that for the other measures, all of which involved production. In addition, the GJ was scored on a different basis from the other tasks (i.e., learners' judgements were scored as 'right' or 'wrong' whereas the production tasks were scored on the basis of their accurate use of *BC1*). Thus, it was possible to achieve 50% on the GJ simply by guessing. This probably explains why the GJ scores in the UGJT were significantly higher than those in the other task measures.

Also, a task where the only production required is that of correcting an ungrammatical sentence (as in the EC) is likely to be easier than a task that requires them to produce complete sentences. This proved to be the case as the learners scored higher on the error correction in the UGJT than on the other production tasks.

Second, time pressure on the learners' response might be a crucial factor for explaining the difference between the two groups of tasks. While both oral tasks (i.e. OPTV and OIT) required instant responses, metalingusitc task (i.e., UGJT including both GJ and EC) allowed for a delayed response – that is, the learners were given as much time as they needed to decide whether each sentence was grammatical and ungrammatical in the UGJT and to correct the sentences they regarded as ungrammatical.

In contrast, the two oral tasks required a more instant response, with less opportunity for learners to search their explicit knowledge of L2 Chinese. This might also have contributed to the lower scores in the OPTV and OIT.

167

Finally, whether or not the learners could focus on form may have affected scores on the different measures. Arguably, the metalinguistic task, UGJT, allowed a greater degree of attention to form than the oral tasks. In particular, the items provided in written form were more stable than those provided in auditory form. In the OPTV learners had to search for the linguistic resources they needed to express their meanings. In the OIT the learners had to process the stimuli they heard which may have directed their attention to meaning, and then had to imitate the sentences in good Chinese. Decoding the linear phonological signals is harder than recognising the written form. However, it is likely that the OIT induced greater attention to form than the OPTV. Given that greater accuracy is likely to ensue when learners are attending to form more closely, I would have expected that the learners would be more accurate in the OIT. However, this was not the case. It would seem then that the differences in the scores for these three tasks cannot be explained entirely by the degree to which they were required to focus on form.

7.3.2.2 Grammaticality judgments versus error correction

The differences between the two metalinguistic measures will be explained in terms of whether the learners could borrow or had to use their own resources. The grammaticality judgements required learners to make judgements about the grammaticality and appropriateness about the sentences. The learners had 50% of chance to get the right answer by guessing. In contrast, in the error correction learners had to rely on their own resources. This may be why the error correction resulted in lower scores than the grammaticality judgments.

7.3.2.3 OPTV versus OIT

Both OPTV and OIT involved oral production. The explanation for the learners' better performance in using *BC1* in the OPTV requires a consideration of the following factors: time pressure, borrowing vs. use of own resources, and processing pressure.

Of the two oral tasks, oral imitation required the more instant responses. The learners were asked to reproduce the sentence under time pressure after listening to the whole sentence. The first clause was repeated in order to prevent learners from rote imitating the aural sounds. In contrast to the response required by the OIT, the response in the OPTV was, to some extent, delayed. When learners watched the video clips in the OPTV, they had to produce an account

of what they had seen 'online'. However, they were permitted to watch them more than once, so as they watched they would have had some opportunity to plan what they wanted to say. Research on pre-task planning (R. Ellis, 2005b; Foster & Skehan, 1996; F. Yuan & R. Ellis, 2003) which draws on Levelt's (1989) model of speaking, shows that pre-task planning can at least sometimes lead to greater accuracy.

A second key difference between the OPTV and the OIT is that the OPTV required learners to use their linguistic resources (i.e., they were given a video prompt to discuss), whereas the OIT provided them with sentences which they had to listen to and then imitate (i.e. potentially they could 'borrow' from the input stimuli). The learners first heard the sentence before they had to produce it in the OIT. Again, one might expect, therefore, that the scores in the OIT would have been higher. However, it can be argued that in fact the OIT placed the students under considerable time pressure as there was little time between hearing a sentence and having to produce a response.

A third explanation for the greater difficulty that the learners experienced with the OIT probably lies in the fact that it placed them under very considerable processing pressure. The OIT required the learners to focus on meaning by asked them to comprehend instantaneous phonological signals and then to re-produce correct sentences within a set time. They were asked to complete these two tasks consecutively. The replaying of background clauses prevented learners from rote memorizing the target clauses. The learners had to complete no less than three receptive stages[33] (cf., Cutler & Clifton, 1999) and five productive procedures[34] (Levelt, Roelofs, & Meyer, 1999) spontaneously. The complex processing procedures may have taxed learners' working memory capacity (Baddeley, 2000). Only the

---

[33] According to Curtler and Clifton (1999), comprehending spoken language involves four stages: *speech decoding* (i.e. selecting discrete speech items from other auditory input), *segmentation* (i.e. segmenting the continuous signal into its component parts), *recognition* (including *word recognition* [e.g. activation of lexical candidates, competition, and retrieval of lexical information] and *utterance interpretation* [i.e. syntactic analysis and thematic processing]), and *integration* (i.e. integrating it into discourse model). The last stage is not relevant to the OIT in this research.

[34] According to Levelt (1989, 1999), speech production follows five procedures: *conceptual preparation*, *grammatical encoding*, *morpho-phonological encoding, phonetic encoding*, *articulation* and *self-perception*. The stages of conceptual preparation and self-perception may not be relevant to the OIT.

speakers who had successfully built up their implicit knowledge of *BC* could successfully complete this task. Therefore, the OIT was the most difficult task.

In addition, the difference between the modalities involved in the two tasks may also explain the lower scores for the OIT. The OIT involved both reception and production, but the OPTV only the latter. Accordingly, the misuse of nouns could be disregarded in scoring the OPTV, but could not in scoring the OIT. This was because misunderstanding any phonological aural form would interfere with the learners' imitation. Thus, the misuse or misunderstanding of other components irrelevant to *BC* did not seriously influence the accuracy scores for the OPTV, but it did for the OIT.

7.3.3 Causes of task variation for *BC1* and *BC2*

In Section 7.2.3, the one-way repeated measures ANOVA, the Friedman Tests and their *post hoc* tests showed a significant difference between the learners' accuracy of use of *BC1* and *BC2* on four measures. However, there are two differences between task variation for *BC1* and *BC2*. First, a statistically significant difference between the OPTV and OIT was evident for *BC1*, but not for *BC2*. Second, the difficulty order of the tasks was OIT < OPTV< UGJT-EC< UGJT-GJ for *BC1*, but OPTV/ OIT < UGJT-EC < UGJT-GJ for *BC2*. More specifically, for *BC1* the OIT was significantly harder than the OPTV, but for *BC2* the OIT and the OPTV were equally difficult. The causes of the significant difference between the OIT and OPTV for *BC1* have been explained in the last section. Here, I will focus on the causes of these differences between *BC1* and *BC2*.

In the OPTV, *BC1* was easier to elicit than *BC2*. This might be partly because *BC1* is subject to both syntactical obligatory constraints and semantic and discourse constraints while *BC2* is only subject to the latter. This may also be because the concrete entities in the video clips might have attracted the learners' attention to *BC1*. In contrast, in the OIT *BC2* was easier to imitate than *BC1*. This might be explained by the following factors: the interaction of the target features and task features, the number of grammatical items, and familiarity of the content of the items. First, *BC2* is shorter than *BC1* in structure (i.e. the complement usually contains at least 3 Chinese characters or syllables in *BC1* but only one or two in *BC2*). Second, one more grammatical item was provided for *BC2* than for *BC1* in the OIT. It was found that for both *BC1* and *BC2* the overused items scored highest, followed by the

grammatical items, with the ungrammatical items lowest. Third, the learners might have been familiar with the content of some *BC2* items or have learned some *BC2* sentences as formulaic expressions. This was supported by the fact that one grammatical *BC2* item (i.e. *Ni yao ba maoyi chuang shang* 'You'd better wear our sweater') was successfully imitated by the learners who did not even attempt any *BC* in the OIT.

The greater avoidance in the OPTV and the ease of imitation in the OIT might explain the fact that the OIT was easier than the OPTV for *BC2* and that the difference between the two tasks was not as significant for *BC2* as it was for *BC1*.

7.3.4 Learners' comments on task difficulty

After the objective analyses of task complexity, the learners' subjective explanations about task difficulty were examined. The majority of learners recognised that the OIT was the most difficult task, though variation did exist among learners. For instance, some learners indicated that all tasks were difficult; some argued that the oral production tasks were the most difficult; and some claimed the UGJT was. Learners' comments on the difficulty of the OIT are discussed from five perspectives.

- **Task design**

A number of learners reported that their "imitation" of the target sentence was disturbed by hearing the repetition of the first clause and the starting signal, "beep".

Jean said, "I think I have remembered the sentence, but once I heard the "ding" [signal sound], my brain went blank".

Han said, "*Wo dong, [keshi] wo wang le* / I understood the sentences, but I forgot them".

Mike said, "I could not stop listening to the first clauses. That has disturbed my memory for the second clauses".

Amy said:

> *Wo keyi gen zhe laoshi shuo, gen zhe dianshi shuo, mei wenti,[keshi] , zhe ge hen nan, wo buneng gai, wo bu zhidao na ge juzi dui bu dui. You de juzi wo mei tingguo, wo xiang na shi xin juzi, wo [hai]mei xue [guo]* / I can imitate what the teacher said in class or TV programs. No problem. It was because they are all correct sentences. However, this [task] is very difficult. I could not correct the sentences because I am not sure if they are correct or not. I

171

thought the sentences that I have never heard might be new expressions that I have not learned.

- **Attention to form**

Many learners said that they could not understand the meaning of the sentence although they imitated part of it. After being told the correct answers, some learners said they knew the words in the items but could not figure out their meaning while listening.

Joe explained, "I could only pay attention to the whole sentence and catch [the] overall meaning, but not pay attention to the small pieces in the sentences. I cannot find the mistakes in the sentences. I didn't know the order was wrong".

Pao reported, "*wo neng mofang, keshi wo bu dong yisi*/ I could imitate (the target clauses), but I don't know their meanings/".

Min said, "I can copy the sounds, but *wo bu ting dong*/ I could not understand/".

Zhung said, "*ting dong de wo ji yisi, ting bu dong wo ji fayin*/ for the sentences that I understood, I tried to memorise their meaning; for those that I could not understand, I tried to copy their pronunciation".

- **Processing pressure**

Some learners complained that the speed of a sentence was too fast and the gaps between items were not long enough for them to finish imitating the target sentence. The time constraints created processing pressure and prevented learners from successfully carrying out the OIT.

Kim said, "*tai kuai, wo bu jizhu/ That was so fast that I could not remember [the target clauses]*". Wu said, "*Wo dong, Wo wang le* / I understood [the sentences], but I forgot/".
Kao, said, "*wo neng zai ting yi ci ma, tai kuai, wo wang le*" /Can I listen to them once again?

[The statements were presented] too fast, so I forgot [them]/.

- **Lack of familiarity with task type and cultural context**

Some learners attributed their poor performance in OIT to lack of familiarity with the task type.

John explained:

> I know every word in the sentence, but I still could not figure out the meaning of the whole sentence. I think that it is because I am not familiar with that cultural context. For example, '*ta ba qian fang zai xie li le* / he hid money in his shoes'. I normally hide money underneath the bed or put it into my wallet, so I could not figure out that the meaning of the sound *xie* was shoes.

Yun said, "A ha! fàngxìn, a , *wo yiwei shi* fàngxīn / I thought it was 'no worries' rather than 'place [the] letter' [at your door]. *Suoyi wo bu mingbai* / That is why I could not figure out the meaning of the sentence".

Some learners claimed that the lack of familiarity with the task type made them anxious and influenced their performance. For instance, learners who were used to rote-learning correct sentences felt too nervous or stressed to instantly complete the two processing stages (i.e. temporarily retaining the phonological signals and reconstructing grammatical ones) at the same time. As Alan said, "ah, *tai lei, xiang tingli kaoshi* , *wo hen/ah*, that made me too tired. It was just like a listening test. I hate it".

- Match of learning and task modalities

After being told the correct answers, some learners said they knew the words in the items but could not figure out their meaning when hearing them. Jon said, "I remembered last semester we did spend a few lessons to learn the *ba* construction. But I was not clear how to use it. I have never used it in my daily communication. Then, it is just gone from my memory."

Pan said, "I might have learned [*ba*] but I forgot".

Park explained:

> *Laoshi jiang guo ba juzi, wo bu dong. Wo yao kao HSK, suoyi wo ziji kan le yufa shu Hanguoyu de, wo hai bu tai mingbai shenme shihou yong* / My teachers have taught me the ba construction in class, but I could not understand. When preparing for the HSK test, I read some Chinese grammar books written in Korean, but I am still not very clear when to use it.

173

Jude, who successfully performed the OIT, said "I remembered that my tutor and I practiced using this structure. We asked each other to move something to somewhere, like put something in the fridge".

Liu said, "*Wo zai yi dui yi keshang lianle ba juzi , wo xianzai shuo de shihou bu xiang yufa, wo juede bu nan*/ I used to practice the *ba* sentences in one-on-one class. I do not think of grammar when I use it now. I don't think this structure is difficult".

In sum, the learners' explanations indicate a number of factors that contributed to the task difficulty of the OIT.

First, task design of the OIT contributed to the task difficultly. In the first place, the learners' rote imitation interfered with the repetition of the prior clauses. In the second place, more than 70% of the items were ungrammatical. This made it difficult for the learners to imitate without understanding. This task design placed relatively high demands on the learners' L2 proficiency.

Second, attention to form cannot facilitate the processing of the OIT. Learners could pay attention to form (phonological signals) in the OIT, but it was difficult to understand the sentences' meaning. This was because homophones are very prevalent in Chinese language. Understanding their meaning, to a large extent, relies upon the background information and context. Therefore, familiarity of context may facilitate this process.

Third, processing pressure made the OIT difficult. Due to limited capacity of short-term memory, the learners' failures may arise from two aspects. On the one hand, the learners could not store the heard original sentences as larger chunks in short-term memory as native speakers could. One the other hand, learners could not instantly complete word recognition and utterance interpretation in the recognition stage of comprehension. In both cases this was probably because the learners generally had relative low proficiency and *BC* had not been successfully internalised as their implicit knowledge. Therefore, the learners needed more time in the process of comprehending spoken phonetic signals. However, time pressure made it impossible and caused the learners' tiredness, stress and nervousness in the processes of decoding, segmenting, and recognizing instantaneous resources.

Fourth, familiarity has been viewed as a key element that determines whether the information in working memory can be stored in long-term memory or retrieved from long-term memory (N. Cowan, 2005). In the SLA literature, familiarity of task type (Bygate, 2001; Gass, Mackey, & Ross-Feldman, 2005; Gass & Varonis, 1994; Robinson, 2001c) and familiarity of task content (Foster & Skehan, 1996; Skehan & Foster, 1997) have been shown to impact on L2 task performance. The learners' explanations here lend support to these findings.

7.4 Conclusion

Chapter 7 has examined the intra-learner variation in the learners' accurate use of *BC*. To this end, it has been established that learners' accuracy scores for *BC* significantly varied according to *BC* type and task type. That is, *BC1* was easier than *BC2* across all four tasks. Accuracy scores for *BC1* significantly differed in the four measures following the difficulty order: OIT > OPTV > UGJT-EC > UGJT-GJ. Variation in the accuracy scores for *BC2* was evident, following the same difficulty order as for BC1 except that the difference between the OIT and the OPTV was not significant.

The *BC* effect was accounted by considering seven dimensions: discourse context, functional value, saliency, regularity, potential for L1 transfer, ease of elicitation and prototypicality. The task variation for *BC1* was explained in terms of task features: nature of the learner's response, focus on form, borrowing vs. own resources, time pressure, and processing pressure. No single factor explains the variation**.** The difference between task variation for *BC1* and *BC2* was explained by both the interaction of the linguistic characteristics of *BC1* and *BC2* and the task features of the OPTV and OIT and the characteristics of the items in the OIT (e.g., the length of the structure, the number of grammatical items, and familiarity of the content of the items). Finally, the learners' comments have borne out the factors explaining the task difficulty of the OIT.

Chapter 8 will investigate the extent to which there is inter-learner variation (i.e., differences among the learners) in the learners' use of *BC* and what sources of this variation are.

CHAPTER 8: INTER-LEARNER VARIATION IN THE INTERLANGUAGE OF *BC*

8.1 Introduction

Chapter 7 examined the intra-learner variation in the learners' use of *BC*. This chapter addresses the second research question regarding inter-learner variation (i.e. variation between learners). This research question is examined in two parts:

2a: To what extent is there inter-learner variation in the learners' use of *BC*?

2b: What are the sources of inter-learner variation?

The chapter begins by introducing data and variables, followed by presenting the results of the statistical analysis and a discussion of the results. It concludes with a summary of the key findings of this chapter.

8.2 Data and variables

The whole learner sample (N = 110) was involved when answering this research question. Two distinct measures (i.e. oral scores and metalinguistic scores) were used as dependent variables for measuring learners' use of *BC1* (i.e., containing a locative nominal phrasal complement) and *BC2* (i.e., containing a directional verbal complement). This decision was made on the basis of the findings of the first research question. That is, while the learners' accuracy scores for *BC1* significantly differed across all four measures in the three tasks, there was no significant difference between their accuracy scores for *BC2* in the oral production and the oral imitation tasks. Also, the learners' metalinguistic judgment and error correction scores were significantly higher than their oral production and oral imitation scores. Therefore, the oral score was calculated by averaging the learners' accuracy scores in the OPTV and the OIT, while the metalinguistic knowledge score was calculated by averaging their grammaticality judgment scores and error correction scores in the UGJT.

Six potential sources of inter-learner variation served as independent variables. These potential sources included one contextual factor – setting, which was defined as the country in which the participants lived when they took part in this study, four individual factors –

starting age, first language (L1), gender, and self-rated proficiency, and one instructional factor – number of years of study.

To explore research question two, three groups of multivariate statistical techniques were conducted. First, the independent samples *t*-test and its non-parametric counterpart, the Mann-Whitney *U* test, were performed to examine group means for the potentially influential factors afforded categorical data (i.e., setting, gender and L1). Second, the Pearson Product-Moment Correlation Coefficients were calculated to examine correlations between the oral/metalinguistic scores and the factors afforded continuous data (i.e., starting age, self-rated proficiency and number of years of study). Finally, four statistically significant factors identified from the above results were used as the independent variables to predict the oral/metalinguistic score for *BC1* and *BC2* using stepwise multiple regression.

8.3 Results – inter-learner variation

8.3.1 Descriptive statistics

The extent of inter-learner variation in the learners' oral/ metalinguistic scores of *BC* was first investigated by means of descriptive statistics.

Table 38: Descriptive Statistics for the Oral/Metalinguistic Scores for *BC1* and *BC2* (n = 110)

Measures	BC type	Mean percentage	Std. Deviation percentage	Skewness	Kurtosis	Range
Oral scores	BC1	34.02	26.30	.24	-1.35	0.00-85.00
	BC2	18.25	19.54	1.45	1.67	0.00-75.00
Metalinguistic	BC1	64.59	20.86	-.59	.17	5.00-100.00
scores	BC2	55.80	20.65	-.37	-.46	00-93.00

Note. *BC1* = A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP.

The descriptive statistics (Table 38) demonstrate notable inter-learner variation in the learners' oral/metalinguistic scores for *BC*. The range of the *BC* scores was large, with the largest range in the metalinguistic scores for *BC1* (5%-100%), and the smallest range in the oral

scores for *BC2* (0-75%). Moreover, the oral scores were lower than the metalinguistic scores for each type of *BC*, with 34.02% and 64.59% for *BC1* and 18.25% and 55.80% for *BC2*.

8.3.2 Inferential statistics

The inter-learner variation was further examined using inferential statistics to investigate the effects of six potentially influential factors.

8.3.2.1 Setting

The learners in this study were recruited in two settings: New Zealand as a foreign language setting and China as a second language setting. The New Zealand (NZ) group comprised 20 learners, and the China (CN) group 90. It was hypothesised that the learners' oral/metalinguistic scores for *BC1* and *BC2* would differ according to each of these settings. Descriptive statistics were first calculated for the two setting groups.

Table 39: Descriptive Statistics for the Setting Groups (n = 110)

Measure	BC type	Setting	N	Mean	Std. Deviation	Range	Skewness	Kurtosis
Oral scores	BC1	NZ	20	6.94	24.40	0-73.75	**1.23**	.12
		CN	90	19.06	11.53	0- 85.00	.10	**-1.33**
	BC2	NZ	20	35.50	21.48	0- 50.00	**3.03**	**10.71**
		CN	90	48.38	19.10	0- 87.50	**1.30**	**1.22**
Metalinguistic scores	BC1	NZ	20	20.76	25.67	5.00- 85.00	-.33	-.34
		CN	90	37.35	20.10	15.00- 100.00	-.59	.37
	BC2	NZ	20	60.31	19.04	0- 80.00	.37	.54
		CN	90	68.19	18.20	10.00- 92.50	-.41	-.36

Note. *BC1* = A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP.

As the results (Table 39) show, the CN group displayed higher oral and metalinguistic scores than the NZ group for both *BC1* and *BC2*. The skewness and kurtosis values in bold demonstrate the oral scores for *BC1* and *BC2* where the assumption of normality was violated. Consequently, the non-parametric Mann-Whitney *U* test was adopted to compare the medians for these four data sets, whereas independent samples *t*-tests were applied to the other four sets of metalinguistic scores.

Table 40: Mann-Whitney U tests for the Setting Groups

Measure	BC	Setting	N	Mean Rank	Median	Z	Exact Sig. (2-tailed)	Exact size
Oral scores	BC1	NZ	20	35.70	10.00	-3.08	.002	.29
		CN	90	59.90	35.00			
	BC2	NZ	20	32.18	6.25	-3.65	.000	.35
		CN	90	60.68	12.50			

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a single or compound directional verb but no NP.

The Mann-Whitney *U* tests (Tables 40) showed that the CN group had significantly higher oral scores than the NZ group for both *BC1* (z = -3.08, *p*<.005) and *BC2* (z = -3.65, *p* <.0005). The effect size *r* was calculated according to the formula: *r = z /* square root of N, where N = total number of cases. The effect size for *BC1* (*r* =.29) would be considered a small effect size, and that for *BC2* (r = .35) a medium one using Cohen's (1988) criteria of .1 = small effect, .3 = medium effect, .5 = large effect.

Table 41: Independent-Samples T-Tests for the Setting Groups

Measure	BC		Levene's Test for Equality of Variances		t-test for Equality of Means				95% CI of the Difference		Effect size
			F	Sig.	t	df	Sig. (2-tailed)	MD	Lower	Upper	
Meta-linguistic scores	*BC1*	Equal variances assumed	.501	.481	-4.11	108.00	.000	-19.82	-29.37	-10.27	.14
	BC2	Equal variances assumed	.014	.906	-5.47	108.00	.000	-24.81	-33.80	-15.81	.22

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a single or compound directional verb but no NP.

The independent samples *t*-tests were then run to compare the metalinguistic scores for *BC1* and *BC2*. The results of which are shown in Table 41. The Levene's tests suggested that equal variances could be assumed since all *p* values were greater than .05.

The independent samples *t*-tests showed that the CN group scored significantly higher than the NZ group in their metalinguistic scores for both *BC1* (*MD* = -19.82%, *t* (108) = -4.11,

$p<.0005$) and *BC2* ($MD = -24.81\%$, t (108) $=-5.47$, $p <.0005$).The effect size eta squared was calculated according to the formula $\eta^2 = \dfrac{t^2}{t^2 + (N1 + N2 - 2)}$ and interpreted using the guidelines proposed by Cohen (1988, pp.284-7): .01 = small effect, .06 = moderate effect, .14 = large effect (see Pallant, 2007, p.236). Thus, the differences between the two setting groups in their metalinguistic scores for both *BC1* (η^2=.22) and *BC2* (η^2=.14) all showed large effect sizes.

In short, the CN group outperformed the NZ group in the oral and metalinguistic tasks for *BC1* and *BC2*.

8.3.2.2 Gender

To probe the hypothesis that gender will have an effect on the oral/metalinguistic scores of *BC*, descriptive statistics was again first calculated.

Table 42: Descriptive Statistics for the Gender Groups

Measure	BC type	Gender	N	Mean	Std. Deviation	Range	Skew-ness	Kurtosis
Oral scores	*BC1*	Female	56	33.48	24.09	.00-77.50	.27	**-1.24**
		Male	54	34.58	28.64	.00-85.00	.21	**-1.51**
	BC2	Female	56	18.93	17.63	.00-62.50	.98	-.02
		Male	54	17.55	21.49	.00-87.50	**1.75**	**2.58**
Metalinguistic scores	*BC1*	Female	56	67.01	18.48	10.00-100.00	-.42	.50
		Male	54	62.08	22.98	5.00-100.00	-.58	-.26
	BC2	Female	56	55.98	21.09	.00-92.50	-.15	-.59
		Male	54	55.60	20.37	5.00-87.50	-.64	-.24

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP.

The descriptive statistics (Table 42) for the gender groups show that the females generally achieved higher mean scores than the males in all measures except in their oral scores for *BC1*. The skewness and kurtosis values in bold indicate that the oral scores for both *BC1* and *BC2* violate the assumption of normal distribution. Thus, the Mann-Whitney *U* test was conducted to compare medians of these scores for the two gender groups, whereas the independent samples *t*-test was performed to compare the means of their metalinguistic scores.

Table 43: Mann-Whitney U Tests for the Gender Groups

Measure	BC type	Gender	N	Mean Rank	Median	Mann-Whitney U	Z	Exact Sig. (2-tailed)	Effect size
Oral scores	BC1	Female	56	55.44	28.13	1508.50	-.02	.99	.00
		Male	54	55.56	35.00				
	BC2	Female	56	58.72	14.38	1331.50	-1.09	.28	.10
		Male	54	52.16	10.00				

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a single or compound directional verb but no NP.

The Mann-Whitney *U* tests (Tables 43) showed the difference between the females and the males in their oral scores for *BC1* (*U* = 1508.50, z = -.02, *p* = .99, *r* =.00) and *BC2* (*U* = 1331.50, z = -1.09, *p* =.28, *r*=.10) was not significant because both *p* values were greater than .05.

Table 44: Independent - Samples T-Tests for the Gender Groups

Measure	BC type		Levene's Test for Equality of Variances		t-test for Equality of Means				95% Confidence Interval of the Difference	
			F	Sig.	t	df	Sig. (2-tailed)	MD	Lower	Upper
Meta-linguistic scores	*BC1*	Equal variances assumed	2.68	.10	1.24	108	.22	4.93	-2.94	12.79
	BC2	Equal variances assumed	.30	.58	.10	108	.92	.38	-7.46	8.22

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a single or compound directional verb but no NP.

The independent samples *t*-tests (Table 44) showed no significant difference between the two gender groups in their metalinguistic scores for both *BC1*, *t* (108) = 1.24, *p* = .22 (two-tailed), and *BC2*, *t* (108) = .10, *p* = .92 (two-tailed).

8.3.2.3 Learners' L1

Descriptive statistics were again first calculated to examine the hypothesis that the L1 will have an effect on the oral/metalinguistic scores.

Table 45: Descriptive Statistics for the L1 Groups

Measure	BC type	L1	N	Mean	Std. Deviation	Range	Skewness	Kurtosis
Oral scores	BC1	English	56	35.02	25.44	.00-85.00	.17	**-1.29**
		Korean	54	32.99	27.37	.00-78.75	.32	**-1.40**
	BC2	English	56	20.29	22.66	.00-87.50	**1.31**	.95
		Korean	54	16.13	15.60	.00-62.50	**1.31**	**1.26**
Metalinguistic scores	BC1	English	56	60.80	23.06	5.00-100.00	-.56	-.08
		Korean	54	68.52	17.67	27.50-100.00	-.24	-.60
	BC2	English	56	50.94	20.77	.00-87.50	-.31	-.31
		Korean	54	60.83	19.46	20.00-92.50	-.44	-.66

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP.

The descriptive statistics (Table 45) show that for both *BC1* and *BC2*, English native speakers (n = 56) exhibited higher oral scores, whereas the Korean learners (n = 54) displayed higher metalinguistic scores.

The skewness and kurtosis values in bold in Table 45 suggest that the oral scores of the two L1 groups are not normally distributed. The Mann-Whitney *U* test was run to compare the medians of these scores for the two L1 groups, while the independent samples *t*-test was run to compare the means of the metalinguistic scores.

Table 46: Mann-Whitney U Test for the L1 Groups

Measure	BC type	L1	N	Mean Rank	Median	Mann-Whitney U	Z	Exact Sig. (2-tailed)
Oral scores	BC1	English	56	56.61	35.00	1450.00	.37	.71
		Korean	54	54.35	26.25			
	BC2	English	56	56.13	11.25	1476.50	-.21	.83
		Korean	54	54.84	12.50			

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP.

The Mann-Whitney *U* Tests (Tables 46) indicated that the differences between the two L1 groups in their oral scores for both *BC1* (z = .37, *p* = 0.71) and *BC2* (z = -.21, *p* = 0.83) were not significant.

Table 47: Independent - Samples T-Tests for the L1 Groups

Measure	BC type		Levene's Test for Equality of Variances		t-test for Equality of Means				95% Confidence Interval of the Difference		Effect size
			F	Sig.	t	df	Sig. (2-tailed)	MD	Lower	Upper	
Metalinguistic scores	BC1	Equal variances assumed	1.99	.16	-1.96	108	.05	-7.71	-15.50	.07	.03
	BC2	Equal variances assumed	.08	.78	-2.58	108	.01	-9.90	-17.51	-2.28	.06

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP.

The independent samples *t*-tests (Table 47) showed that the Korean learners significantly outperformed the English learners in the metalinguistic scores for *BC2*, *t* (108) = -2.58, *p*<.05, with a moderate effect size (η^2 = .06) and for *BC1*, *t* (108) = -1.96, *p* = .05, with a small effect size (η^2 = .06) (see Section 8.3.2.1 for the formula and guidelines).

In summary, the Korean learners outperformed the English learners in the metalinguistic scores for *BC1* and *BC2*. However, there was no significant difference between their oral scores for *BC1* and *BC2*.

8.3.2.4 Starting age

To test the hypothesis that starting age will have an effect on the oral/metalinguistic scores for *BC*; the Pearson Product-Moment Correlation Coefficient was calculated. This method was selected because there was no basis for assigning the whole sample into different groups according to starting age.
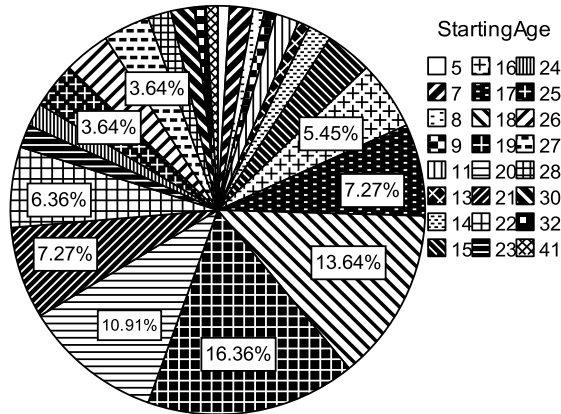
Figure 7: Starting age distribution of the learner sample (n = 110)

The 110 learners began learning Chinese between the ages of 5 to 41, with a mean of 19.6 (*SD* = 5.14). As Figure 7 shows, the largest proportions clustered at age 18 (13.64 %), 19 (16.36%), and 20 (10.91%), with only 8 (7.27%) learners starting learning Chinese before age 13.
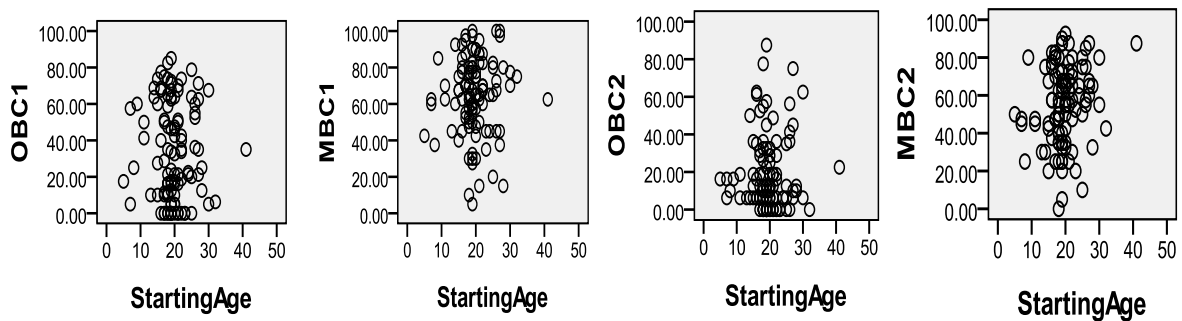
Figure 8: Scatterplot for the oral/metalinguistic scores for *BC1* and *BC2* and starting age

Prior to running the Pearson correlation coefficients, a preliminary analysis was conducted to ensure that the normality, linearity, and homoscedasticity assumptions were not violated. The scatterplots (Figure 8) show the results for the linearity assumption, and they illustrate that no curvilinear relationship exists between the oral/metalinguistic scores for *BC1* and *BC2* and starting age.

Table 48: Pearson Correlations between the Oral/Metalinguistic Scores and Starting Age (n = 110)

Measure	BC type	Starting age		
		Pearson Correlation (r)	Coefficient of determination(r^2)	Sig. (2-tailed)
Oral scores	*BC1*	-.04	0.00	.70
	BC2	.05	0.00	.59
Metalinguistic	*BC1*	.04	0.00	.68
scores	*BC2*	.20˙	0.04	.04

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP.

The results of the Pearson correlation coefficients (Table 48) show that most of the coefficients were weak and non-significant, but a significant, albeit very weak, correlation was observed between starting age and the metalinguistic scores for *BC2*, $r = .20$, $p < .05$. The coefficient of determination ($r^2 =.04$) indicates that starting age can only explain 4% of variance in the learners' metalinguistic scores for *BC2*. The positive direction of the relationship suggests that later beginners scored higher than earlier beginners.

8.3.2.5 Self-rated L2 proficiency

The learners' L2 proficiency scores were collected from the interview with the participants regarding their personal backgrounds. The learners were required to self-rate their proficiency of Chinese in terms of four language skills (i.e. listening, speaking, reading and writing) on a five point Likert scale (i.e. Excellent = 5, Very Good = 4, Good = 3, Average = 2, Poor = 1) (see Question 4 of the background questionnaire in Appendix 15). A self-rated oral proficiency score was calculated by combining listening and speaking scores and a self-rated written proficiency score was calculated by combining reading and writing scores, with a maximum possible score of 10 for each. A self-rated total proficiency score was calculated by combing the above oral and written proficiency scores, with a maximum possible score of 20.

After inspecting scatterplots for a preliminary test, the Pearson Product-Moment Correlation Coefficient was calculated to examine the hypothesis that the oral/metalinguistic scores are related to self-rated proficiency scores.

The results (Table 49) indicate that the self-rated oral proficiency scores were significantly correlated with all measures of *BC1* and *BC2*. All the correlations are of medium strength according to Cohen's (1988, pp.79-81) guidelines (i.e. small $r =.10$ to $.29$; medium $r = .30$ to $.49$; large $r = .50$ to 1.0). Self-rated oral proficiency scores significantly correlated with the oral/ metalinguistic scores for *BC1* and *BC2*, with the highest correlation in the metalinguistic scores for *BC2*, $r =.49$, $p<.01$, and the lowest correlation in the metalinguistic scores for *BC1*, $r =.33$, $p<.01$. The written proficiency scores were significantly correlated with all measures except for the oral scores for *BC1*. The self-rated total proficiency scores were significantly correlated with all measures of *BC1* and *BC2*, with the highest correlation in the metalinguistic scores for *BC2* and the lowest in oral scores for *BC1*. In sum, the oral/metalinguistic scores for the two types of *BC* significantly correlated with the self-rated oral, written and total proficiency scores except there was a non-significant correlation between the oral scores for *BC2* and the self-rated written proficiency scores.

Table 49: Pearson Correlations between the Oral/Metalinguistic Scores and Self-rated Proficiency Scores (n = 110)

Measure	BC type	Self-rated Oral Proficiency	Self-rated Written Proficiency	Self-rated Total Proficiency
		Pearson Correlation (r)	Pearson Correlation (r)	Pearson Correlation (r)
Oral scores	*BC1*	.47**	.25**	.48**
	BC2	.38**	.16	.36**
Metalinguistic	*BC1*	.33**	.34**	.42**
scores	*BC2*	.49**	.29**	.51**

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP. **.Correlation is significant at the 0.01 level (2-tailed). *. Correlation is significant at the 0.05 level (2-tailed).

8.3.2.6 Number of years of study

The number of years of study – the only instructional variable in this study – was measured by years of formal instruction. The months were expressed as years with one decimal place. To investigate the hypothesis that there will be a relationship between the oral/metalinguistic scores and years of study, the Pearson product-moment correlation coefficient was calculated. This method was selected due to the characteristics of the distribution of years of study.
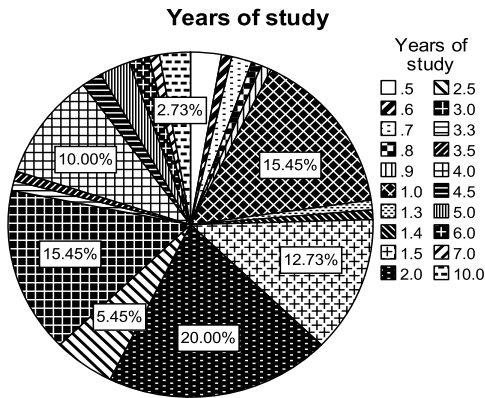
Figure 9: Distribution of years of study

Two and half years ($M = 2.5$, $SD = 1.8$) was the average length of time learners had studied Chinese, with a range of .5 -10 years[35]. As Figure 9 shows, the learners who had studied Chinese for between 1 and 3 years accounted for 71% of the sample population; 20% had studied for 2 years and 15% for both 1 and 3 years respectively.
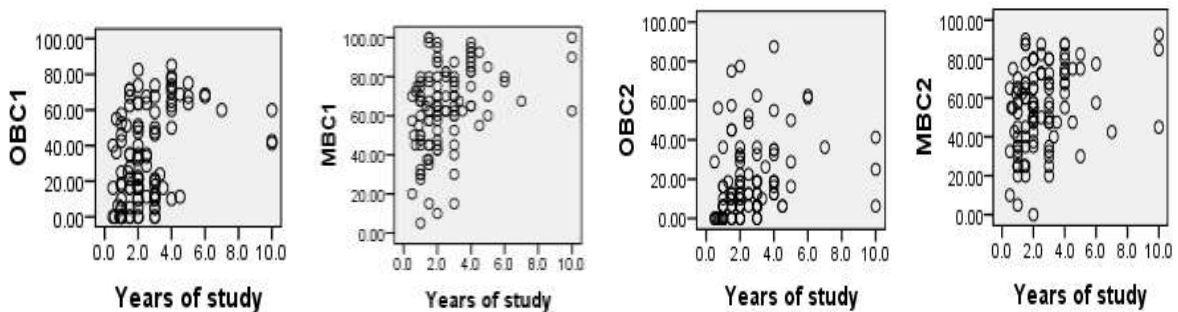
Figure 10: Scatterplots for years of study and the oral/metalinguistic scores for *BC1* and *BC2*

The scatterplots (Figure 10) show no curvilinear relationship between the oral/metalinguistic score and years of study, so the assumption of linearity was met.

The Pearson correlation coefficients (Table 50) show that number of years of study was positively and significantly correlated with all measures of *BC1* and *BC2*. All the correlations showed medium strength according to Cohen's (1988) guidelines (see Section 8.3.2.6), with the highest coefficient in the correlation with the oral scores for *BC1*, $r = .42$, $p < .01$, $r^2 = .18$ and the lowest in the correlation for both oral and metalinguistic scores for *BC2*, $r = .30$, p

---

[35] Generally speaking, only learners who had learned Chinese for at least 1 year were recruited, but two English speaking learners who took instructed Chinese courses for less than one year but still self-studied Chinese in Beijing were considered to qualify.

$< .01$, $r^2 = .09$). The positive directions of the correlations indicate that the learners' oral and metalinguistic scores for both *BC1* and *BC2* increased with years of study.

Table 50: Pearson Correlations between the Oral/Metalinguistic Scores for *BC1* and *BC2* and Years of Study (n = 110)

Measure	BC type	Years of study		
		Pearson correlation (*r*)	Coefficient of determination (r^2)	Sig. (2-tailed)
Oral scores	*BC1*	.42**	0.18	.000
	BC2	.30**	0.09	.000
Metalinguistic scores	*BC1*	.35**	0.12	.001
	BC2	.30**	0.09	.001

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP. **.Correlation is significant at the 0.01 level (2-tailed).

8.4 Sources of inter-learner variation

8.4.1 Variables

In investigating how well the various factors predict the learners' oral/metalinguistic knowledge scores for *BC1* and *BC2*, standard multiple regression and stepwise regression were run. Both tests yielded the same results. Only the results of the stepwise regression are reported below since the choice of predictive variables is carried out by an automated statistical procedure.

Only four significant potential factors (i.e. setting, L1, starting age, and years of study) were used as independent variables. The other factors were not included either because they manifested no significant relationship with the learners' oral/metalinguistic scores for *BC1* and *BC2* (i.e. gender), or because they were not independent and objective predictors of the learners' performance (i.e., the self-rated oral, written, and total proficiency scores). In addition, the three self-rated proficiency scores correlated with each other, violating the assumption of multicollinearity for running a multiple regression. Given that the self-rated total proficiency strongly correlated with the oral/metalinguistic scores (Table 50) and that both oral/metalinguistic scores and proficiency scores are measures of language knowledge, a second stepwise regression was run to test how well the four independent variables predict the self-rated total proficiency.

8.4.2 Assumption Testing

Prior to running a multiple regression, six assumptions were tested: sample size, absence of outliers among the dependent and independent variables; absence of multicollinearity and singularity; absence of outliers in the solution; normality, linearity and homoscedasticity of residuals; and independence of errors (see Tabachnick & Fidell, 2007).

First, according to Tabachnic and Fidell (2007), a standard multiple regression requires at least 104+m cases (where *m* is the number of independent variables), whereas a stepwise multiple regression requires 40m to achieve the same desired power. Thus, the sample size of this study (N =110) satisfies the assumption for running a standard multiple regression. Since both regressions yielded the same results, we assume that this assumption was met.

Second, outliers were checked by inspecting the Manalanobis distances. Four outliers greater than 18.47, (the critical Chi-square values for 4 independent variables shown in Pallant [2007, p.157]) were recoded to fit in the general range since they were just above the cut-off point.

Table 51: Correlation Matrix for Four Independence Variables

Independent variables	Setting	Starting age	L1
Starting age	-0.27	-	-
L1	-0.06	0.08	-
Years of study	-0.08	0.22	0.01

Note. L1 = first language, n = 110, **, p < .01

Third, absence of multicollinearity and singularity was tested by inspecting the correlation matrix for the four independent variables (Table 51). None of the correlation coefficients was higher than .50, suggesting that there was no multicollinearity between independent variables. In addition, the Tolerance values for the independent variables ranged from .8 to .9, which were higher than the cut-off point of .10; the VIF values ranged of from 1.0 to 1.2 well below the cut-off of 10. Therefore, the assumption concerning the absence of multicollinearity was met.

189
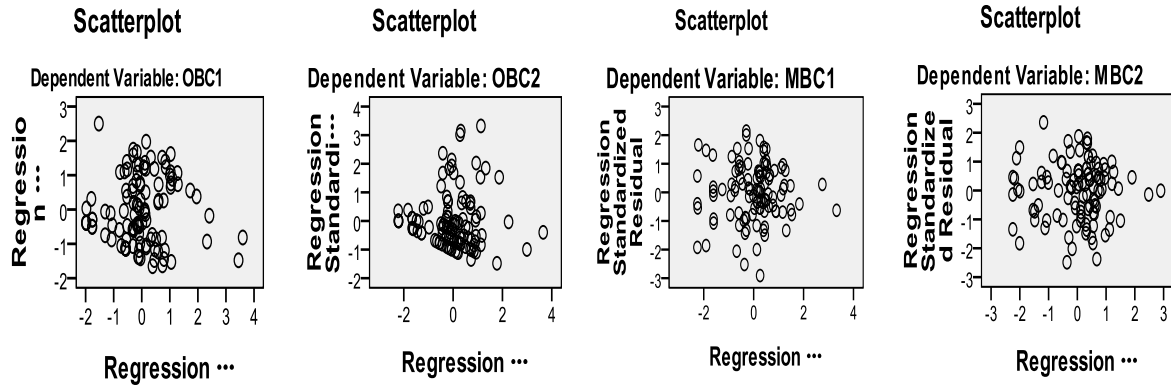
Figure 11: Residual plots of the oral/metalinguistic scores for *BC1* and *BC2* on the 4 independent variables (setting, L1, starting age and years of study)

Fourth, the assumption of absence of outliers in the solution was tested by producing the residual plots of each of the 4 dependent variables on the 4 independent variables (Figure 11). In Figure 11, all residuals fall inside the standardized residuals of ±3.3 (the criteria for a power of $p<.001$ in the Mahalanobis distance according to Tabachnick and Fidell [2007]) except for two outliers (3.32, and 3.4) just being outside. After recoding these outliers, this assumption was met.

Fifth, the assumptions of normality, linearity, and homoscedasticity were examined by inspecting normal P-P plots of the regression standard residuals. As Figure 12 shows, the assumptions of normality, linearity, and homoscedasticity were generally met in all four data sets except for a little curvilinear correlation in the oral scores for *BC2*. This reflected the fact that a large number of learners did not use *BC2* in the two oral tasks, the OPTV and the OIT. Due to a large number of zero values, it was impossible to use a logarithmic transformation. Nor was the curvilinear correlation improved after undertaking a square root transformation. Consequently, the original oral scores for *BC2* were adopted in the standard multiple regression since no evident outlier was present after recoding.
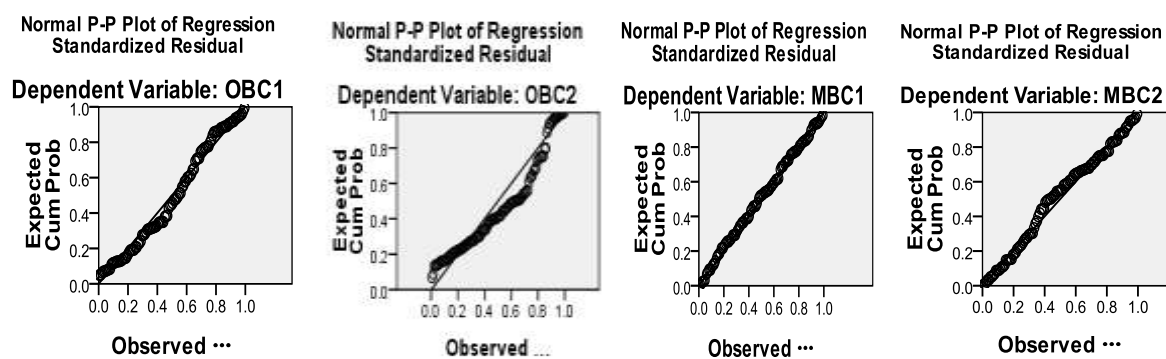
Figure 12: Normal P-P plots of regression standard residual of the three measures for *BC1* and *BC2*

Finally, the independence of errors was assumed since both dependent and independent variables were neither sensitive to, nor associated with, the order of cases (e.g. the function of time or distance data, which normally might cause the violation of this assumption according to Tabachnick and Fidell, 2007).

In summary, the six assumptions for running standard multiple regression were generally met.

8.4.3 Results

Four stepwise multiple regressions were conducted between each of the four dependent variables (i.e. the oral/metalinguistic scores for *BC1* and *BC2*) and the four independent variables (i.e. setting, the L1, starting age, and years of study), with 95% confidence intervals.

The results (Table 52) show that all stepwise regression models were significant beyond the .001 level ($p<.001$ for F values), but the best predictors varied depending upon which dependent variables were considered. In the case of oral scores, two predictors, years of study and setting, explained 21% of variance in the oral scores for *BC1*, and 15% of variance in the oral scores of *BC2*, with years of study having a greater magnitude of effect. In the case of metalinguistic scores, three predictors, setting, years of study, and the L1, predicted 26% of the variance in the metalinguistic score for *BC1*, and four predictors including starting age predicted 35% of the variance in the metalinguistic scores for *BC2*, with setting being the primary contributor as shown by beta.

Table 52: Stepwise Regression Analyses for the Oral/Metalinguistic Scores for *BC1* and *BC2* against Setting, Starting age, L1, and Years of Study (n = 110)

Dependent variables	Predictors	Beta	Adjusted R square	df1,df2	F
Oral scores for *BC1*	Years of study	.39***	.21	2,107	15.47***
	Setting	.26***			
Oral scores for *BC2*	Years of study	.30***	.15	2,107	10.43***
	Setting	.27***			
Metalinguistic scores for *BC1*	Setting	.35***	.26	3,106	13.69***
	Years of study	.34***			
	L1	.17*			
Metalinguistic scores for *BC2*	Setting	.40***	.35	4,105	15.96***
	Years of study	.33***			
	L1	.24***			
	Starting age	.18*			

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP; L1= first language; *p<.05; **p<.01; ***p<.001.

It should be noted that, however, a low adjusted R^2 value (15-35%) across the four regression models indicated that these models could only account for a small portion of the variance. Thus, a number of other unknown factors might have contributed to the variation in the two measures for *BC1* and *BC2*.

8.4.4 Regression for predicting the self-rated proficiency

To examine how well the four independent variables predict the self-rated total proficiency score, a second stepwise regression was run after testing the assumptions. The scatterplot and Normal P-P plot of regression standardized residual are shown in Figure 13.
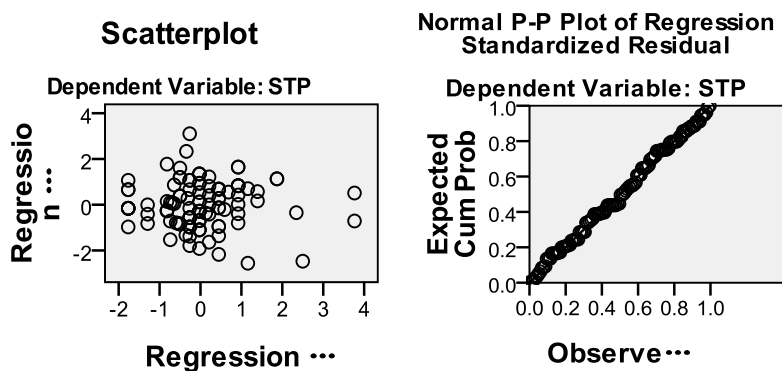
Figure 13: Scatterplot and Normal P-P plot of regression standardized residual between the self-rated total proficiency (STP) and the four independent variables

Table 53: Stepwise Regression Analyses for the Self-rated Total Proficiency against Setting, L1, Starting age, and Years of Study (n = 110)

Dependent Variables	Predictors	β	Adjusted R square	df1,df2	F
STP	Years of study	.41***	.22	2,107	16.03***
	Setting	.24**			

Note. BC1= A (locative) nominal *BC* whose complement contains a NP; BC2 = A directional verbal *BC* whose complement contains a directional verb but no NP; STP= self-rated total proficiency; *p*<.05; **p*<.01; ***p*<.001.

After recoding the two outliers outside ± 3.3 shown in the scatterplot, a stepwise regression was carried out. The results (Table 53) indicate that years of study and setting explained 22% of variance in the self-rated total proficiency score, with years of study showing a greater magnitude of impact.

8.5 Discussion

The main purpose of the second research question was to investigate the extent to which there is inter-learner variation in the learners' oral/metalinguistic scores for *BC1* and *BC2* and what the sources of the variation are. The analysis conducted has achieved these goals to a certain extent. First, descriptive statistics (Table 38) showed clear inter-learner variation in the oral/metalinguistic scores for *BC1* and *BC2*. Second, the independent samples *t*-tests and the Mann-Whitney *U* Tests suggested that setting and the L1, but not gender, had significant effects on the learners' oral and metalinguistic knowledge for *BC1* and *BC2*. The Pearson product-moment correlation coefficients suggested that years of study and self-rated proficiency scores significantly correlated with the oral/metalinguistic scores for *BC1* and *BC2* except for a non-significant relationship between the self-rated written scores and the oral scores for *BC2*. Starting age only correlated with the metalinguistic scores for *BC2*. Finally, the stepwise multiple regression analyses (Table 52) revealed that setting and years of study were the best predictors of the variation in the oral and metalinguistic scores for both *BC1* and *BC2*. The L1 best predicted the variation in the metalinguistic scores for *BC1* and *BC2*, and starting age only predicted the variation in the metalinguistic scores for *BC2*. These results are summarised in Table 54.

Table 54: Sources of Inter-learner Variation

Dependent variables	Predictors
Oral scores for BC1	Years of study > Setting
Oral scores for BC2	Years of study > Setting
Metalinguistic scores for BC1	Setting >Years of study > L1
Metalinguistic scores for BC2	Setting >Years of study > L1 > Starting age

Note. *BC1*= A (locative) nominal *BC* whose complement contains a NP; *BC2* = A directional verbal *BC* whose complement contains a directional verb but no NP.

These findings will be discussed in the order of the magnitude of their effects on the oral or metalinguistic scores for *BC*.

8.5.1 Sources of the variation in the oral scores for *BC1* and *BC2*

The Pearson correlation coefficients (Table 49) revealed that the number of years of study significantly related to the oral scores for both *BC1* and *BC2* with medium strength. The directions of correlation indicated that the longer the study, the better the oral scores. Likewise, the Mann-Whitney *U* tests (Table 40) showed that setting significantly contributed to the variation in the oral scores for both *BC1* and *BC2*, with the China group exhibiting higher oral scores. Finally, the stepwise regression analyses (Table 52) not only confirmed the above results but also suggested that years of study was a better predictor for these oral scores than setting (see Tables 52 and 54). The discussion of these findings will focus on answering the following three questions:

1) *Why did years of study have such an effect on oral scores for BC1 and BC2?*

2) *Why did setting have such an effect on oral scores for BC1 and BC2?*

3) *Why did years of study have a greater influence on oral scores than setting?*

1) Why did years of study have such an effect on oral scores for BC1 and BC2?
The oral scores for *BC1* and *BC2* were analysed for accuracy in using *BC* in the two oral tasks, the OPTV and the OIT, where both speaking and listening oral skills were involved. First, the development of the oral skill required to accurately use *BC* takes a long time. From a pedagogical perspective, *BC*, a non-canonical word order involving many unique properties in Chinese, is usually taught at the late elementary stage in the general Teaching Chinese as a Foreign/Second Language curriculum or presented late in the elementary textbooks that the participants use. For this reason, the learners with a longer period of study have more opportunities to receive input containing *BC*. Arguably, frequency and amount of input will

increase with the length of formal instruction. Some studies have provided evidence that frequency of input could facilitate L2 learning and improve learners' accuracy (e.g. N. Ellis, 2002; N. Ellis & Ferreira–Junior, 2009; Gass, 1997; Gass & Mackey, 2002; Tarone, 2002)

Furthermore, *BC* involves many other features of Chinese (e.g. directional complement, unique prepositional structures, semantic and discourse function of ba-NP, semantic function of the whole structure). Therefore, it is hard for learners to acquire all these features at an early stage of study, particularly to spontaneously use *BC* in both obligatory and optional contexts. Building up this implicit linguistic competence requires learners' exposure to native speakers' speech and incidental acquisition through practice (Paradis, 2009). The oral competence or procedural knowledge involved in oral communication requires learners to master a certain amount of vocabulary and linguistic knowledge and to be able to automatically access these elements based on phonological memory (cf. O'Brien, Segalowitz, Freed, & Collentine, 2007) in specific contexts. This all requires a long period of practice under formal instruction. Thus, the length of formal instruction should be considered a significant factor in promoting learners' oral competence in using *BC*.

However, this finding does not imply that learners' oral competence for using *BC1* and *BC2* will automatically improve with an increase in the number of years of instruction. As Ellis (2009b) found, there was no relationship between years of formal instruction and all his test measures except for that between years of study and the scores for ungrammatical sentences in the UGJT. The differences between the findings of the current study and of R. Ellis's can be attributed to a number of instructional factors (e.g. goals of learning and teaching, instructors, teaching methods, teaching materials, duration and intensity of learning) and the educational setting. In the current study, the majority of the learners (88%) lived in China, a second language environment. Their purpose for learning Chinese was to develop their communicative abilities in using both oral and written Chinese. They learned Chinese in intensive Chinese language programmes or through private tutoring by native Chinese-speaking professional instructors, where oral conversation in Chinese consisted of the main component of the learning process. Arguably, this encouraged the development of implicit or procedural knowledge rather than declarative knowledge, a view supported by the fact that they had probably not received clear explicit instruction in the use of *BC*. The target language, Chinese, was used in every situation and in all teaching materials. The majority of the learners studied Chinese full-time. Thus, although the duration of their learning was not long

(i.e. with an average of 2.5 years), the intensity of learning was very great. Through extensive exposure to the target language, the learners might have been able to build up their oral capability to use *BC*. In contrast, most learners in Ellis's study (70.5%) studied English mostly in China, a foreign language context. The general purpose of learning English there is to achieve good grades or to pass the university entrance examination. The dominant instructional method is focus-on-forms and the teaching language is primarily learners' L1 (Chinese). Thus, the learners are more likely to have developed comprehension skills than oral communication skills. Although the learners experienced on average of 10 years of formal instruction, it was still hard for them to develop procedural knowledge in the foreign language environment without any opportunity to practice in an everyday context. Therefore, the number of years of study in a foreign language context may not necessarily correlate with the learners' oral competence.

2) Why did setting have such an effect on oral scores for BC1 and BC2?

Setting is the second source of inter-leaner variation in the oral scores for *BC1* and *BC2*. The effect of setting on oral scores can be explained by language contact and intensity of learning. Clearly, the learners in China had more opportunities to obtain procedural knowledge of *BC* through contact with the target language and to maintain this knowledge through practice in everyday situations. As Curtain (1997) puts it, the more years devoted to learning a language and the more opportunities available to use it in everyday situations, the greater the proficiency achieved. Most importantly, *BC1* and *BC2* are frequently used by native speakers in their daily life. The learners could receive more input containing *BC* in communicative contexts with native speakers, which would facilitate the development of their oral competence to use *BC1* and *BC2*.

Moreover, the intensity of Chinese learning by the two setting groups differed. The New Zealand group studied Chinese in undergraduate papers with limited learning time. A Chinese course in an academic year consisted of 120 hours' class time (5 hours per week over 24 weeks). The time spent on learning *BC* would be no more than one hour over three years' study. As many New Zealand learners reported, they had forgotten *BC* because they never heard nor used it in everyday situations. In contrast, the learners in the China group generally spent 15-20 hours per week learning Chinese in either intensive immersion or one-on-one programmes. According to the background information they provided, they spent an average

of 10 hours on self-study after class. All the learners in China regularly communicated with native speakers, as well as learning Chinese through multimedia programmes.

3) Why did years of study have a greater influence on oral scores than setting?

The setting in this study did not exactly match the learners' learning environments. For instance, 7 out of 20 (35%) learners in the New Zealand group had learned Chinese in China at some point, whereas 60 out of 90 (67 %) learners in the China group had learned Chinese overseas before coming to China. The majority of the learners, 97 out of 110 (88%), in this study experienced learning Chinese in China. In the New Zealand group, the learners who accurately used *BC1* and *BC2* in the oral tasks generally had learned Chinese for a longer period (i.e. by either starting learning Chinese in primary or secondary school in New Zealand or by having studied Chinese in China for a certain period) than their counterparts who started learning Chinese only at university. In the China group, not all the learners successfully produced *BC* in the oral tasks, particularly those who had studied Chinese for a relatively short period. Thus, years of study had a greater effect on oral scores for *BC1* and *BC2* than setting.

8.5.2 Sources of the variation in the metalinguistic scores for *BC1* and *BC2*

The independent samples *t* - tests (Table 41) revealed that setting had significant effects on the metalingustic scores for *BC1* and *BC2* with large effect sizes. The Pearson Correlation Coefficients (Table 50) showed that the metalinguistic scores for *BC1* and *BC2* correlated with years of study to a medium degree. The independent samples *t* - tests (Table 47) indicated that the L1 had a moderate effect on the metalinguistic scores for *BC1* and a small effect on the metalinguistic scores for *BC2*. The Pearson correlation (Table 48) indicated a weak correlation between starting age and the metalinguistic scores for *BC2*, with late starters exhibiting higher scores. The stepwise regression (Table 52) confirmed that all above four independent variables predicted the variance in the metalinguistic scores for *BC* in the following order of magnitude of effect size: setting > years of study > L1> starting age (see Table 54). The discussion of these results will consider the following points: A) *Why did setting have a stronger effect on the metalinguistic scores for BC than years of study?* B) *Why did the L1 have an effect on the metalingusitic scores for BC1 and BC2 but only on the oral scores?* C) *Why did the starting age have an effect on metalinguisitic scores for BC2?*

8.5.2.1 Explanations for why setting has stronger effects on the metalinguistic scores for *BC* than years of study

The same points that explained why setting and years of study had such an effect on the oral scores for *BC1* and *BC2* also apply to the metalinguistic knowledge scores for *BC1* and *BC2*. However, it is notable that the magnitude of these two sources of variation is reversed.

The metalinguistic scores for *BC1* and *BC2* were obtained from the accuracy scores on the grammaticality judgment and error correction in the UGJT. The learners' ability to recognize Chinese characters may have influenced their performance in these two tasks since there are many homophones in Chinese. As indicated above, the NZ group generally received limited classroom training. Chinese language is only one of four or five papers that they learned for a semester. According to the learners' self-report in the interview about their personal background, they spent only an average of half an hour per week on self-learning after class, which included the time spent on remembering and writing Chinese characters. Therefore, the learners were only exposed to limited Chinese characters, lexicon, grammar, and texts. In contrast, the learners in the China group were surrounded by Chinese characters. Most of them were enrolled in intensive Chinese programmes, which required the learners to grasp about 60-100 words per week, equal to the work load of the New Zealand group for several weeks.

The UGJT in this study required the learners to judge the appropriateness of each sentence after reading it once without time pressure. However, the learners' performance varied according to their proficiency. The more proficient learners made judgments quickly once they finished reading each sentence. They generally claimed that their judgments were based on intuition or grammar rules which they could not verbalise. The low proficiency learners took longer to make judgements and might have tried to access their metalinguistic knowledge. According to their claims, they generally made judgments based on guessing or grammar rules (which were usually incorrect). It has been generally agreed that intuitions about a language can be built up through extensive reading. Thus, the China group, exposed to more vocabulary, more grammar and more reading texts, may have benefited where intuitive knowledge of *BC* is concerned. In respect of metalinguistic knowledge for *BC*, the China group generally outperformed the New Zealand group. This is demonstrated by the fact that nearly 50% of the learners (44/90) in the China group had taken the standard proficiency

test, *HSK*, and the remaining learners in that group were preparing to take it. Metalinguistic knowledge is a key component of the HSK and knowledge about *BC* is one of the key grammar features tested.

That the effect of years of study on the metalinguistic scores for *BC1* and *BC2* was not as strong as setting can be explained in two ways. First, in both setting groups, some learners (particularly English native speakers) started learning reading and writing Chinese characters after they had learned speaking with the assistance of Romanised orthography –*Pinyin* – because recognising Chinese characters is regarded as a challenge when learning Chinese. Second, the learners with a longer period of study in a foreign language setting did not learn as much metalinguistic knowledge of *BC* as their counterparts in the China group. Therefore, the setting had a greater effect on the metalinguistic scores than the number of years of study.

8.5.2.2 Explanations for why the L1 has effects on metalinguistic scores but not on oral scores

The L1 might have played a role in the learner's metalinguistic knowledge. Since there is no exact equivalent structure of *BC* in either Korean or English, speakers in neither of these languages have an advantage in learning *BC*. The Korean learners' better metalinguistic scores can be attributed to narrower distance between the L1 and the L2 and the learning methods that the learners adopted.

First, the closer distance between Korean and Chinese language might have contributed to the Korean learners' better metalinguistic knowledge. The classical Korean writing system borrowed traditional Chinese characters, and loan words in Korean retain similar pronunciation to Chinese words. Therefore, the Korean learners might have had an advantage in reading and writing Chinese characters and pronouncing some Chinese words. In the UGJT, the items were presented in both Chinese characters and *Pinyin*. The ability to recognise Chinese characters may have facilitated understanding the sentences due to the fact that there are a large number of homophones in Chinese. In contrast, reading and writing pictorial characters is one of the obstacles for English-speaking learners in learning Chinese. Many of the English speaking learners had to rely on Pinyin to complete the metalinguistic tasks. Therefore, the Korean learners' advantage in reading Chinese words might have facilitated their understanding of the items in the tasks.

199

A second reason may derive from the learning methods that the two groups of learners had experienced. Based on the information obtained in background interview, the Korean learners spent more time on reading and writing than on speaking. A few of them even taught themselves Chinese by reading written materials when living in Korea. In contrast, the English-speaking learners showed more interest in learning how to speak Chinese relying more on Romanised phonological symbols than on Chinese characters.

The different learning procedures adopted by the L1 learners can be attributed to two facts. First, the Korean learners' approach might have been influenced by the nature of the L1 and the Asian traditional educational model. For instance, to acquire a language using pictorial writing symbols (e.g. Chinese characters, Japanese *kanji*, Korean *hangul*), learners in fact have to learn two different systems (oral and written). Because there is a lack of a clear relationship between written symbols and their actual pronunciation, learning these languages requires a great deal of practice in reading and writing. In contrast, when learning a language using alphabetic writing system (e.g., English), it seems relatively easy to read if one can speak the language. So the requirements for learning the L1 might have influenced the learners' strategies for learning L2 Chinese.

Second, the different learning methods may also be determined by the learners' learning purposes. The primary purpose for learning Chinese was employment for the Korean learners but personal interest for English speaking learners**,** although both L1 groups expected to achieve high academic scores in their Chinese programmes. For the Korean learners, the best way to achieve their goals was to receive degrees from universities either in China or in Korea. To do so, they had to pass the entrance examinations of the universities in China and **t**he standard Chinese proficiency test, *HSK*, in which written skills are foremost, or achieve high academic scores in Chinese courses at universities. The majority of the Korean learners in the China group studied full-time and invested a large amount of time in self-studying Chinese by reading and writing after class. These factors may explain why the Korean learners developed better metalinguistic knowledge of *BC*. In contrast, the English speaking participants tended to develop oral competence prior to written skills to minimize their disadvantage in reading and writing Chinese characters. Many of them even withdrew from intensive Chinese classes where there were a lot of Korean learners and enrolled in one-on-one oral programmes or learned from private professional tutors. In addition, all English-speaking learners in China worked full-time or part-time while studying Chinese, except for

those who were enrolled in short-term intensive study abroad programmes. Most of them learned Chinese for interest or to survive in the local Chinese society. Consequently, they focused on improving their oral communicational skills by speaking only Chinese or by living with native speakers of Chinese who cannot speak English.

8.5.2.3 Explanations for why starting age only has an effect on metalinguisitic scores for *BC2*

The finding that the late starters exhibited higher metalinguistic scores for *BC2* is counter intuitive. To explain this result, the following characteristics of the early and late starters and their Chinese studies need to be addressed: learning environment, intensity of study, motivation, and formal instruction**.** As indicated in Section 8.3.2.4, in this study only 8 learners started learning Chinese before age 13, and 7 of them started learning Chinese 1-2 hours per week at primary or in Sunday school in foreign language environments. Most of them claimed that, although Chinese was taught for interest, they were forced to study Chinese by their parents and were not interested in learning it when they were very young. They did not learn Chinese intensively. The late starters in this study consisted of 10 learners who started learning Chinese from age 27 – 30 with one starting at age 41. Nine of them had been living in China since they started learning Chinese in intensive programmes in China. These mature students were highly motivated to learn Chinese by either living with native speakers or studying with one-on-one professional tutors. Their intensive study in the target language setting might have improved their metalinguistic knowledge of *BC*. Their close contact with native speakers might have developed their abilities to use *BC2* in discourse contexts. This may explain why the late starters had better metalinguistic scores for *BC2*.

8.5.3 The relationship between performance and the self-rated proficiency

The learners' self-rated total proficiency correlated with their oral/metalinguistic scores for *BC* (Table 49). The stepwise regression analyses (Table 53) demonstrated that years of study and setting were the best predictors of the self-rated total proficiency.

This result is unsurprising because both the oral/metalinguistic scores and the self-rated proficiency scores are measures of language knowledge. Interestingly, the relatively strong correlation between the oral/ metalinguistic scores and self-rated proficiency indicates that

whether the learners could accurately use *BC*, a single grammar feature, could itself reflect their proficiency. Because the *BC* involves many unique properties in Chinese (e.g. syntactic, semantic, and discourse functions), learners cannot acquire all of them without having already achieved a high level of general proficiency. The above results also suggest that the learners were able to correctly estimate their proficiency based on their usual performance, learning experience, teachers' comments, and overall test achievement. Because this research was done anonymously, they quite happily provided the researcher with an honest self-assessment of their proficiency in the one-on-one interview.

8.5.4 Non-significant sources of variation in the oral/metalinguistic scores

The Mann-Whitney *U* tests and independent samples *t*-tests in Section 8.3.2.3 and the regression analyses (Table 53) revealed that gender was not a significant source of variation in the learners' oral/metalinguistic scores. Existing evidence for a gender effect mainly lies in certain aspects in phonology (e.g. Eisenstein, 1982; Labov, 1994, 2001) interactional acts (e.g. Gass, 1997) or overall L2 achievement (Boyle, 1987). The use of *BC*, a Chinese grammatical structure, by native speakers of Chinese may differ according to their native dialects (e.g., Du, 2004) but not in terms of gender. Clearly, gender is not a source of variation in the learners' oral ability or their metalinguistic knowledge of *BC* either.

8.6 Conclusion

This chapter has shown that there is clear inter-learner variation in the learners' oral and metalinguistic scores for using *BC1* and *BC2*. The statistical analyses suggest that the number of years of study and setting are the key sources of variation in the learners' oral and metalinguistic scores for both types of *BC*, that the L1 is a source of variation in the metalinguistic scores for both types of *BC*, and that starting age is only a source of variation in the metalinguistic scores for *BC2*. Gender had no effect on the learners' use of *BC*. Self-rated proficiency strongly correlated with both oral and metalinguistic scores for *BC*.

CHAPTER 9: CONCLUSION

9.1 Introduction

The overall aim of this study was to investigate whether there was variation in the English and Korean learners' interlanguage of *BC* and, if so, what the sources of the variation were. Eight potential sources of variability in interlanguage were selected for study on the basis of a number of theories and empirical findings.

The study involved the following steps. First, two types of target *BC* were selected on the basis of five selection criteria (i.e., frequency, productivity, prototypicality, order of instruction, and acquisition order) (see Section 4.3.2) and the findings of two preliminary studies with native speakers of Chinese in NZ (n1 = 10; n2 = 20). Second, the procedures and items of five data collection instruments (i.e., background information questionnaire, OPTV, OIT, UGJT, and a follow-up interview) were piloted on native speakers of Chinese in NZ (n = 22). The items that successfully elicited the two types of *BC* in the three tasks were selected and the procedures were amended. Third, the revised instruments and selected items were tested on both L2 learners in NZ (n = 20) and in China (n = 90). The background information questionnaire was completed orally to collect learners' personal information on their age, gender, L1, starting age, self-rated proficiency, and learning context, etc. The learners' knowledge of the two types of BC was measured in terms of both oral production and metalinguistic understanding. A range of parametric analyses (e.g., paired and independent samples *t*-test, One-way repeated measures ANOVA, Pearson Correlation Coefficients, and Stepwise regression analyses) and non-parametric statistical analyses (e.g., Wilcoxon signed rank test, Mann-Whiteney *U* test, and Friedman test) were employed. These quantitative methods were supplemented with qualitative data– the learners' explanations about their perceptions and performance in the tasks –collected from a follow-up interview.

This chapter concludes this thesis with a summary of the main findings, a discussion of theoretical and practical implications, limitations, and suggestions for future research.

9.2 Summary of findings

This study investigated the variation in the accuracy of use of *BC* by two groups of learners (n = 110). One group (n = 20) studied Chinese in New Zealand and spoke either English (n = 11) or Korean (n = 9) as their L1. The other group (n = 90) was learning Chinese in China and also spoke either English (n = 45) or Korean (n = 45) as their L1. The study investigated two kinds of variability in the learners' use of *BC*- intra learner variation and inter-learner variation. The main findings are summarised below in terms of these two types of variability.

Intra-learner variability:

Intra-learner variability in the accuracy of use of *BC* was predicted by two factors: linguistic difficulty and tasks.

1. Linguistic difficulty explained the intra-learner variation in the accuracy of use of *BC1* and *BC2*. That is, *BC1* which is subject to both syntactic and functional constraints received higher scores than *BC2* which is subject to only functional constraints. The difference in linguistic difficulty was accounted for by seven factors: discourse context, functional value, saliency, regularity, potential for L1 transfer, ease of elicitation, and prototypicality.

2. The scores for *BC1* and *BC2* also varied according to task. That is, the two measures (i.e., grammaticality judgments and error correction) in the metalinguistic task (the UGJT) resulted in markedly higher scores for both *BC1* and *BC2* than the oral production and imitation measures in the two oral tasks (i.e., the OPTV and the OIT). In the UGCT, the grammaticality judgment scores were also significantly higher than the error correction scores for both *BC1* and *BC2*. In the two oral tasks, the oral production scores in the OPTV were significantly higher than the oral imitation scores but only for *BC1*; no significant difference was found between the scores in these two tasks for *BC2*. The overall task effect was explained by five factors: the nature of learners' response, focus on form, borrowing versus using own resources, time pressure, and processing pressure. The difference in the scores of the oral tasks for *BC1* and *BC2* was explained from two perspectives: 1) The video clips in the OPTV favoured *BC1* in which the final destination is indicated by a nominal phrase, while the OIT favoured *BC2* which involves a shorter structure. 2) The qualitative interview data indicated that the learners were more familiar with the items in *BC2* than *BC1*

in the OIT and therefore reproduced them as formulaic expressions. Overall interviews bore out the inherent difficulty of the OIT.

Inter-learner variability:

Inter-learner variability in the use of *BC* was investigated in terms of six potential sources of variation (i.e., number of years of study, setting, L1, starting age, self-rated L2 proficiency, and gender). It was found that the variation in the oral and metalinguistic scores of *BC1* and *BC2* were predicted by four of these six potential sources.

1. The number of years of study was the strongest predictor of variation in the accuracy scores for the oral and metalinguistic tasks for both *BC1* and *BC2*. Positive relationships between the number of years of study and both oral and metalinguistic abilities were found. This relationship can be explained by the facts that 1) the frequency of input increases with the number of years of study, and that 2) acquiring *BC* requires a relatively long time of study.

2. Setting strongly predicted the variation in the oral and metalinguistic scores for both *BC1* and *BC2*. The China group in the second language setting outperformed the New Zealand group in the foreign language setting. This finding was explained by the fact that the learners in the foreign language setting had less opportunity for contact with Chinese.

3. The learners' L1 was the only factor that accounted for differences in the metalinguistic scores for *BC1* and *BC2*. That is, Korean learners outperformed English-speaking learners but only in the metalinguistic tasks. This can be explained by the closer distance between the Korean and Chinese written systems than between the English and Chinese systems, and also by the different purposes and methods that Korean and English learners adopted to learn Chinese.

4. Starting age only predicted the variance only in the metalinguistic scores for *BC2*. The late starters excelled the early starters. Older learners are more likely to develop metalinguistic skills than younger learners. However, the effect of starting age was confounded with that of setting. That is, the learners in the China group started learning Chinese later than the learners in the NZ group. The learners in the second language setting generally took intensive Chinese language lessons to prepare for the demanding

Standard Chinese proficiency exams (e.g., HSK) and admission exams for universities in China.

5. Self-rated proficiency significantly correlated with both oral and metalinguistic scores for *BC1* and *BC2*. Somewhat surprisingly, the learners' accurate use of a single grammar feature, *BC*, predicted their overall proficiency.

6. Gender had no effect on learners' oral and metalinguistic scores for both *BC1* and *BC2*.

In summary, the inter-learner variability in the oral scores for *BC1* and *BC2* were predicted primarily by the number of years of study, and then by setting. The variation in the metalinguistic scores for *BC1* and *BC2* were predicted by setting, and then by the number of years of study and L1. Starting age only predicted the variation in the metalinguistic scores for *BC2*. The stronger effects of setting on metalinguistic scores than on oral scores was accounted for by the fact that learners in the second language setting had more exposure to Chinese characters and were studying intensively towards Chinese written examinations.

9.3 Implications

The current study has a number of theoretical and practical implications.

9.3.1 Theoretical implications

This study has five theoretical implications: 1) it contributes to our understanding of the relationship between systematicity and variability; 2) it addresses the objective difficulty of grammatical structures; 3) it helps to explain task-induced variation; 4) it demonstrates the effect of multiple sources of inter-learner variation on oral and metalinguistic ability, respectively; and 5) it provides an explanation of the functional constraints on the use of *BC*.

9.3.1.1 Revealing the relationship between systematicity and variability

The first implication of this study is that it has provided evidence for the relationship between the two characteristics of interlanguage – systematicity and variability (R. Ellis, 1994, 2008c). Early acquisition studies in SLA displayed great interest in the systematicity of interlanguage and showed that the development of syntactic structures follows a consistent pattern

irrespective of learners' L1, task, and setting. Nevertheless, extensive variability in learner language was also documented. Variability has therefore received growing attention since the onset of SLA (e.g., R. Ellis, 1987c; Larsen-Freeman, 1976; Tarone, 1983, 1985, 1988; Tarone & Parrish, 1988). Research into variability has shown that the variability in interlanguage is influenced by multiple contextual factors which include both internal factors (i.e., various linguistic contexts) and external factors (see Chapter 2). However, the combined influences of psycholinguistic sources, learning difficulty of the target feature, and individual factors on different types of knowledge have not been considered previously. In addition, the relationship between systematicity and variability has not been clearly defined or experimentally tested.

The current study found that the learners' use of *BC* followed the same order across all tasks, that is, the *BC1* was consistently easier than *BC2*. This is clear evidence for systematicity of interlanguage. However, this study also found that the systematicity was influenced by a psycholinguistic factor (i.e., task), individual factors (e.g., starting age, the number of years of study, and L1), and a social factor (i.e., setting). Most importantly, the effects of these factors varied depending upon whether oral or metalinguistic ability was measured. For example, the learners' L1 had an effect on metalinguistic scores but not on oral scores (i.e., Korean learners outperformed English speaker learners in metalinguistic ability but not in the ability to use *BC* orally). In addition, the effects of some factors also depended upon the target feature. For example, starting age only had an effect on the metalinguistic scores for *BC2* (i.e., late starters scored higher than early starters for *BC2* in the metalinguistic task). These findings help our understanding of the complex factors that account for variability.

9.3.1.2 Testing the criteria for determining objective difficulty

The examination of the actual learning difficulty of the two types of *BC* has enabled the adequacy of existing criteria for predicting objective difficulty to be tested. This study found that *BC1* which is subject to both syntactic and functional constraints is easier to acquire than *BC2* which is only subject to functional constraints. This suggests that structures that are only subject to functional constraints will be more difficult than those that are subject to formal syntactic constraints. This provides empirical evidence for Z. Han's (2008) speculation that meaning (or function) is more complex than form in form-meaning relations. The study also lends support to the claim that the prototypical function is acquired earlier than peripheral

function of a form (R. Ellis, 2006; G. Hu, 2002) or that learners first use the most frequent, prototypical, and generic exemplar of a structure (N. Ellis, 2009; N. Ellis & Ferreira–Junior, 2009).

The study identified a number of other factors to account for the learning difficulty of *BC1* and *BC2*. For example, the learners' higher accuracy scores for *BC1* could be explained by comparing the perceptual salience of the complements of *BC1* and *BC2*. That is, the complement of *BC1* contains more phones and syllables and is pronounced normally with more stress than that of *BC2*.

The study also expanded the determinants of objective difficulty by adding discourse context as a factor. The lower accuracy scores for *BC2* were mainly due to the difficulty in identifying the discourse context for using *BC2* rather than that any problem with composing a syntactically correct structure.

Most importantly, this study indicated that these criteria for predicating the objective difficulty of linguistic features were applicable to grammar features in L2 Chinese.

9.3.1.3 Explanations for task-induced variation

The current study found that learners' performance differed according to tasks. The distinctive effects of the two types of tasks (i.e., oral vs. metalinguistic tasks) was explained by the nature of the response required, time pressure, and the possibility of focusing on form. Clearly, the two oral tasks (i.e., OPTV and OIT) were more demanding than the metalinguistic task (i.e., the UGJT) which included two measures: grammaticality judgments and error correction. This was because the oral tasks required the production of complete sentences and instant responses and provided less opportunity for learners to focus on linguistic form. The difference between the two metalinguistic tasks was explained by whether the learners were able to borrow from the input data or had to use their own resources. The difference between the two oral tasks was accounted for by time pressure and processing pressure. In the OIT, the learners were under time pressure to immediately reproduce correct sentences after listening to the whole sentence and then the first clause of the sentence, whereas in the OPTV the learners were able to conduct 'online' planning while watching the video clips. In addition, the OIT involved more processing (i.e., *speech*

decoding, *segmentation*, *recognition*, and *utterance interpretation*) (see, Cutler & Clifton, 1999) than the OPTV.

Task variation has been explained by at least six psycholinguistic models reviewed in the literature (see Chapter 2, pp. 22-30). However, the findings of this study suggest that the task-induced variation should not be attributed to a single source such as attention to form (J. H. Hulstijn & W. Hulstijn, 1984; Tarone, 1985), monitoring (Kormos, 2000; Krashen, 1980, 1981; M. Schmidt, 1980), style-shifting, or discourse conditions (Tarone, 1979, 1982, 1983, 1985, 1987, 1988) but rather to multiple factors. These factors include the nature of the response required by tasks, time pressure, and the processing procedures involved in the tasks.

The OIT was harder than the OPTV. This finding does not accord with the prediction of Tarone's (1983) capability continuum model. Tarone's style-shifting model claims that the styles produced by learners in response to different tasks may be ranged along a continuum depending on the degree of attention to language form required. This model predicts that the learners' accuracy scores for the OIT should fall between those of the OPTV and the UGJT which tap more 'vernacular' and more 'careful' styles, respectively. However, this study found that the OIT was the most challenging task even though the learners were allowed to listen to the stimuli a second time due to their relatively low proficiency of Chinese. The difficulty of the OIT was attributed to the processing pressure required by the task and the task design– the number of ungrammatical or unacceptable items and the clarity of the discourse context of the items.

The results of the study suggest that it is necessary to employ multiple types of tasks to examine learners' vernacular style rather than just interview data as suggested by Tarone (1983, 1985). This is because learners' ability to use a linguistic feature differs in terms of the nature of tasks, and different tasks tap into different abilities. Thus, it is clearly desirable to examine oral and metalinguistic abilities separately when studying the acquisition of L2 grammatical structures.

209

9.3.1.4 Multiple sources of inter-learner variation on oral and metalinguistic ability

The current study has filled a gap in the literature of SLA by investigating the effects of multiple sources of inter-learner variation on oral and metalinguistic tasks (See section 9.1 for a summary of the findings). Although oral production and metalinguistic tasks have been widely used to examine variability in interlangauge, few studies have considered them as measures of different underlying knowledge and examined the effects of the sources of inter-learner variation on the development of these two types of language use.

Furthermore, this research has filled another gap by examining the interlanguage of Chinese. Although the sources of variability have been investigated in the interlanguage of a few Indo-European languages such as English, Spanish, and French (see Chapters 2 and 3), to my knowledge, little research has been done on L2 Chinese.

9.3.1.5 The linguistic characteristics of *BC*

Over the past half century, researchers in Chinese linguistics have made great effort to describe the overall function of the different types *BC* by analyzing samples provided by native speakers or based on the intuitions of linguists. However, it seems that exceptions are unavoidable. This study defined the function of two types of *BC* on the basis of the analyses of native speakers' data in the pilot studies, in corpora, and in the literature. The empirical evidence showed that the distinction between *BC1* (which is subject to both syntactic and functional constrains) and *BC2* (which is subject to only functional constrains) was crucial in explaining the learners' use of these two types of *BC*. This suggests that it may be important to distinguish the prototypical versus peripheral types of *BC* and the syntactically categorical versus variable types of *BC*.

9.3.2 Methodological implications

This study has two methodological implications: 1) it has explored models for investigating variability in interlanguage; and 2) it developed an instrument (i.e., OPTV) to elicit focused oral production data.

9.3.2.1 Exploring models for investigating variability

This study attempted a new approach to examining variability by combining both psycholinguistic and sociolinguistic models. Research into interlanguage variability is typically based on just one of these two theoretical paradigms. The former is concerned with the influence of task, monitoring, and planning on interlanguage (see, R. Ellis, 2008c; Kormos, 1999, 2000; Skehan, 1996, 1998). As R. Ellis (2008c) states, 'a full account of variability in learner language must consider psycholinguistic sources. The key mechanism is attention.' (p. 150). The latter considers both internal factors (i.e., linguistic context) and external factors (e.g., social context, age, gender. and proficiency) (see, Bayley, 1996; N. Ellis & Larsen-Freeman, 2006; Romaine, 2003; Young, 1996). There is a need for an approach that examines multiple sources by drawing on both psycholinguistic and sociolinguistic paradigms. The current study attempted such an approach, to compensate for the drawbacks of the existing approaches.

The development of models for investigating variability in interlanguage has involved two stages. In the first stage, researchers sought for 'a single overarching explanation' for interlanguage variation (see, Bayley, 2005; Beebe, 1980; R. Ellis, 1987b; Selinker & Douglas, 1985; Tarone, 1985). In the second stage, variationists, particularly those following the multi-factor model, made an effort to explore the relative strength of the influences of multiple contextual factors (see, Bayley, 2005; Bayley & Preston, 1996; Young, 1989, 1991; Young & Bayley, 1996). As indicated above, this model only considers contextual factors but not psycholinguistic aspects of language processing.

Further, there are some limitations in implementing the multi-factor model. First, this model cannot be easily used to investigate grammatical structures whose linguistic contexts are unclear (e.g., *BC*). Second, the employment of VARBRUL (a common statistical tool used to investigate variability) ideally requires a large number of tokens collected from sociolinguistic interviews – elicited data using metalinguistic tasks are not suitable. In particular, for a target feature which L2 learners generally avoid using (e.g., *BC2*), the data will be limited. Moreover, the need for a large number of tokens also restricts the number of the subjects that can be investigated (e.g., normally around 20 learners were adopted in the studies following this model). This limits the generalisation of the findings.

To avoid the above restrictions, the current study employed a number of different statistical analyses to investigate the influences of multiple sources of variation. A psycholinguistic factor, task, was included besides individual factors and a sociolinguistic factor, setting. Task was considered a psycholinguistic factor rather than a sociolinguistic factor in this study because the various tasks were designed to tap different types of knowledge. Furthermore, by using stepwise regression analyses, the current model investigated the relative strengths of the effects of multiple sources of inter-learner variation (e.g., the number of years of study, setting, L1 and starting age) on both oral and metalinguistic abilities in using the target feature (i.e., *BC1* vs. *BC2*). The interaction between the sources of intra-learner and inter-learner variation was clearly demonstrated. In addition, the current model involved a relatively large sample of L2 learners (n = 110) drawn from diverse institutions. This makes the generalisation of the findings more reliable.

9.3.2.2 Developing an instrument to elicit focused oral production data

The current study developed an instrument to elicit *BC*, a language feature which is relatively hard to elicit. The difficulty in eliciting *BC* is due to a number of factors. First, *BC* is used under certain conditions and does not occur frequently in daily conversation. This makes it very difficult to collect naturally occurring data. Second, L2 learners generally tend to avoid using *BC*, and so it is difficult to elicit the targeted *BC*. Third, *BC1* and *BC2* involve a consideration of causer, known causee, action, and effect in the same structure, and so cannot be easily elicited by using static pictures. Fourth, participants' oral production prompted by interview topics resulted in varying numbers of obligatory contexts for *BC,* making comparison problematic.

Using silent films or video clips to collect data is not uncommon in the SLA literature (Chafe, 1980; Du, 2004; H. G. Jin, 1992, 1993; Skehan & Foster, 1997, 1999; Tarone, 1985). However, these studies either used existing films to elicit narrative production, or used video clips which are not suitable for eliciting *BC*. In addition, although Du (2004) also designed video clips to elicit *BC*, the rationale for producing the video clips was the definiteness of *ba-*NP which is not a necessary constraint for using *BC*.

The current study developed the OPTV to elicit *BC1* and *BC2*. The OPTV was trialled on native speakers of Chinese to establish obligatory occasions for using *BC*. The questions used

to prompt learner responses were also carefully designed to elicit responses that included causer, the known causee (stuff), action, and effect in one sentence using *BC*.

9.3.3 Pedagogical implications

The findings of this study can be used as basis for developing a pedagogical grammar of *BC*, for designing a task-based teaching curriculum, and for establishing programs that take account of individual learner factors.

9.3.3.1 Developing pedagogical grammar of *BC*

The difference in the functional constraints of two types of *BC* needs to be taken into account when developing a pedagogical grammar of *BC*. This study found that functional value and discourse context were two important factors in determining linguistic difficulty. The explanations of the functional constraints of *BC*, particularly *BC2*, in this study may help to formulate explicit explanations of *BC* and thus help learners to develop metalinguistic understanding of the two *BC* structures.

This study classified the two types of *BC* according to whether the post-verbal constituent contains an NP or not. The results showed that accuracy scores for *BC1* were significantly higher than those for *BC2*. This suggests that this difference effectively distinguishes the two types of *BC* and the difference in the difficulty learners experience in learning them. Thus, applying this classification in designing a curriculum for TCSOL may assist L2 learners' to learn *BC*.

This study also shed light on (see Chapter 4) the relationship of the three basic word order structures in Chinese, SVO, OSV and *BC*. That is, the definiteness of O in the three word order structures differs, and so determines their use in the discourse. In addition, this study found that SVO structures were used by native speakers of Chinese to emphasize events denoted by action verbs, while *BCs* were used to stress the effect of the action denoted by the main verb on the entity known to the speaker. Explicitly indicating these conditions for the use of SVO and *BC* in a pedagogical grammar may facilitate teaching and learning of word order structures in Chinese.

9.3.3.2 Teaching and presenting order of *BC*

The acquisition order that the current study found is different from the presentation order in current textbooks. This study found that *BC1* was acquired earlier than *BC2*, whereas the textbooks (reviewed in Chapter 4) present either both *BC1* and *BC2* at the same time (Yuehua Liu, Yao, Shi, & Bi, 2005; Sanders & Yao, 2009; J. Xu & Ren, 2004; N. Yao, Lee, & Sanders, 2009; T.-c. Yao, et al., 2005) or *BC2* first (Kang & Lai, 1999; X. Liu, 2003). *BC2* seems simpler than *BC1* in syntactic structure. However, *BC2* has been shown to be harder to acquire if the criterion for acquisition is spontaneous and accurate use. Based on this finding, this study suggests that *BC1* constitutes a prototypical type of *BC* and therefore should be taught earlier than *BC2*, which is a peripheral type. Textbook writers need to take account of this acquisition order when deciding the presentation order.

In addition, textbook writers should consider the functional meaning of grammar features such as *BC* when preparing teaching materials. The current textbooks tend to introduce *BC* in terms of syntactic formal structures and by explaining the properties of each component of *BC* (e.g., definite or specific *ba*-NP, the semantic function of 'disposal' of the main verb, and obligatory complement). There is a lack of explicit explanation of the discourse conditions that govern when *BC* should be used. If explicit explanation**s** of the conditions of use of *BC* are provided (see Section 4.2.3 and Section 9.3.3.1), learners will develop clear metalinguistic knowledge of *BC*, particularly *BC2*.

9.3.3.3 Designing a task-based curriculum and teaching techniques to elicit BC

The finding that oral and metalinguistic abilities are different skills suggests that teachers should consider assisting learners to develop both types of skills through different kinds of activities. To develop learners' metalinguistic knowledge of *BC*, traditional teaching tasks – grammaticality judgments and error correction– will be helpful. To develop learners' oral ability, task-based teaching methods need to be employed. For example, the tailored video clips and prompt questions used in the OPTV could be used in classroom teaching. In class, the teacher can ask students to perform the task in pairs. Student A is allowed to watch the video clips, but Student B is not. Student B can ask Student A the questions about the content of the video clips using all kinds of questions. After they finish talking about the video clips, the pairs need to report to the class, with Student A repeating the questions and Student B giving the answers. Other students could watch video clips and check if these students had

encoded all the information displayed in the video clips. In the end, the teacher should compare the questions which have elicited *BC* with those which have elicited SVO, and list the different types of information encoded by *BC* and SVO sentences. The learners in this study indicated that they found the OPTV a novel and interesting activity.

9.4 Limitations

Like all empirical studies, this study has a number of limitations. The limitations mainly lie in sampling, task design, and data analysis.

9.4.1 Sampling

The major limitation of this study lies in sampling the participants. First, the native speakers of Chinese were recruited in New Zealand. Ideally, the control group should be comparable to the non-native speaker group (see, L. White & Genesee, 1996). Native controls should be monolingual Chinese speakers in China. According to the findings of my pilot studies with native speakers and Du's (2004) finding, native speakers use *BC* variably according to their native dialects. Thus, native controls should be monolingual speakers from Mandarin speaking regions in China. However, for logistical reasons this was not possible.

Second, there were a few limitations in sampling the learner population.

1) The sample size of the NZ group was small. This was constrained by the limited number of Chinese learners in NZ.

2) The sample sizes of the NZ group (n = 20) and the CN group (n = 90) were not balanced. In addition, in China the overall population of Korean learners of Chinese was much larger than that of English speakers. Although the current study made every effort to recruit the same number of learners with relatively equal levels of Chinese proficiency, this was not possible.

3) The number of years that the learners had been studying was not long enough. *BC* is a late acquired linguistic feature, so ideally learners with a relatively advanced proficiency level in Chinese needed to be recruited. However, it was very hard to find English speaking learners

who had studied Chinese for a long period and reached an advanced proficiency level. For this reason, the current study was unable to examine learners' ultimate achievement of *BC*.

4) Another limitation with respect to learners' proficiency level was that the relatively low proficiency of the learners may have influenced their ability to perform the tasks, particularly, the most demanding task, the OIT.

9.4.2 Task design

The number of sentences for each type of *BC* was limited (no more than 10). This is because it was necessary to control the vocabulary level of the items designed to measure knowledge of *BC* to take account of the low proficiency levels of the learners.

The tasks were completed in a fixed order (i.e., OPTV < OIT < UGJT-GJ < UGJT-EC). Ideally the order of the task should have been counterbalanced to control for any possible task effect. However, it was necessary for the learners to do the oral tasks first as the two metalinguistic tasks may have made them aware of the need to use *ba*.

The greater difference in the accuracy scores of *BC1* and *BC2* in the OPTV than in the other tasks might have been because *BC1* is easier to elicit by video clips than *BC2*. It seems that the visual images in the OPTV video clips favoured the use of *BC1* which contains a nominal destination or location rather than *BC2* which contains a direction of the movement.

In the OIT, the learners were allowed to listen to the stimuli a second time due to their relatively low proficiency. So the learners were able to pay more attention while listening. As a result, their imitation reflected a style closer to a 'careful' style than a 'vernacular style'. Therefore, the results of this study were unable to test Tarone's style-shifting continuum model.

9.4.3 Data analysis

There are three limitations in the way the data were analysed. First, target-like analysis was employed to analyse the data of both *BC1* and *BC2*. Second, the lower accuracy of *BC2* was due to avoidance rather than incorrect use of *BC2*. This led to a biased distribution of *BC2*

and required the use of non-parametric statistics. Finally, the effects of setting and starting age in this study cannot be clearly distinguished. For instance, a few learners in NZ had a short exposure to China, while most of the learners in China had prior learning experience in a foreign language setting. The effects of starting age were confounded with setting because some learners started learning Chinese in the second language setting but others in a foreign language setting.

9.5 Future research

This study also suggests a few directions for future research.

9.5.1 Difference between NS and NNS in their use of *BC*

To reveal the causes of variation in the use of *BC* by native and non-native speakers, future research will need to collect data from both an oral production task such as the OPTI and a metalinguistic knowledge task and also conduct stimulated recall to develop an understanding of why speakers sometimes use *BC* and sometimes do not.

9.5.2 Investigating the acquisition of *BC* in different settings

An interesting topic for future research could be to compare the use of *BC2* which is subject to variable and functional constraints by advanced learners in a foreign language setting with those in the second language setting. The finding may contribute to our current understanding of the effect of learning contexts in SLA.

It would also be interesting to investigate the effect of instructional factors (e.g., preferred teaching methods/actual teaching methods experienced) and age-related factors (e.g., age of arrival, length of exposure, and length of residence). These sources may also contribute to explanations of the variation in the learners' use of *BC*.

9.5.3 Designing a valid and reliable self-rating proficiency instrument

This study found that self-rated proficiency is viable as a replacement for measured proficiency under the following conditions: 1) The information was obtained in a one-on-one interview before learners performed the tasks. 2) The learners were informed that the scores

obtained from their own ratings would not affect their course grades. Self-rating instruments are rarely used in investigating learners of L2 Chinese but they have considerable potential.

This study suggests that a self-rated proficiency questionnaire should include questions about self-rated proficiency in reading, writing, listening, and speaking. It would also be an idea to ask questions about factors which can have significant effect on learning, for example, contextual factors (e.g., exposure to native speakers of the target language, exposure to TV and radio, the length of immersion in a second language environment), and age–related individual factors (e.g., L1, starting age, length of classroom learning, and length of self-study).

9.5.4 Examining the functional constraints of *BC* using function-form analysis

Finally, using function-form analysis to analyse qualitative data of *BC1* and *BC2* may help to provide a clear explanation of the development of the acquisition of *BC* by English and Korean learners. This is because the primary difference between *BC1* and *BC2* lies in the fact that *BC1* is subject to both syntactic and functional constraints while *BC2* is subject to functional constraints only. Using a function-form analysis can effectively reveal interlanguage development by examining how linguistic forms are related to specific functions in particular discourse contexts. Oral production data in the OPTV is best suited for a function-form analysis. First, the functions to be investigated in each context should be determined. Second, the different forms used to accomplish that function will be identified and their frequency tabulated. By comparing the frequency of the forms that were used by learners at different proficiency levels, we can obtain an idea of the learners' developmental route in using different forms (e.g., *BC1* or *BC2,* or SVO structures) to realise a specific function.

APPENDICES

Appendix 1: The classifications of *BC* in the previous studies

Criterion of Classification	Representative	Data	Categories
Structural type	X. Li *et al.* (1990)	843 BC tokens from 70 children 2-5 years old	9 types 17 subtypes
	Jingling Liu (1996)	Structural syllabus for L2 learners	3 general types
	H. Jin (1992)	Based on Li and Thompson (1981) and Tao (1987,1990)	5 types 1.S *be* NP V le 2.S *be* NP1 $V_{ditransitive}$ NP2 (NP2 is Sb.) 3.S *ba* NP V NM NM(no.+measure word) 4. S *ba* NP1 V NP2 (NP2 is part of NP1) 5. S *ba* NP1 V NP2 (NP1 is a noun indicating a location)
	Zhao et al. (1997)	*BC* tokens in the textbooks for primary school pupils in China	32 subtypes
Structure + Semantic	W. Lü (1994)	1094 *BC* tokens from modern Chinese literature	6 types 19 subtypes
	L. Jin (1997)	Based on studies of Cui (1995) and Hsueh (1994)	3 types 1. A causes B to change 2. When A acts on B, A or B stays in some situation. 3. A conducts an action on B
	Cui (1995)	1265 BC tokens from the early modern novel 'Dream of the Red Chamber' 361 BC tokens from modern fiction	2 types (result type) (situation type) 9 subtypes
	W. Zhang (2001)	2160 *BC* tokens from the China Daily newspaper A typical *ba* sentence is the process of spatial displacement of an object under the action of an external force	Four variant forms, namely, relevant, equivalent, changing and resultative schemata, can be derived from the schema representing this process by means of metaphorical mapping.

Semantic types	Xiong (1996)	199 BC tokens from L2 Chinese textbooks	4 types
			Related object + ba subject + new
			1. location
			2. owner
			3. situation/ thought
			4. characteristics
	Chen (2005)	Based on 7710 BC tokens collected from two corpora of Modern Chinese in Taiwan	Frequency order of 9 semantic types of *BC*
			1. RVC (a verb or stative verb as a post-verbal constituent) (35.85%)
			2. Target one (*zai* 'at'/ *dao* 'to'/ *gei* 'to' animate +NP) (31.88%)
			3. Directive verb (16%)
			Target two (*cheng* 'into' / *zuo*/*wei* 'as'+NP)(5.22%)
			5. Aspect marker (3.67%)
			6. De CSC (2.63%)
			7. Others (1.97%)
			8. Scopes (1.7%)
			9. Verb reduplication (1.06%)

Appendix 2: The verbs and stative verb used as a RC of *BC* in LCMC

A total of 139 verbs serve as the resultative complement of *BC* collected from the written data corresponding to the spoken style in LCMC. They are shown below with the frequency of occurrence followed.

Stative verb	Frequency	Stative verb	Frequency	Stative verb	Frequency
kāi 开 'open'	31	zhù 住 'hold'	19	diào 掉 'drop'	18
zǒu 走 'walk, away'	12	sǐ 死 'die'	6	wán 完 'complete'	5
guāng 光 'finish'	8	duàn 断 'break'	4	dǎo 倒 'fall'	3
xǐng 醒 'wake'	3	huó 活 'live'	3	tòu 透 'thorough, penetrate'	3
jìn 尽 'complete'	2	chuān 穿 'through'	2	miè 灭 'extinguish'	2
pò 破 'break'	2	tānhuàn 瘫痪 'break down, palsy'	1	sàn 散 'loose'	1
shā 杀 'kill'	1	tā 塌 'collapse'	1	shāng 伤 'hurt'	1
dǒng 懂 'understand'	1	pǎo 跑 'run'	1	kuì 溃 'burst'	1
kū 哭 'cry'	1	tuō 脱 'cast'	1	dào 到 'arrive'	1
fān 翻 'turn over'	1	zuì 醉 'drunk'	1	zháo 着 'burn'	1
diū 丢 'lose'	1				

Appendix 3: The stative verbs used as a RC of *BC* in LCMC

A total of 107 verbs serving as the resultative complement of *BC* in LCMC are shown below and followed by its frequency of appearance.

Stative verb	Frequency	Stative verb	Frequency	Stative verb	Frequency
hǎo 好 'good'	35	jìng 净 'clean'	7	qīngchǔ 清 楚 'clear'	6
gānjìng 干 净 'clean'	5	píng 平 'flat'	5	gān 干 'dry'	3
dà 大 'big'	4	luàn 乱 'confused, in a mess'	2	suì 碎 'broken, fragmentary'	2
sǐ 死 'inflexible'	2	huài 坏 'bad'	2	xiǎo 小 'small'	1
yǎ 哑 'mute, dumb'	1	gāo 高 'tall, high'	1	hún 混 'confused',	1
zhāshi 扎 实 'sturdy'	1	zhèng 正 'upright'	1	shī 湿 'wet'	1
làn 烂 'mashed, mushy'	1	yǎn 严 'tight'	1	jǐn 紧 'taut, tight'	1
liàng 亮 'bright'	1	yùn 晕 'dizzy'	1	shú 熟 'cooked'	1
biǎn 扁 'flat'	1	pò 破 'damaged'	1	qí 齐 'neat'	1
shí 实 'solid'	1	dāi 呆 'dull'	1	qīng 青 'blue'	1
duì 对 'right'	1	cuò 错 'wrong'	1	hóng 红 'red'	1
sōng 松 'loose'	1	zhǔn 准 'accurate'	1	míngbái 明 白 'clear'	1
wěn 稳 'steady'	1	huāng 慌 'flurried'	1	lèng 愣 'absent-minded'	1
zāng 脏 'dirty'	1	bào 倒 'inverted'	1	láo 牢 'firm, durable'	1
kōng 空 'empty'	1	huài 坏 'break down'	1	fēng 疯 'crazy'	1
qīng 清 'distinct'	1				

Appendix 4: Frequency of BA patterns based on production data in Cheung (1992)

	RVC	DE	PP	LE	ERR	FRG	TOTAL
5 yr	45 (32%)	3 (2%)	78 (55%)	8 (5%)	13 (9%)	3 (2%)	142
6 yr	37 (22%)	2 (1%)	105 (62%)	9 (5%)	22 (13%)	2 (2%)	169
Adult	98 (27%)		260 (72%)	45 (12%)	-	2 (1%)	360

Note. RVC= verb is resultative verb compound (e.g. stative verb hao 'properly' in 'ta ba yifu gua hao'/

he hung the clothes); DE= De introduces a modifier clause; PP= a locative phrase is used after the

complex verb; Le= the use of aspectual marker le; ERR= errors in using BA; FRG= sentence

fragments in using BA.

Appendix 5: BA patterns in Cheung's (1992) found by analysing the spontaneous speech data in Tse *et al.* (1991)

Group	Pse VP	V-comp	VV	RVC	DE	3 Arg	LE	ZHE	NEG	ERR	FRG	Total BA	Total Utt
2 yr	3	1	3	49	2	0	13	0	0	17	13	92	9547
3 yr	3	1	3	52	1	0	9	0	0	22	9	96	8819
4 yr	2	3	5	148	4	6	20	2	5	25	34	244	14743
5 yr	0	3	0	100	4	5	22	2	1	22	34	183	10050

Note. Pse-vp=pseudo-verb phrase '*ta ba ta zenmeyang*'; v-comp=complement clause '*ta ba yifu na le qu xi*';
VV=verb reduplication to express the sketchiness; ZHE= durative marker *zhe*; NEG= perfective negation *meiyou*;
3 Arg= BC involving 3 argument predicates such as bitransitive verbs;

Appendix 6: Map of the area of Mandarin Dialects in Mainland China
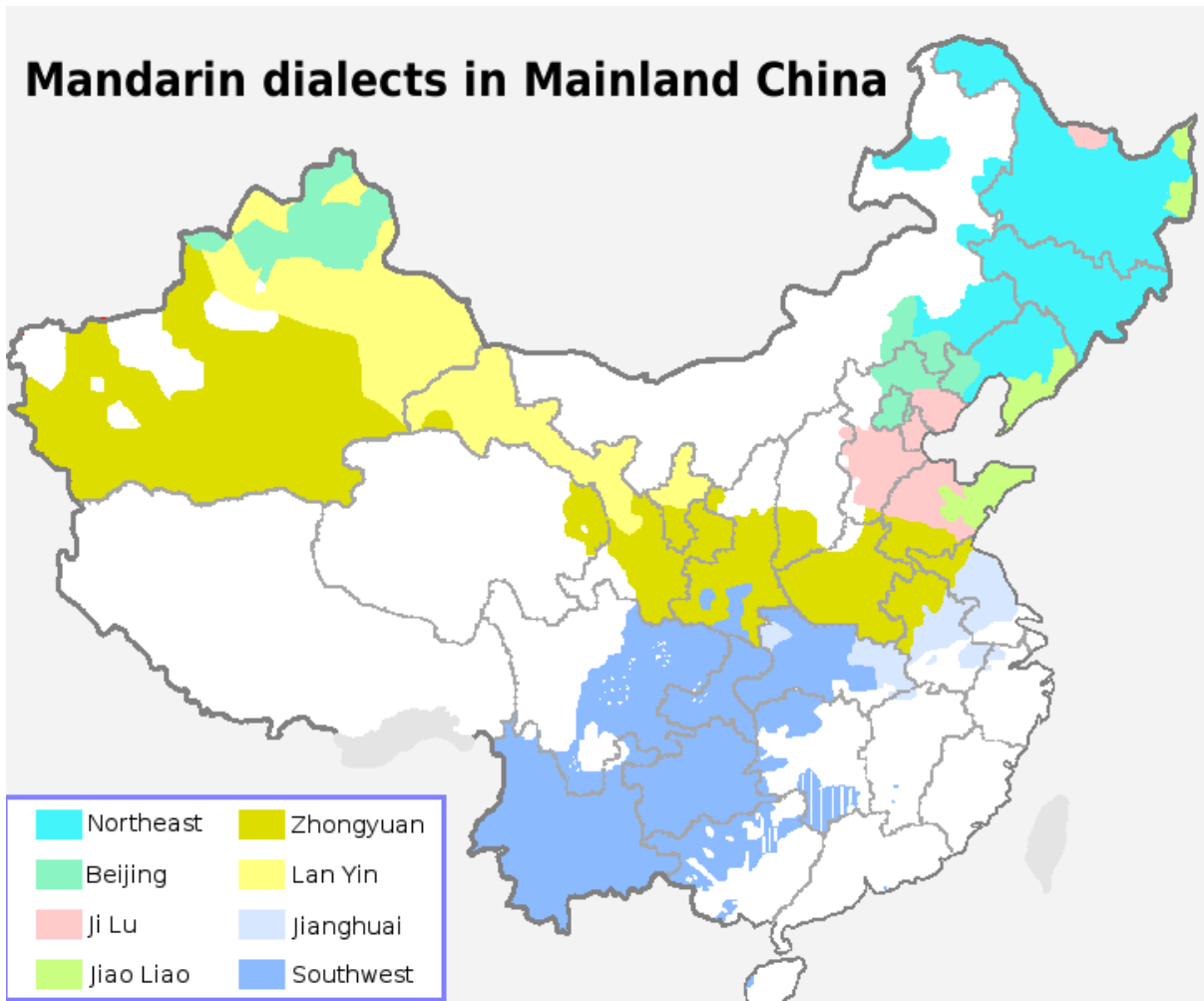
http://en.wikipedia.org/wiki/File:Madarin_in_Chinese_Mainland_EN.PNG

Appendix 7: Participant information sheet (English and Chinese version)

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

<div align="right">

Department of Applied Language Studies and Linguistics
Fisher Building
18 Waterloo Quadrant,
Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360
Private Bag 92019
Auckland, New Zealand

</div>

Participant Information Sheet (PIS)

To: All participants

Title: Variability in the use of some Chinese grammar feature by L2 learners

I am a doctoral student at The University of Auckland working in the area of Language Teaching and Learning. I am conducting this research to investigate what sources contribute to the variability in the use of some Chinese grammar feature by English and Korean speaking learners of Chinese. I attempt to help second language learners to learn this grammar feature effectively through the findings.

Who will be invited to this research?
English or Korean native speakers who have learned Chinese in classroom more than one year are expected to participate in my research. I would very much like your help.

What will I do?
I will have a one-on-one interview with you outside of class. You are required to fill in a background questionnaire and conduct the following six tasks orally in approximately 2 hours.
- An oral production task prompted by interview topics;
- An oral production task prompted by video clips;
- An oral imitation task;
- A untimed grammaticality judgement task;
- A metalinguistic knowledge task;
- A stimulated recall

How will it benefit you?
Through carrying out these tasks, you can get a chance to practice your oral Chinese and my professional feedback on your oral production. It will facilitate your learning of Chinese.

What will I do with the information collected from you?
I will audio-record the interviews. It will be only me who will transcribe all the audio recordings. All the audio tapes, discs and transcripts will be stored in a locked cabinet at university for a period of six years. After completing this study, I will store them in a locked cabinet for a period of six years for peer reviewed publication and further research purposes only. Afterwards, I will destroy all the information by deleting the files and shredding the hardcopies.

You are voluntarily making a decision whether or not to participate in this research study. Whether you decide to take part in or not will not affect your relationship with me and the School and your grades in any way. If you want to withdraw your data after completing the interview, you may do this

within one month without giving any reason. Just let me know and I will remove any information about you.

If you want to help me in my study, please sign the Consent Form. I will not use your real name in my thesis and my publications, so no one will know who you are.

Thank you very much for making this study possible. If you have any queries or wish to know more information please contact me:

Xiaoping Gao
Department of Applied Language Studies and Linguistics
The University of Auckland
Private Bag 92019
Auckland
New Zealand
Email: gxp1999@gmail.com; or gxia002@aucklanduni.ac.nz
Phone number: 0064 9373 7599 ext.86338

My main supervisor: Professor Rod Ellis
Department of Applied Language Studies and Linguistics
The University of Auckland
Private Bag 92019
Auckland
New Zealand
Email: r.ellis@auckland.ac.nz
Tel: 0064 9373 7599 ext. 84876

Head of Department: Associate professor John Read
Department of Applied Language Studies and Linguistics
The University of Auckland
Private Bag 92019
Auckland
New Zealand
Email: ja.read@auckland.ac.nz
Tel: 0064 9373 7599 ext. 87673

For any queries regarding ethical concerns please contact:
The Chair, The University of Auckland Human Participants Ethics Committee (UAHPEC),
Room 005, Alfred Nathan House, 24 Princes Street, The University of Auckland,
Tel: 373-7599 ext. 87830

Reference No. 2008/205 Approved by The University of Auckland Human Participants

Ethics Committee on 11 July 2008 for a period of six years.

Department of Applied Language Studies and Linguistics
Fisher Building
18 Waterloo Quadrant,
Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360
Private Bag 92019
Auckland, New Zealand

调研情况说明

致： 参加者

课题： 二语学习者使用汉语某语法结构中介语的变异性研究

我是奥克兰大学第二语言教学方向博士研究生，我从事的这项研究主要调查第二语言学习者使用汉语某语法结构时的差异及其成因，希望研究结果能帮助二语学习者有效学习汉语。

如您的母语是英语或韩语，并且已在校学习中文一年以上，我希望您能参与这项调查。

调查将在课外以一对一谈话的方式进行，您需完成一份背景资料调查和六项任务：话题交流、录像导述、口头模仿、不限时语法判断、语法知识测试、和刺激性回忆。整个调查需要 1 到 2 个小时。

您可通过完成这些任务练习口语，并得到专业性指导和 40 元补贴。

谈话过程会被录音，并转写成文字。所有录音文档和转录文本都将在奥克兰大学加密保存 6 年，再由我个人保存 6 年，作为发表文章和进一步研究之用。之后，所有的信息将会被清除。

参加这项研究纯属自愿。您在调查中的表现与在校成绩无关。您可以在参加面谈后一个月内要求无条件退出，取消您提供的语料。

如您同意参加本调查，请在调研同意书上签字。您的真实姓名将不会在论文和发表的文章中出现。

感谢您参与这项调查，如有问题或需更多了解，请联系研究者：

高小平 Xiaoping Gao
Department of Applied Language Studies and Linguistics
应用语言研究及语言学系
The University of Auckland
奥克兰大学
Private Bag 92019
私人邮箱 92019
Auckland 奥克兰
New Zealand 新西兰
Email: gxp1999@gmail.com; or gxia002@aucklanduni.ac.nz
Phone number: 0064 9373 7599 ext.86338
中国联系电话：13011010255

导师: Rod Ellis 教授
Department of Applied Language Studies and Linguistics

应用语言研究及语言学系
The University of Auckland
奥克兰大学
Private Bag 92019
私人邮箱 92019
Auckland 奥克兰
New Zealand 新西兰
Email: r.ellis@auckland.ac.nz
Tel: 0064 9373 7599 ext. 84876

系主任: John Read 副教授
Department of Applied Language Studies and Linguistics
应用语言研究及语言学系
The University of Auckland
奥克兰大学
Private Bag 92019
私人邮箱 92019
Auckland 奥克兰
New Zealand 新西兰
Email: ja.read@auckland.ac.nz
Tel: 0064 9373 7599 ext. 87673

如有有关道德规范问题，请联系：
奥克兰大学人文调研道德规范委员会(UAHPEC)主席
Room 005, Alfred Nathan House, 24 Princes Street, The University of Auckland,
Tel: 373-7599 ext. 87830

Appendix 8: Consent form (English and Chinese version)

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND

<div align="right">

Department of Applied Language Studies and Linguistics
Fisher Building
18 Waterloo Quadrant,
Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360
Private Bag 92019
Auckland, New Zealand

</div>

<div align="center">

**Consent Form
(For all participants)**

</div>

Title of project: Variability in the use of some Chinese grammar feature by L2 learners

Researcher: Xiaoping Gao

Statements:

* This research will keep your personal information secret. The final report on this research will not use your name.

* If there is a study relationship between you and the researcher, your participation or non-participation will not affect your relationship with the researcher and your study.

* Your speech produced when conducting the tasks will be audio-recorded and accessed only by the researcher.

* The data gathered for this project will be saved onto disks and stored in a locked cabinet at university for a period of six years. After this, it will remain in the researcher's home office for six years for the peer reviewed publication and research that be further developed. Further approvals will be sought prior to commencement of the future research.

* This consent form will be stored at the University separately from the data for six years and destroyed afterwards.

* You will not be provided with the data produced in this project unless you request it.

* You will not be offered the opportunity to edit the transcripts of the recordings.

* You may request a summary of the results by contacting the researcher.

* You have the right to withdraw your information or data within one month without giving a reason.

Declaration:

I, (name)_____ agree to take part in this research after reading and understanding all of the statements.

<div align="center">Signed: _____   Date: _____</div>

Reference No. 2008/205 Approved by The University of Auckland Human Participants

Ethics Committee on 11 July 2008 for a period of six years.

Department of Applied Language Studies and Linguistics
Fisher Building
18 Waterloo Quadrant,
Levels 8 & 9
Auckland, New Zealand
Telephone 64 9 3737599 ext. 82001
Facsimile 64 9 308 2360
Private Bag 92019
Auckland, New Zealand

调研同意书

（参加者）

✸ **研究课题: 外语学习者使用汉语某语法结构的变异性**

✸ **研究者姓名**: 高小平

✸ **内容：**

✸ 本研究的最终报告将确保您个人信息的机密性和匿名性。本研究成果发表时，不会使用您的真实姓名，您所提供的一切信息都将保密，且仅供本研究使用。

✸ 如您与研究者有师生关系，无论您是否参加本研究都不影响师生关系，也不会对您的学习造成任何影响。

✸ 您在参与调查时将被数码录音机录音。

✸ 您在调查中产生的原始语料将存入光盘，在奥克兰大学办公室保存 6 年,之后由研究者个人保存 10 年，以供今后研究和发表文章时使用。

✸ 参加者同意表将在研究者办公室保存 6 年后销毁。

✸ 除非特别要求，否则调查中所获得的语料将不提供给参加者。

✸ 本研究不提供参加者修改其录音文本的机会。

✸ 参加者可向研究者索要最终的研究结果。

✸ 参加者有权在参与调查后一个月内撤回所提供的信息和数据。

声明：我_____（楷书）已仔细阅读以上内容，同意参加这项调查。

签名：_____日期: _____

奥克兰大学人文调研道德规范委员会调研许可编号：2008/205，自 2008 年 7 月 11 日生效，有效期 6 年

Appendix 9: Background information questionnaire for native speakers of Chinese (English and Chinese version)

Background questionnaire for native speakers of Chinese

1. Please write your age and circle your gender:

_____

A. Male   B. Female

2. Please fill in the blanks by using the provided choices. (Please specify the particular dialect or language).

The language and dialect that I mostly use at home is _____.

The language and dialect that I mostly use at school or at work place in China is

_____.

The language and dialect that I use mostly use in New Zealand is _____.

    A. Chinese, Mandarin

    B. Chinese, dialect , _____

    C. English,

    D. Another language, _____

3. What is your major?

_____

4. How long have you lived in New Zealand?

_____

5. Who do you live with?

    A. Members of your family

    B. Other Chinese speaking people

    C. By yourself

    D. With English speaking people

☻Thank you very much for participating in this project! ☻

汉语母语者语言背景调查

姓名:＿＿＿＿＿＿＿＿   **编号：**＿＿＿＿＿＿＿＿＿

1. 请注明您的性别和年龄

    A． 男   B.女

    年龄：＿＿＿＿＿＿＿

2. 请在下列四项中选择，并在横向线上填写特定的方言。

    您在家里最常用的语言或方言是＿＿＿＿＿＿

    在中国时，您在学校、工作单位或社会上最常用的语言是＿＿＿＿＿＿

    在新西兰，您最常用的语言或方言是＿＿＿＿＿

        E. 汉语普通话

        F. 汉语方言 ＿＿＿＿＿＿（请指明特定的方言或语言）

        G. 英语

        H. 其他语言＿＿＿＿＿＿（请指明特定的语言）

3. 您的专业是什么？

        ＿＿＿＿＿＿

4. 您来新西兰多久了？你在别的英语国家住过吗？

        ＿＿＿＿＿＿＿＿＿

5. 您跟谁一起住？

    A. 家人 （您家人说什么语言或方言 ＿＿＿＿＿＿ ）

    B. 其他说中文的人

    C. 自己

    D. 说英文的人

☺感谢您参与这项调查 ☺

Appendix 10: The list of actress's actions in the OPTV

No.	The actress's actions	Type
1	tā bǎ shūbāo fàng dàole zhuōzishàng 她 把 书 包　放　到 了　桌 子 上　。 'She put her bag on the table.'	(BC1)
2	tā bǎyǐzi fàng dàole zhuōzǐ xiàmian 她 把椅子　放　到 了　桌 子 下 面　。 'She put all the chairs underneath the table.'	(BC1)
3	tā bǎ shū cóngzhuōzishàng náqǐlái fàng dàoleshūjiàshàng 她 把 书　从　桌 子 上　拿起来，放　到 了书架　上　。 'She took a book from the table，and put it on the bookcase.'	(BC2)
4	tā bǎ píngmù shōu shàngqù le 她 把 屏 幕　收　上 去 了。 'She closed the screen.'	(Distractor in the main study)
5	tā bǎshū cóngshūjiàshàng náxiàlái yòu fànghuíqùle 她 把 书　从　书架　上　拿下来，又　放 回 去了。 'She took a book from the bookshelf, and put it back.'	(BC2)
6	tā bǎwēibōlú dǎkāi bǎ fàn fàng dàolewēibōlúlǐ 她 把 微 波 炉 打 开，把　饭　放　到 了 微 波 炉里。 She opened the microwave, put her lunch into it and closed the door.	(BC1)
7	tābǎwēibōlúdǎkāi bǎfànnáchūlái fàngdàolezhuōzǐshàng 她把 微 波 炉打开，把饭拿 出 来，放　到 了 桌 子 上　。 'She opened the door of the microwave, took the heated meal out, and put it on the table.'	(BC2)
8	tā bǎ bàozhǐ rēng dàolelājīxiāng lǐ 她 把　报 纸 扔　到 了垃圾 箱　里。 'She threw a newspaper into a rubbish bin.'	(BC1)
9	tā bǎshū cóngzhuōzǐshàng náqǐlái kànlekàn yòufànghuíqùle 她 把 书　从　桌 子 上　拿起来，看 了看，又 放 回 去了。 'She picked up a book from the desk, had a quick look, and then put it back.'	(Distractor in the main study)
10	tā bǎyīzhāngkǎpiàn tiēdàolebǎnshàng yòubǎtā náxiàláile 她 把一 张　卡 片 贴到了板 上　，又 把它 拿下来了。 'She pasted a card on the wall, and then took it off.'	(BC1& BC2)
11	tā nále nàběnshū bǎtā fàngdàoleshūbāolǐ 她 拿 了 那本 书，把它　放 到 了书 包里。 'She picked up the book, and put it into a bag.'	(BC1)

12   tā bǎbēizilǐdeshuǐ cóngyīgèbēizi dǎojìnlìngyīgèbēizi
     她 把杯子里的水　从　一个杯子　倒 进 另一个杯子。

     'She poured the hot water from one cup into the other one.'   (BC1)

13   tā bǎyǐzi bāndào biédefángjiānle   (BC1)
     她 把椅子　搬　到　别的 房　间 了。

     'She moved the chair to another room.'

14   tā yòubǎyǐzi bānhuílaíle   (Distractor in
     她 又 把椅子　搬 回 来了。   the main

     'She moved that chair back.'   study)

15   tā bǎshǒujī fàngjìnbāolǐ　ránhòu　yòubǎtānáchūlái　kànlekàn   (Distractor in
     她 把 手 机　放 进 包 里，然 后，　又 把 它拿 出 来，看 了 看。   the main

     'She put her mobile phone into her bag, and then took it out.'   study)

16   tā bǎbǐcónghézilǐ náchūlái　zàizhǐshàng xiěleyíxià　yòu bǎtāfàng   (BC2)
     她 把笔 从　盒子里 拿 出 来，在 纸　上　写 了一下，又　把它 放

     huíhézilǐle
     回 盒子里了。

     'She took the pen from the box, wrote something on the paper, and then put it

     back.'

Appendix 11: The instruction, training examples and stimuli in the OIT

Instruction：

This oral imitation test contains five training examples and twenty statements. Listen to each statement, and then imitate the final clause in the statement following a beep.

Training examples:

1. 吃饭以前，你要洗手先。 (Ungrammatical)

   吃饭以前，<u>你要先洗手</u>。

   Before eating, you'd better wash your hands.

2. 我不能去你家，因为，我没有车。(Grammatical)

   我不能去你家，因为，<u>我没有车</u>。

   I cannot go to your home, because I haven't got a car.

3. 虽然我不会中文，可是，我有很多中国朋友。　　(Grammatical)

   虽然我不会中文，可是，<u>我有很多中国朋友</u>。

   Although I cannot speak Chinese, I have many Chinese friends.

4. 我的女朋友是中国人，她很漂亮和真聪明。(Ungrammatical)

   我的女朋友是中国人，<u>她很漂亮也很聪明。</u>

   My girl friend is Chinese. She is tall and smart (or intelligent?).

5. 我喜欢看小说，也电视。

   我喜欢看小说<u>和电视</u>。 或者<u>也喜欢看电视</u>。(Ungrammatical)

   I like reading novels and watching TV.

The stimuli:

No.	Stimulus	Type
1	nǐyàoshishànglóu wǒdeshūbāonáshàngqùba *你 要 是 上 楼，<u>我的书包拿 上 去吧.</u> 'If you go upstairs, please take my bag up.'	(UGBC2)
2	tāméiyǒuqián kěshì háiyàobǎháizisòngdàoguówàiqù 她 没 有 钱，可是，<u>还 要 把 孩 子 送 到 国 外 去</u>. 'Although she is not rich, she still wants to support her kids to study overseas.'	(GBC1)
3	nǐbúzhīdàoma wǒbǎzhèjiànshìzhīdàozàijiāle *你不 知 道 吗？<u>我把 这 件 事 知 道 在 家 了</u>。 'Don't you know? I knew this matter when I was home.'	(OUBC1)
4	jīntiānwánglǎoshībìngle tāmíngtiān cái gěiwǒmenshàngkè 今 天 王 老师 病 了，<u>她 明 天 才 给 我 们 上 课</u>。 'Ms. Wang is sick today. She will give us a lecture tomorrow.'	(GSVO)
5	péngyǒusònggěiwǒyīběnshū kěshì wǒwàngtāzàigōnggòngqìchēshàng *朋 友 送 给我一本书，可是，<u>我 忘 它 在 公 共 汽 车 上</u> <u>了</u>。 'A friend of mine gave me a book, but I left it on the bus.'	(UGSVO → BC1)
6	wàimiànhěnlěng kuàibǎmáoyīchuānshàng 外 面 很 冷，<u>快 把 毛衣 穿 上 </u>。 'It is cold outside. Please put your sweater on.'	(GBC2)
7	wǒsònggěimāmāyìxiēhuā tāxiǎngfàng huāzài tāfángjiānlǐ *我 送 给妈妈一些花，<u>她 想 放 花 在 她 房 间 里</u>。 'I bought my mum some flowers. She wanted to put them in her room.'	(UGSVO → BC1)
8	wǒmenyàoshàngkèle nǐbǎtóngxuémenjiàojìnláihǎoma 我 们 要 上 课了，<u>你把 同 学 们 叫 进 来 好 吗</u>？ 'It is time for class. Can you call your classmates in?'	(GBC2)
9	zuótiānnǐbùzàijiā wǒfàngnǐdexìnzàinǐjiāménkǒule *昨 天 你不在 家，<u>我 放 你的信 在你家 门 口了</u>。 'You were not in yesterday. I put your letter at your door.'	(UGSVO→ BC1)
10	zhèxiēdōngxīwǒbúyàole nǐrēngtāmenchūqùba *这 些 东 西我不要了，<u>你 扔 它们 出 去吧</u>。 'This is rubbish. Please throw it out.'	(UGSVO→ BC2)
11	nàgecàihuàile nǐbǎdǎotādàocèsuǒlǐba *那个 菜 坏 了，<u>你把 倒 它到 厕 所 里吧</u>。 'That dish has gone off. Please pour it into the toilet.'	(UGBC1)

12   zàitúshūguǎnkànwánshū   yàofàngshūhuíqù   (UGSVO→
*在 图 书 馆 看 完 书，<u>要 放 书 回 去</u>。   BC2)

'When you finish reading (a book) at the library, you'd better put it

back.'

13   tāméináhǎo   diàotādeshūdàoshuǐlǐle   (UGSVO→
*他 没 拿 好，<u>掉 他 的 书 到 水 里 了</u>。   BC1)

'He dropped his book into the water because he was not holding it

carefully enough.'

14   zhègezhuōzǐtàidà   nǐnéngnátājìnláima   (UGSVO→
*这 个 桌 子 太 大，<u>你 能 拿 它 进 来 吗？</u>   BC2)

'This table is too big. Can you move it in?'

15   zhèshìyàobúshìtáng   nǐzěnmedōuchīxiàqùle   (GSVO)
这 是 药 不 是 糖，<u>你 怎 么 都 吃 下 去</u>了？

'This is medicine not lollies, how could you take all of them?'

16   xiàyǔle   nǐbǎnáyīfuhuíláiba   (UGBC2)
*下 雨 了，<u>你把拿衣服回 来 吧</u>!

'It is raining. Please bring the clothes in quickly.'

17   tādānxīnyǒuréntōutādeqián   jiùfàngqiándàotādexiélǐle   (UGSVO→
*她 担 心 有 人 偷 她 的 钱，<u>就 放 钱 到 她 的 鞋 里 了。</u>   BC1)

'She was afraid that someone would steal her money,so she hid it in her

shoes.'

18   yīxiàkè   tóngxuéjiùbǎjiàoshìpǎolechūqù   (OUBC2)
*一 下课，<u>同 学 就把教 室 跑 了 出 去</u>。

'Once the class was over, the students ran out of the classroom.'

19   zuótiānwǒbìngle   shìlǎoshīsòngwǒhuíjiāde   (GSVO)
昨 天 我 病 了，<u>是 老 师 送 我 回 家 的</u>。

'I was sick yesterday. It was my teacher who drove me back home.'

20   wǒxǐhuanchīzhōngguócài   kěshì   búhuìzuòzhōngguócài   (GSVO)
我 喜 欢 吃 中 国 菜，可 是，<u>不 会 做 中 国 菜</u>。

'I like Chinese food, but I don't know how to cook.'

Note. G= grammatical, UG=ungrammatical, OU=overused, *= ungrammatical sentence

Appendix 12: The instruction, training examples and stimuli in the UGJT

The instruction in the UGJT:

This UGJT test contains four training examples and thirty items. Following the training examples, please read out aloud each item, and then judge whether the underlined part in each sentence is grammatical or ungrammatical. If you think it is ungrammatical, please correct it orally. Please also indicate whether you have judged the sentences by feel or rule.

这项语法判断测试一个有 4 道例题和 30 道题目。请先读一遍每一个句子，然后说说句子中划线的部分对不对，如果不对，请改正，并说出你判断和改正的理由，是根据感觉还是语法规则？

Training examples： 例如：

1. Suīrán zhǐyǒu bā suì, kěshì, tā xué Zhōngwén yǐjīng liǎngnián le.

   虽然只有 8 岁，可是，<u>他学中文已经两年了。</u>

   <u>Grammatical.</u>   I judged this sentence

   <u>(a) by feel</u> (b) by rule (The rule is _____ .) (c) not sure

2. Yīnwèi wǒ yào qù Zhōngguó, suǒyǐ, wǒ xiang xué Zhōngwén.

   因为我要去中国，<u>**所以，我想学中文**</u>。

   <u>Grammatical.</u>   I judged this sentence

   <u>(a) by feel</u>   (b) by rule (The rule is _____ .) (c) not sure

3. Tāmen shì wǒ hǎo de péngyǒu..

   他们是我<u>好的朋友</u>.

   Ungrammatical. 他们是我的好朋友。 I judged this sentence

   (a) by feel   <u>**(b) by rule (The rule is   .)**</u> (c) not sure

4. Wǒ yǒu èr ge dìdi.

   我有<u>二个弟弟</u>.

   Ungrammatical. 我有两个弟弟。 I judged this sentence

   (a) by feel   <u>**(b) by rule (The rule is   .)**</u> (c) not sure

The stimuli in the UGJT:

kǎoshì yǐqián wǒ zǒngshì hěn máng
1. 考试 以 前，<u>我 总 是 很 忙</u>。 (Distractor)

'I am always busy before the exams.'

wǒ xiǎng kàn tā de shū kěshì tā bǎ shū ná dào xuéxiào qù le
2. 我 想 看 她 的 书，可是，<u>她 把 书 拿 到 学 校 去 了</u>。(GBC1)

'I want (to) read her book, but she took it to school.'

zhè běn zìdiǎn tài dà le nǐ bǎ yìběn xiǎo de dài shàng ba
3.* 这 本 字 典 太 大 了，<u>你 把 一 本 小 的 带 上 吧</u>（UGBC2）

'This dictionary is too big. You'd better take a small one with you.'

tīngshuō zhōngguócài hěn hǎochī kěshì wǒ méi chī le guò
4. [36]*听 说 中 国菜 很 好吃，可是，<u>我 没 吃 了/ 过</u>。(Distractor)

'I have heard that Chinese food is delicious, but I have never tasted.'

wǒ yǒu hěnduō xīnzhàopiàn wǒ xiǎng bǎ tāmen dào diànnǎo shàng
5.*我 有 很 多 新 照 片，<u>我 想 把 它 们 到 电 脑 上</u>。 (UGBC1)

'I have many new pictures, so I want to upload them to my computer'.

tā bǎ yíjiàn yīfu ná qǐlái kànlekàn yòufàngxiàle
6. <u>他 把 一 件 衣 服 拿 起来</u>，看 了 看，又 放 下 了。（GBC2）

'He picked up a jacket, had a look at it, and then put it down.'

zuótiān kǎoshì de shíhòu wǒ wǎn lái le
7.* 昨 天 考 试 的 时 候，<u>我 晚 来 了</u>。(Distractor)

I was late at yesterday's test.

zhè běn shū zhēn hǎo tā wǒ yào dài huí guó
8.*这 本 书 真 好，<u>它 我 要 带 回 国</u>。 (UGSVO→ BC1)

This book is really good. I am going to take it back to my home country with me.

tā bù qǐchuáng nǐ néng bǎ jiào qǐlái tāma
9.*他 不 起 床，<u>你 能 把 叫 起来 他 吗</u>? (UGBC2)

He would not get up. Can you wake him up?

lái běijīng yǐhòu wǒ jiàn le yìxiē zhōngguó péngyǒu
10. 来 北 京 以后，<u>我 见 了一些 中 国 朋 友</u>。(Distractor)

'I have met a few Chinese friends since I came to Beijing.'

jīntiān wǒ bù huí jiā wǒ bǎ fàn chī zài xuéxiào
11.*今天 我 不 回 家，<u>我 把 饭 吃 在 学 校</u>。 (OUBC1)

'I am not going home today. I will have my lunch at school.'

---

[36] This item was changed to grammatical one by changing 了(le) into 过(guò) in the main study.

wǒ sòng gěi tā yīběnshū kěshì tā yòu bǎ nà běn shū sòng huílái le
12. 我 送 给 她 一本 书，可是，她 又 把 那 本 书 送 回来了。(GBC2)

'I sent her a book, but she sent it back to me.'

suīrán wǒ xiànzài xué zhōngwén kěshì wǒbù méiyǒu yīběnzhōngwén shū
13.[37] *虽然 我 现在 学 中 文，可是，我不/ 没 有 一 本 中 文 书。(Distractor)

'Although I am studying Chinese now, I don't have a Chinese book.'

wǒ qù jiàn tāmāmā kěshì tā yǐjīng bǎ sòng māmā dào fēijīchǎng le
14. *我 去 见 他妈妈，可是，他 已经 把 送 妈妈 到 飞机 场 了。(UGBC1)

'I went to see his mum, but he had dropped his mum off at the airport.'

wǒ kànjiàn nǐ tōu le wǒde qián bǎ ná qián chūlái
15. *我 看见 你 偷 了 我的 钱，把 拿 钱 出来。（UGBC2）

'I saw you stealing money from my wallet. Hand it back to me.'

wǒ jīntiān tài máng míngtiān qù nǐ jiā
16. 我 今天 太 忙，明 天 去 你家。(Distractor)

'I am very busy today. I will visit you tomorrow.'

tā nà běn shū fàng zài zhuōzǐ shàng jiù zǒu le
17. *他 那 本 书 放 在 桌子 上，就 走 了。(UGSVO→ BC1）

'He put that book on the desk and then left.'

zhè bǎ yǐzi huài le nǐ ná tā chūqù ba
18. *这 把 椅子 坏 了，你 拿 它 出去 吧。（UGSVO→ BC2)

'This chair is broken. Please take it out.'

wǒ gānglái zhèlǐ méiyǒuduōpéngyǒu
19. *我 刚来 这里，没有 多朋 友。(Distractor)

'I just arrived, so I have few friends.'

wǒ bìngle nǐ néng bǎ wǒ dào yīyuàn qù ma
20. *我 病了，你 能 把 我 到 医 院 去 吗？(UGBC1)

'I am sick. Could you please take me to the hospital?'

chúle wǒ yǐwài dàjiā dōu bǎ shān pá shàngqù le
21. *除了 我 以外，大家 都 把 山 爬 上 去 了。(OUBC2)

'Everyone but I climbed up the hill.'

wǒ shì hǎo xuéshēng wǒ cónglái méiyǒu bú qù shàngkè
22. 我 是 好 学 生，我 从来 没有 不 去 上课。(Distractor)

'I am a good student, and I have never been absent from class'

tài wǎn le nǐ bǎ nǚpéngyǒu sòng zàijiā ba
23. 太 晚 了,你 把 女 朋 友 送 在家 吧。(UGBC1)

---

bù méi
[37] This item was changed to a grammatical one in the main study by changing 不 into 没.

'It is getting late. You'd better walk your girlfriend home.'

tā tài shēngqìle   diànshì cóng lóushàng rēng xiàqù le
24. 他 太 生 气了， <u>电 视 从 楼 上 扔 下去 了</u>。(UGSVO→ BC2)

'He was so angry that he threw the TV down the stairs.'

wǒ xiǎng qù zhōngguó kěshì wǒ méiyǒu qián yě méiyǒu shíjiān
25. 我 想 去 中 国，可是， <u>我 没 有 钱 也 没 有 时 间</u>。 (Distractor)

'I want to go to China, but I have neither the money nor the time.'

jīntiānchīfàn deshíhòu wǒ bǎ yī gèwǎn diào dào dìshàng le
26. 今 天 吃 饭 的时 候， <u>我 把 一 个 碗 掉 到 地 上 了</u>。（GBC1）

'I dropped a bowl on the floor at lunch time'.

bān jiā de shíhòu tā yígèrén jiù bǎ zhuōzǐ qǐláile
27. 搬 家 的 时 候，<u>他 一个人 就 把 桌 子 起来 了</u>。(UGBC2)

'When we moved house, he moved the table himself'.

zài xīnxīlán   kěyǐ chīdào zhōngguócài
28. 在 新 西 兰， <u>可以 吃 到 中 国菜</u>。(Distractor)

'In New Zealand, there are places where you can eat Chinese food.'

zhèběnshūtàidàle hěnnánbǎfàngjìnzhèmexiǎodebā
29.* 这 本 书 太 大 了，<u>很 难 把 放 进 这 么 小 的包里</u>。(UGBC1)

'This book is too big to fit into such a small bag.'

nǐ xuéguònàge hànzìle nǐ néngbǎtāxiě chū ma
30.*你 学 过 那个 汉 字 了，<u>你 能 把它 写 出 吗？</u> （UGBC2）

'You have learned that Chinese characters. Can you write it ?'

Appendix 13: Follow-up interview questions

1. 你 觉 得 哪 项 任 务 最 难 ? 为 什 么 ?

   Which task do you think was the most difficult for you? Why?

2. 你 觉 得 哪 项 任 务 最 容 易 ? 为 什 么 ?

   Which task do you think was the easiest for you? Why?

3. 你最喜欢做哪项任务？为什么？

   Which task was your favourite? Why?

4. 你觉得自己的哪项任务做得最好？为什么？

   Which task do you think you did best? Why?

5. 你知道这些任务要考你什么吗？

   Do you know what target feature these tasks tested?

Appendix 14: Background information sheet for L2 learners (English and Chinese versions)

Background questionnaire for non-native speakers of Chinese

Name:

Please circle the most suitable answer for each of the following questions or fill an answer in each blank.

1. What is your age and gender?   2. What is your ethnicity?

Age: _____   _____

A. Male   B. Female

3. What is your native language?

A. English   B. Other language _____

4. Rate your ability in using Chinese by putting ticks into the table below for each skill.

Skills	Excellent	Very Good	Good	Average	Poor
Speaking					
Listening					
Reading					
Writing					

4. What other languages can you speak?

_____

5. What is your purpose for learning Chinese? Circle your choice(s).

A. Towards obtaining a degree.
B. To enhance my employment opportunities
C. Personal interest
D. To visit China
E. Other   _____

6. Have you visited China? If 'yes' for how long?
Yes/ No

How long? _____

7. Where have you studied Chinese previously and for how long?

Place   Estimate of total number of hours spent studying

_____   _____

_____   _____

8. How old were you when you started your Chinese lesson?

_____

244

What Chinese courses are you currently enrolled in?

_____

9. How many hours do you spend on learning Chinese outside class every week?

A. Less than 1 hour   B.1-2hours (including 2 hours)

C. 2-3hours (including 3hours)   D. More than 3 hours_____

10. Do you speak Chinese with native speakers outside class? If 'yes' how many hours each week?

Yes/ No   Number of hours _____

12. Which methods do you think are most effective for learning Chinese class?

Tick or circle your choice(s).

A. Watching video programs

B. Reading

C. Listening to audio-tapes

D. Listening to instructors' explanation of grammar features

E. Doing written exercises by yourself (e.g. doing translation or doing written exercises in the textbook)

F. Doing oral exercises by yourself (e.g. reading after the instructor or audio tape)

G. Doing written tasks in pairs (e.g. checking language partners' answers)

H. Doing oral tasks in pairs (e.g. completing a task by using new items)

I. Doing oral tasks in groups (e.g. discussing a topic in a group)

J. Doing written tasks in groups (e.g. dictation)

13. Have you ever taken a standard Chinese proficiency examination? If so, please specify the name of the exam and your scores.

 Yes   **No**

The name of the exam: _____

Proficiency level: _____

Total score: _____

Listening comprehension score: _____Grammar score:_____

Reading comprehension score: _____ Comprehensive test score:_____

Composition score: _____

Oral test score: _____

14. What level of Chinese do you expect to be able to finally achieve?

A. native speaker level

B. advanced learner level

C. intermediate level

D. beginner level only.

☺Thank you very much for participating in this project! ☺

汉语非母语学习者背景调查问卷

Background questionnaire for non-native speakers of Chinese

姓名:

请挑选或填写适合您的选项。

1. 请写出您的性别和年龄

年龄: _____

A. 男　　　　B. 女

2. 您的母语是什么?

A. 英语　　B. 韩语

3. 请在适合您言语技能水平的地方打勾?

言语技能	优秀	很好	好	一般	差
说					
听					
读					
写					

4. 您还会说什么其它语言?

_____

5. 您学汉语的目的是什么? 请圈出合适的选项， 可以多选。

A. 为获得学位

B. 为增加就业机会

C. 个人兴趣

D. 去中国

E. 其他

_____

6. 您去过中国吗? 如果去过，请说明有多长时间?

去过 / 没有

多长时间? _____

7. 以前您学过中文吗? 学过多长时间?

地点　　　　　估计的所花的小时数

_____　　　_____

_____　　　_____

8. 现在您学习什么中文课程？

——————————————————

9.课后您每星期花多长时间学习中文？

A. 不到 1 小时   B.1-2 个小时（包括 2 小时）

C. 2-3 小时 (包括 3 小时)   D. 超过 3 小时_____

10. 课后您跟中国人说中文吗？估计每星期说多少小时？

是/ 否

说中文的小时数_____

12. 您认为哪种方法学中文最有效? 可多选。

A. 看电视节目

B. 阅读伴有插图的文字

C. 听录音

D. 听老师讲解语法点

E. 自己做书面练习 （如：翻译 或 做其他书面练习）

F. 自己做口头表达练习 （如：跟老师或跟录音朗读）

G. 做两人小组书面练习 （如：互相检查书写错误）

H. 做两人小组口头表达练习 （如：两人用新学的词语做交际练习）

 I. 大组书写练习 （如：听写，连句成段书写练习）

J. 大组口头讨论 （如：多人讨论一个话题， 辩论）

13. 您希望自己的汉语能 达到什么样的水平?

E. 跟中国人一样

F. 学习者高级水平

G. 学习者中级水平

H. 学习者初级水平

☺感谢您参与这项调查 ☺

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning, 59*(2), 249-306.

Adamson, H. D. (1988). *Variation theory and second language acquisition*. Washington, DC: Georgetown University Press.

Adamson, H. D., & Elliott Jr, O. P. (1997). Sources of variation in interlanguage. *Annual Review of Applied Linguistics (IRAL), 35*(2), 87-98.

Adamson, H. D., & Regan, V. M. (1991). The acquisition of community speech norms by Asian immigrants learning English as a second language. *Studies in Second Language Acquisition, 13*(1), 1-22.

Adjemian, C. (1976). On the nature of interlanguage systems. *Language Learning, 26*(2), 297-320.

Adjemian, C. (1981). La specificite de I ' interlangue et l' idealisation des langues secondes. In J. Gueron (Ed.), *Grammaire Generaive: Theorie et Methologies*. Paris.

Andersen, R. W. (1984). The one-to-one principle of interlanguage construction. *Language Learning, 34*, 77-95.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA, London: Harvard University Press.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum.

Arditty, J., & Perdue, C. (1979). Variabilite et connaissances en langue etrangere *Encrages* (pp. 32-43). Vincennes: Universite de Paris VIII.

Arthur, B. (1980). Gauging the boundaries of second language competence: A study of learner judgments. *Language learning, 30*, 177-194.

Bachman, L. F. (1988). Language testing-SLA research interfaces. *Annual Review of Applied Linguistics (IRAL), 9*(1), 193-209.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science, 4*(11), 417-423.

Bailey, K. M., Madden, C., & Krashen, S. D. (1974). Is there a "natrual sequence" in adult second language learning? *Language Learning, 24*(2), 235-243.

Bardovi-Harlig, K. (1987). Markedness and salience in second language acquisition. *Language Learning, 37*(3), 385-407.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning and use*. Oxford: Blackwell.

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*(2), 233-262.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 629-654.

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173-218). New York: Cambridge University Press.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 157-193). Hillsdale, NJ: Erlbaum.

Bayley, R. (1991). *Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking*. Unpublished doctorial disseration, Stanford University, United States -- California.

Bayley, R. (1996). Competing constraints on variation in the speech of adult Chinese learners of English. In R. Bayley & D. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 97–120). Amsterdam; Philadelphia: John Benjamins.

Bayley, R. (2005). Second language acquisition and sociolinguistic variation. *Intercultural Communication Studies, 14*(2), 1-15.

Bayley, R., & Lucas, C. (2007). *Sociolinguistic variation: Theories, methods, and applications*. Cambridge; New York: Cambridge University Press.

Bayley, R., & Preston, D. (1996). *Second language acquisition and linguistic variation*. Amsterdam; Philadelphia: John Benjamins.

Bayley, R., & Schecter, S. R. (2003). *Language socialization in bilingual and multilingual societies*. Clevedon; Buffalo: Multilingual Matters.

Beebe, L. (1980). Sociolinguistic variation and style shifting in second language acquisition. *Language Learning, 30*(2), 433-447.

Beijing Daxue, X. H. J. (1993). *Xiandai Hanyu. Mordern Chinese* (1 ed.). Beijing: The Commercial Press.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*(1), 69-83.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning, 29*(1), 81-103.

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied Linguistics, 3*(3), 181-206.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*(2), 157-168.

Bialystok, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research, 13*(2), 116-137.

Bialystok, E., & Hakuta, K. (1994). *In other words: The science and psychology of second-language acquisition*. New York: Basic Books.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, N.J.: Erlbaum.

Bialystok, E., & Miller, B. (1999). The problem of age in second-language acquisition: Influences from language, structure, and task. *Bilingualism: Language and Cognition, 2*(02), 127-145.

Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. In D. L. Rorrest-Pressley, G. E. Mackinnon & T. G. Waller (Eds.), *Metacognition, cognition, and human performance: Vol.1. Theoretical perspectives* (pp. 207-252). San Diego, CA: Academic Press.

Birdsong, D. (1989). Metalinguistic performance and interlanguage competence. *Berlin ua: Springer*.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language, 68*(4), 706-755.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56*(s1), 9-49.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language, 44*(2), 235-249.

Bley-Vroman, R., & Chaudron, C. (1990). Second language processing of subordinate clauses and anaphora—first language and universal influences: A review of Flynn's research. *Language Learning, 40*(2), 245-285.

Bley-Vroman, R., & Chaudron, C. (1994). Elicited imitation as a measure of second-language competence. In S. M. Gass, E. Tarone & A. D. Cohen (Eds.), *Research methodology*

in second-language acquisition (pp. 245–261). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bornstein, M. (1989). Sensitive periods in development: Structural characteristics and causal interpretations. *Psychological Bulletin, 105*(2), 179-197.

Boyle, J. P. (1987). Sex differences in listening vocabulary. *Language Learning, 37*(2), 273-284.

Brecht, R. D., Davidson, D. E., & Ginsberg, R. B. (1991). On evaluating language proficiency gain in study abroad environments: An empirical study of American students of Russian. *Washington, DC: US Department of Education, American Council of Teachers of Russian, and the National Foreign Language Center.(ERIC Document Reproduction Service No. ED 350 855).*

Brecht, R. D., Davidson, D. E., & Ginsberg, R. B. (1995). Predictors of foreign language gain during study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 37-66). Amsterdam: John Benjamins.

Bresnan, J. (2001). *Lexical-Functional Syntax*. Oxford: Blackwell.

Bresnan, J. (Ed.). (1982). *The mental representation of grammatical relations*. Cambridge: Mass: MIT Press.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al.(1989). *Journal of Experimental Psychology: General, 120*(3), 316-323.

Brooks, N. (1960). *Language and language learning*. New York: Harcourt Brace and World.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Bruhn de Garavito, J. (1999). *The syntax of Spanish multifunctional clitics and near-native competence.* Unpublished Ph. D. thesis, McGill University.

Burstall, C. (1975). Factors affecting foreign-language learning: A consideration of some relevant research findings. *Language Teaching and Linguistic Abstracts, 8*, 105-125.

Burt, M., Dulay, H., & Hernández- Chávez, E. (1975). *Bilingual Syntax Measure I*. New York: Harcourt, Brace and Jovanovich.

Bygate, M. (1999). Task as context for the framing, reframing and unframing of language. *System, 27*(1), 33-48.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 23-48). Harlow: Pearson Education Limited.

Bygate, M., Skehan, P., & Swain, M. (2001). *Researching pedagogic tasks: Second language learning, teaching, and testing*. Harlow: Pearson Education.

Carroll, J. B. (1967). Foreign language proficiency levels attained by language majors near graduation from college. *Foreign Language Annals, 1*(2), 131-151.

Celaya, M. L., Torras, M. R., & Pérez Vidal, C. (2001). *Short and mid-term effects of an earlier start: An analysis of EFL written production.* Amsterdam: John Benjamins.

Cenoz, J., Hufeisen, B., & Jessner, U. (2001). *Looking beyond second language acquisition: Studies in tri- and multilingualism.* Tübingen: Stauffenburg.

Chafe, W. L. (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

Chao, Y. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Charters, A. H. (2005). *The second language acquisition of Mandarin nominal syntax.* Unpublished doctoral thesis, The University of Auckland, Auckland, New Zealand.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning, 33*(3), 343-377.

Chaudron, C. (2003). Data collection in SLA research. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 762-828). Malden, Oxford, Carlton: Wiley-Blackwell.

Chen, L. (2005). *A pedagogical grammar of the Chinese disposal (Ba) construction.* Unpublished Master's thesis, National Taiwan Normal University.

Chen, P. (2003). Indefinite determiner introducing definite referent: a special use of ' yi 'one'+ classifier'in Chinese. *Lingua, 113*(12), 1169-1184.

Cheng, L. (2006). Riben liuxuesheng baziju xide qingkuang kaocha yu tanxi. An analysis of the acquisition of the "ba" sentence by Japanese students. *Yunnan shifan daxue xuebao. Journal of Yunnan Normal University: Teaching and Research of Chinese as a Foreign Language*(3), 41-46.

Cheung, H. (1992). *The acquisition of ba in Mandarin. Unpublished doctoral thesis.* Unpublished doctoral dissertation, University of Kansas.

Christie, K., & Lantolf, J. P. (1992). The ontological status of learner grammaticality judgments to UG approaches to L2 acquisition. *Rassegna Italiana di Linguistica Applicata, 25*(3), 31-57.

Churchill, E. (2006). Variability in the study abroad classroom and learner competence. In M. A. DuFon (Ed.), *Language learners in study abroad contexts* (pp. 203–227). Buffalo: Multilingual Matters.

Coakes, S. J., Steed, L. G., & Price, J. C. (2008). *SPSS : Analysis without anguish : Version 15.0 for Windows*. Milton, Qld.: John Wiley & Sons Australia.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coleman, J. A. (1998). Language learning and study abroad: The European perspective. *Frontiers, 4*, 167-203.

Collentine, J. (2004). The effects of learning contexts on morphosyntactic and lexical development. *Studies in Second Language Acquisition, 26*(2), 227-248.

Collentine, J. (2009). Study abroad research: Findings, implications, and future directions. In M. H. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 218-233). Chichester, UK; Malden, MA: Wiley-Blackwell.

Collentine, J., & Freed, B. F. (2004). Learning context and its effects on second language acquisition: Introduction. *Studies in Second Language Acquisition, 26*(2), 153-171.

Colombo, J. (1982). The critical period concept: Research, methodology, and theoretical issues. *Psychological Bulletin, 91*(2), 260-275.

Cook, V. J. (1991). The poverty of the stimulus argument and multi-competence. *Second Language Research, 7*(2), 103-117.

Cook, V. J. (2003). *Effects of the second language on the first*. Clevedon; Buffalo: Multilingual Matters.

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161-169.

Corder, S. P. (1973). The elicitation of interlanguage. In J. Svartik (Ed.), *Errata: Papers in Error Analysis*. Lund: EWK Gleerup.

Corder, S. P. (1976). *The study of interlanguage.* Paper presented at the The Fourth International Conference of Applied Linguistics, Munich, Hochschiulverlag.

Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.

Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.

Cowan, R., & Hatasa, Y. A. (1994). Investigating the validity and reliability of native speaker and second-language learner judgments about sentences. In S. M. Gass, E. Tarone & A. D. Cohen (Eds.), *Research Methodology in Second Language Acquisition* (pp. 287-302). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition, 11*(4), 367-383.

Crookes, G., & Gass, S. M. (1993). *Tasks and language learning: Integrating theory and practice.* Clevedon: Multilingual Matters.

Cui, X. (1995). "Ba" ziju de ruogan jufa yuyi wenti. Some Syntactic and Semantic Puzzles Concerning the Ba-Construction. *Chinese Teaching in the World, 33*(3), 12-21.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*(2), 175-187.

Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada: A reassessment. *Applied Linguistics, 2*(2).

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children*. Cambridge: Cambridge University Press.

Curtain, H. (1997). Early start language programs, Unpublished paper. Madison, WI: Author.

Curtiss, S. (1977). *Genie: A linguistic study of a modern day "wild child"*. New York: Academic Press.

Cutler, A., & Clifton, C. E. (1999). Comprehending spoken language: A blueprint of the listener. In C. Brown & P. Hagoort (Eds.), *Neurocognition of Language* (pp. 123-166). Oxford: Oxford University Press.

Davies, W. D., & Kaplan, T. I. (1998). Native speaker vs. L2 learner grammaticality judgements. *Applied Linguistics, 19*(2), 183.

de Graaff, R. (1997). The experanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*(2), 249-276.

de Villiers, J. G. (1980). The process of rule learning in child speech: A new look. In K. E. Nelson (Ed.), *Children's language* (Vol. 2, pp. 1-44). New York: Gardner.

DeKeyser, R. M. (1986). *From learning to acquisition? Foreign language development in a U.S. classroom and during a semester abroad.* Unpublished doctoral dissertation, Stanford University, United States - CA.

DeKeyser, R. M. (1990). From learning to acquisition? Monitoring in the classroom and abroad. *Hispania, 73*(1), 238-247.

DeKeyser, R. M. (1991). The semester overseas: What difference does it make? *ADFL Bulletin, 22*(2), 42-48.

DeKeyser, R. M. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition* (pp. 42-63). New York: Cambridge University Press.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499-533.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Malden, Oxford, Carlton: Wiley-Blackwell.

DeKeyser, R. M. (2005). *Grammatical development in language learning*. Malden, MA: Blackwell Pub.

DeKeyser, R. M., & Juffs, A. (2005). Cognitive considerations in L2 learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 437-454). Mahwah, NJ: Lawrence Erlbaum Associates.

DeKeyser, R. M., Alfi-Shabtay, I., Ravid, D., & Shi, M. (2005). The role of salience in the acquisition of Hebrew as a second language. Manuscript in preparation.

DeKeyser, R. M., Ravid, D., & Alfi-Shabtay, I. (2005). Cross-linguistic evidence for the nature of age effects in second language acquisition. Manuscript in preparation.

Dewaele, J.-M. (1998). The effect of gender on the choice of speech style. *ITL Review of Applied Linguistics*, 9-25.

Dewaele, J.-M. (2004a). Retention or omission of the *ne* in advanced French interlanguage: The variable effect of extralinguistic factors. *Journal of Sociolinguistics, 8*(3), 433-450.

Dewaele, J.-M. (2004b). Vous or tu? Native and non-native speakers of French on a sociolinguistic tightrope. *International Review of Applied Linguistics in Language Teaching (IRAL), 42*(4), 383-402.

Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition, 4*(03), 275-297.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition, 26*(2), 249-274.

Díaz-Campos, M. (2006). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In C. A. Klee & T. L. Face (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 26-39). Somerville, MA: Cascadilla Proceedings Project.

Dickerson, L. (1974). *Internal and external patterning of phonological variability in the speech of Japanese learners of English.* Unpublished doctoral dissertation, University of Illinois, Urbana.

Dickerson, L. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly, 9*(4), 401-407.

Donovan, L. A., & MacIntyre, P. D. (2004). Age and sex differences in willingness to communicate, communication apprehension, and self-perceived competence. *Communication Research Reports, 21*(4), 420-427.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*(4), 431-469.

Douglas, D. (1986). Communicative competence and the tests of oral skills. In C. Stansfield (Ed.), *TEFOL research reports, 21. Toward communicative competence testing: Proceedings of the Second TEFOL Invitational Conference Princeton* (pp. 157-175). NJ: Educational Testing Service.

Du, H. (2004). *The acquisition of the Chinese ba-construction by adult second language learners.* Unpublished doctoral dissertation, The University of Arizona, United States -- Arizona.

DuFon, M. A., & Churchill, E. (Eds.). (2006). *Language learners in study abroad contexts.* Buffalo: Multilingual Matters.

Dulay, H., & Burt, M. (1973). Should we teach children syntax? *Language Learning, 23*(2), 245-258.

Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*(1), 37-53.

Dulay, H., & Burt, M. (1975). Creative constrution in second language learning and teaching. In M. Burt & H. Dulay (Eds.), *TESOL '75* (pp. 21-32). Washington DC: TESOL.

Dyson, P. (1988). *The effect on linguistic competence of the year spent abroad by students studying French, German and Spanish at degree level.* Oxford: Oxford University Language Teaching Centre.

Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning, 32*(2), 367-391.

Ekstrand, L. (1976). Age and length of residence as variables related to the adjustment of migrant children, with special reference to second language learning. In G. Nickel

(Ed.), *Proceedings of the Fourth International Congress of Applied Linguistics* (Vol. 3, pp. 179-197). Stuttgart: Hochschulverlag.

Ellis, N. (1996a). Analyzing language sequence in the sequence of language acquisition: Some comments on Major and Ioup. *Studies in Second Language Acquisition, 18*(3), 361-368.

Ellis, N. (1996b). Sequencing in SLA. *Studies in Second Language Acquisition, 18*(1), 91-126.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143-188.

Ellis, N. (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). Malden, Oxford, Carlton: Blackwell Publishing.

Ellis, N. (2009). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 139-158). Malden, Oxford, Chichester: Blackwell Publishing.

Ellis, N., & Ferreira–Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal, 93*(3), 370-385.

Ellis, N., & Laporte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. *Tutorials in bilingualism: Psycholinguistic perspectives*, 53–83.

Ellis, N., & Larsen-Freeman, D. (2006). Language emergence: implications for Applied Linguistics - Introduction to the special issue. *Applied Linguistics, 27*(4), 558.

Ellis, R. (1984). *Classroom second language development: A study of classroom interaction and language acquisition*. Oxford; New York: Pergamon Press Ltd.

Ellis, R. (1985). Sources of variability in interlanguage. *Applied linguistics, 6* (2), 118-131

Ellis, R. (1987a). Contextual variability in second language acquisition and the relevancy of language teaching. In R. Ellis (Ed.), *Second Language Acquisition in Context* (pp. 179-194). Cambridge: Prentice-Hall International Ltd.

Ellis, R. (1987b). Interlanguage variability in narrative discourse. *Studies in Second Language Acquisition, 9*, 1-19.

Ellis, R. (1987c). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*(1), 1-19.

Ellis, R. (1987d). *Second language acquisition in context*. Englewood Cliffs, NJ: Prentice-Hall International.

Ellis, R. (1988). The effects of linguistic environment on the second language acquisition of grammatical rules. *Applied Linguistics, 9*(3), 257-274.

Ellis, R. (1989). Sources of intra-learner variability in language use and their relationship to second language acquisition. In S. Gass, C. Madden, D. Preston & L. Selinker (Eds.), *Variation in second language acquisition: Psycholinguistic issues* (pp. 22-45). Clevedon; Philadelphia: Multilingual Matters.

Ellis, R. (1991). Grammatically judgments and second language acquisition. *Studies in Second Language Acquisition, 13*(2), 161-186.

Ellis, R. (1992). *Second language acquisition and language pedagogy*. Clevedon; Philadelphia: Multilingual Matters.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly 27*(1, Spring), 91-113.

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

Ellis, R. (2004). The definition and measurement of explicit knowledge. *Language Learning, 54*(2), 227-275.

Ellis, R. (2005a). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*(2), 141-172.

Ellis, R. (2005b). *Planning and task performance in a second language*. Philadelphia: John Benjamins.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics, 27*(3), 431-463.

Ellis, R. (2008a). Explicit knowledge and second language learning and pedagogy. *Encyclopedia of language and education (6): Knowledge about language*, 1901-1911.

Ellis, R. (2008b). Investigating grammatical difficulty in second language learning: implications for second language acquisition research and language testing. *International Journal of Applied Linguistics, 18*(1), 4-22.

Ellis, R. (2008c). *The study of second language acquisition (2nd, edn)*. Oxford: Oxford University Press.

Ellis, R. (2009a). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-26). Bristol; Buffalo: Multilingual Matters.

Ellis, R. (2009b). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Bristol; Buffalo: Multilingual Matters.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language* (1st ed.). Oxford: Oxford University Press

Ellis, R., & Rathbone, M. (1987). *The acquisition of German in a classroom context*. Mimeograph, London: Ealing College of Higher Education, London.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*(1), 59-84.

Ellis, R., Loewen, S., Elder, C., Erlam, R., Philp, J., & Reinders, H. (Eds.). (2009). *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol, UK; Buffalo [N.Y.]: Multilingual Matters.

Erbaugh, M. (1976). The Chinese Pear Stories. A collection of 21 narratives. Unpublished transcription.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowlege: An empirical validation study. *Applied Linguistics, 27*(3), 464-491.

Færch, C., & Kasper, G. (1981). *Procedural knowledge as a component of foreign language learners' communicative competence*. Paper presented at the Kommunikation im (Sprach-) Unterricht.

Færch, C., & Kasper, G. (1984). Pragmatic knowledge: Rules and procedures. *Applied Linguistics, 5*(3), 214.

Færch, C., & Kasper, G. (1986). One learner-two languages: Investigating types of interlanguage knowledge *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies* (pp. 211–227): Gunter Verlag Tübingen.

Færch, C., Haastrup, K., & Philipson, R. (1984). *Learner language and language learning*. Conpenhagen, Gyldendal and Clevedon: Multilingual Matters.

Fathman, A. (1975). The relationship between age and second language productive ability. *Language Learning, 28*(2), 245-253.

Felix, S. W. (1981). The effect of formal instruction on second language acquisition. *Language Learning, 31*(1), 87-112.

Feng, S. (1999). *Prosodic syntax in Chinese*. Shanghai: Shanghai Educational Publishing House.

Feng, S. (2001). Prosodically constrained bare-verb in ba constructions. *Journal of Chinese Linguistics, 29*(2), 243-280.

Feng, S. (2002). *Prosodic syntax and morphology in Chinese*. München: Lincom Europa.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*(5), 3125-3134.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104.

Flynn, S. (1984). A universal in L2 acquisition based on a PBD typology. In F. R. Eckman, L. H. Bell & D. Nelson (Eds.), *Universals of second language acquisition* (pp. 75-87). Rowley, Massachusetts: Newbury House.

Flynn, S. (1986). Production vs. comprehension: Differences in underlying competences. *Studies in Second Language Acquisition, 8*(2), 135-164.

Flynn, S. (1987a). Contrast and construction in a parameter-setting model of second language acquisition. *Language Learning, 36*(1), 1-27.

Flynn, S. (1987b). *A parameter-setting model of L2 acquisition: Experimental studies in anaphora*. Dordrecht, The Netherlands: D. Reidel.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323.

Freed, B. F. (1995a). Language learning and study abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 3-35). Amsterdam; Philadelphia: John Benjamins.

Freed, B. F. (1995b). *Second language acquisition in a study abroad context*. Amsterdam; Philadelphia: John Benjamins.

Freed, B. F. (1998). An overview of issues and research in language learning in a study abroad setting. *Frontiers, 4*, 31-60.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26*(2), 275-301.

Gao, X. [Xiaodan] (2005). *Noun phrase morphemes and topic development in L2 Mandarin Chinese: A processability perspective.* Unpublished doctoral thesis, Victoria University of Wellington, Wellington.

Gao, X. [Xiaoping] (1999). *Liuxuesheng baziju xide guocheng kaocha jiqi dui jiaoxue de qishi. The acquisition order of the Chinese ba constuctions by L2 learners and its pedagogical implications.* Unpublished M.A. thesis, Peking University, Beijing.

Gao, X. [Xiaoping] (2008). Xinxilan xuesheng ba ziju xide diaocha. The acquisition of the Chinese ba construction by New Zealand learners. *Duiwai hanyu yanjiu. Research on Chinese as Second Language, 4*, 156-166.

García Lecumberri, M. L., & Gallardo, F. (2003). English FL sounds in school learners of different ages. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 115–135). Clevedon; Buffalo: Multilingual Matters.

García Mayo, M. P. (1999). *Grammaticality judgements, metalinguistic awareness and the age factor in EFL. Lleida*. Lleida: Universitat de Lleida.

García Mayo, M. P. (2003). Age, length of exposure and grammaticality judgements in the acquisition of English as a foreign language. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 94 ‑ 114). Clevedon; Buffalo: Multilingual Matters.

Gass, S. M. (1994). The reliability of second language grammaticality judgments. In S. M. Gass, E. Tarone & A. D. Cohen (Eds.), *Research Methodology in Second Language Acquisition* (pp. 303-322). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Gass, S. M., & Mackey, A. (2002). Frequency effects and second language acquisition *Studies in Second Language Acquisition 24*(2), 249-260.

Gass, S. M., & Selinker, L. (1992). *Language transfer in language learning* (Rev. ed.). Amsterdam; Philadelphia: John Benjamins.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: An introductory course* (2nd ed.). Mahwah, N.J.: L. Erlbaum Associates.

Gass, S. M., & Selinker, L. (Eds.). (1983). *Language transfer in language learning*. Rowley, MA: Newbury House.

Gass, S. M., & Varonis, E. M. (1986). Sex differences in NNS/ NNS interactions. In R. R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 327-351). Rowley, MA: Newbury House.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition, 16*(3), 283-302.

Gass, S. M., Mackey, A., & Ross-Feldman, L. (2005). Task based Interactions in classroom and laboratory settings. *Language learning, 55*(4), 575-611.

Geeslin, K. L. (2003). A comparison of copula choice: Native Spanish speakers and advanced learners. *Language Learning, 53*(4), 703-764.

Geeslin, K. L., & Guijarro-Fuentes, P. (2005). The acquisition of copula choice in instructed Spanish: The role of individual characteristics. In D. Eddington (Ed.), *Selected proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 66-77). Somerville, MA: Cascadilla Proceedings Project.

Godfrey, D. (1980). A discourse analysis of tense in adult ESL monologues. In D. Larsen-Freeman (Ed.), *Discourse Analysis in Second Language Research*. Rowley, Mass: Newbury House.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the 'natural order of L2 morpheme acquisition' in English: A meta-analysis of multiple determinants. *Language Learning 51*(1), 1-50.

Golonka, E. (2006). Predictors revised: Linguistic knowledge and metalinguistic awareness in second language gain in Russian. *The Modern Language Journal, 90*(4), 496-505.

Gore, J. E. (2005). *Dominant beliefs and alternative voices: Discourse, belief, and gender in American study abroad*. New York: Routledge.

Goss, N., Zhang, Y.-H., & Lantolf, J. P. (1994). Two heads may be better than one: Mental activity in second language grammaticality judgements. In E. Tarone, S. M. Gass & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 263-286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Green, P. S., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics, 13*, 168-183.

Guijarro-Fuentes, P., & Geeslin, K. (2003). Age-related factors in copula choice in steady state L2 Spanish grammars. *Revista española de lingüística aplicada (RESLA)*(16), 83-110.

Hakuta, K. (1974). Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning, 24*(2), 287-297.

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence. *Psychological Science, 14*(1), 31.

Han, Y. (2000). Grammaticality judgment tests: How reliable and valid are they? *Applied Language Learning, 11*(1), 177-204.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge, and general language proficiency. *Language Teaching Research, 2*, 1-23.

Han, Z. (2008). On the role of meaning in focus on form. In Z. Han (Ed.), *Understanding second language process*. Clevedon, Buffulo,Toronto: Multilingual matters Ltd.

Harley, B. (1985). *Second language proficiency and classroom treatment in early French immersion*. Paper presented at the FIPLV/Eurocentres Symposium on Error in Foreign Language Learning: Analysis and Treatment.

Hashimoto, A. Y. (1971). *Mandarin syntactic structures*. Paper presented at the Unicorn 8 (Chinese Linguistics Project and Seminar).

Hedgcock, J. (1993). Well-formed vs. ill-formed strings in L2 metalingual tasks: specifying features of grammaticality judgements. *Second Language Research, 9*(1), 1.

Heidelberger Forschungsprojekt. (1978). The acquisition of German syntax by foreign migrant workers. In D. Sankoff (Ed.), *Linguistic variation: Model and methods*. New York: Academic Press.

Hisama, T. (1995). Language acquisition or language learning? *Kansai Gaidai University Journal of Inquiry and Research, 62*, 1-13.

Ho, Y. (1993). *Aspects of discourse structure in Mandarin Chinese*. Lewiston, NY: Mellen University Press.

Howard, M. (2001). The effects of study abroad on the L2 learner's structural skills: Evidence from advanced learners of French. *Eurosla yearbook, 1*(1), 123-141.

Hsueh, F.-S. (1987). Shilun baziju de yuyi tezheng. A tentative discussion of the semantic features of *Ba* sentences. *Yuyan jiaoxue yu yanjiu. Language Teaching and Studies* (1).

Hsueh, F.-S. (1989). The structural meaning of ba and bei construction in Mandarin Chinese: Do they really mean disposal and passive? In J. H.-Y. Tai & F.-S. Hsueh (Eds.), *Functionalism and Chinese Grammar*. Ohio: The Ohio state univerity.

Hsueh, F.-S. (1994). Baziju he beiziju de jiegou yiyi. The structural meanings of the ba construction and the bei construction. In J. H.-Y. Tai & F.-S. Hsueh (Eds.), *Gongneng zhuyi yu hanyu yufa. Functionalism and Chinese Grammar*. Beijing: Beijing yuyan daxue chubanshe. Beijing Language Institute Press.

Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *Studies in Second Language Acquisition, 24*(3), 347-386.

Hu, W. (2004). Ye tan ba ziju de yufa yiyi. Revisiting the grammatical meaning of the Ba-construction. *Contemporay Research in Modern Chinese 6*(October), 11-23.

Hu, W. (2005). Ye tan ba ziju de yufa yiyi. Revisiting the grammatical meaning of the Ba-construction. *Yuyan yanjiu. Stuidies in Language and Linguistics, 25*(2), 21-28.

Hu, W. (2010). Baziju yufa yiyi zai baziju jiegou ju zhong de bu junheng biaoxian. The uneven realization of the grammatical meaning of the Ba-construction among Ba-sentences. *Yuyan yanjiu. Studies in Language and Linguistics, 30*(1), 45-50.

Hu, X. [Xiaoling], & Liu, C. (2007). Restrictive relative clauses in English and Korean learners' second language Chinese. *Second Language Research, 23*(3), 263.

Huang, Y., & Yang, S. (2004). Hanyu zuowei di er yuyan de baziju xide yanjiu. The L2 Acquisition of the Chinese Ba-Construction. *Shijie Hanyu Jiaoxue. Chinese Teaching in the World*(1), 49-59.

Huebner, T. (1995). The effects of overseas language programs: report on a case study of an intensive Japanese course. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 171-193). Amsterdam; Philadelphia: John Benjamins.

Hukuta, K., & D'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13*(1), 72.

Hulstijn, J. H., & de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review, 11*, 97-112.

Hulstijn, J. H., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*(34), 23-43.

Hyltenstam, K. (1977). Implicational patterns in interlanguage syntax variation. *Language Learning, 27*(2), 383-411.

Hyltenstam, K. (1983). *Data types and second language variability*: Stockholm University.

Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. *Advances in Psychology, 83*(3), 351-368.

Hyltenstam, K., & Pienemann, M. (1985). *Modelling and assessing second language acquisition*. Clevedon: Multilingual Matters.

Ionin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning, 60*(4), 877-925.

Isabelli, C. A., & Nishida, C. (2005). Development of the Spanish subjunctive in a nine-month study-abroad setting. In D. Eddington (Ed.), *Selected proceedings of the 6th conference on the acquisition of Spanish as first and second languages* (pp. 78–91). Somerville, MA: Cascadilla Proceedings Project.

Iwakiri, M. (1993). Effects of a study abroad program on the English development of Japanese college students. *Kagoshima University Bulletin of English Education, 24*, 41-60.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York; London: Routledge.

Jin, H. G. (1989). *Language universals, language typology and L2 acquisition: A cross-sectional study of adult English speakers learning Chinese.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, United States -- Illinois.

Jin, H. G. (1992). Pragmaticization and the L2 acquisition of Chinese BA constructions. *Journal of the Chinese Language Teachers Association, 27*(3), 33-52.

Jin, H. G. (1993). Cong hanyu bazijiegou kan yuyan fenlei guilü zai di er yuyan xide guocheng zhong de zuoyong. Effects of typological transfer on the BA constructions in SLA. *Yuyan jiaoxue yu yanjiu. Language Teaching and Linguistic Studies* (2), 50-64.

Jin, L. (1997). Baziju de jufa yuyi yujing tezheng. The syntactic, semantic and contextual characteristics of the ba construction. *Zhongguo Yuwen. Chinese Language*(6), 415-423.

Jing-Schmidt, Z. (2005). *Dramatized discourse: the Mandarin Chinese Ba-construction*. Amsterdam, Netherlands: John Benjamins.

Jing-Schmidt, Z. & Tao, H. (2009). The Mandarin Disposal Constructions: Usage and Development. Language and Linguistics 10 (1) 29-58.

Johnson, J. S. (1992). Critical period effects in second language acquisition: The effect of written versus auditory materials on the assessment of grammatical competence. *Language Learning, 42*(2), 217-248.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*(1), 60-99.

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition, 39*(3), 215-258.

Judd, E. (1978). Language policy and TESOL: Socio-political factors and their influences on the profession. In C. Blatchford & J. Schachter (Eds.), *On TESOL '78: Policies, programs and practices*. Washington, D.C.: TESOL.

Kang, Y., & Lai, S. (1999). *Hanyu Huihua 301 Ju. Conversational Chinese 301*. Beijing: Beijing Language and Culture University Press.

Kaplan, R., & Bresnan, J. (1982). Lexical-Functional Grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations*. Cambridge: The MIT Press.

Katamba, F. (2004). *Morphology : critical concepts in linguistics*. London ; New York: Routledge.

Kawaguchi, S. (2005). Argument structure and syntactic development in Japanese as a second language. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory*. Melbourn: John Benjamins Publishing Co.

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working papers on Bilingualism, 15*, 59-92.

Kellerman, E. (1986). An eye for an eye: Crosslinguistic constraints on the development of the L2 lexicon. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 35-48). Oxford: Pergamon Institute of English.

Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics, 15*, 125-150.

Kellerman, E. (2001). New uses for old language: Cross-linguistic and crossgestural influence in the narratives of non-native speakers. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 170–191). Clevedon; Buffalo: Multilingual Matters.

Kellerman, E., & Sharwood Smith, M. (1986). *Crosslinguistic influence in second language acquisition*. Oxford: Pergamon Institute of English.

Kinginger, C. (2007). Research on language education through programmes of study and residence abroad: Recent publications. *International Journal of Bilingual Education and Bilingualism, 10*(1), 104-113.

Kinnear, P. R., & Gray, C. D. (2000). *SPSS for Windows made simple: Release 10* ([New ed.). Hove: Psychology Press.

Kinnear, P. R., & Gray, C. D. (2009). *SPSS 16 made simple*. Hove; New York: Psychology Press.

Klein, E. C. (1995). Second versus third language acquisition: Is there a difference? *Language Learning, 45*(3), 419-466.

Klein, W., & Dittmar, N. (1979). *Developing grammars: The acquisition of German syntax by foreign workers*: Springer Verlag.

Kominski, R. (1989). *How good is 'How well'? An examination of the census English-speaking ability question*.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning, 49*(2), 303-342.

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning, 50*(2), 343-384.

Krashen, S. D. (1977). Some issues relating to the Monitor Model. In H. Brown, C. Yorio & R. Crymes (Eds.), *On TESOL'77*. Washington DC: TESOL.

Krashen, S. D. (1978). Individual variation in the use of the monitor. In W. Ritchie (Ed.), *Second Language Acquisition Research*. New York: Academic Press.

Krashen, S. D. (1980). The theoretical and practical relevance of simple codes in second language acquisition. In R. Scarcella & S. D. Krashen (Eds.), *Research in second language acquisition* (pp. 7-18). Rowley, MA: Newbury House.

Krashen, S. D. (1981). *Second language acquisition and second language learning* (1st ed.). Oxford; New York: Oxford University Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed.). Oxford; New York: Pergamon.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York: Longman.

Krashen, S. D., Butler, J., Birnbaum, R., & Robertson, J. (1978). Two studies in language acquisition and language learning. *LTL: Review of Applied Linguistics, 39*, 73-92.

Krashen, S. D., Long, M. A., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, 573-582.

Kruijff, G.-J. M., Kruijff-Korbayová, I., Bateman, J., & Teich, E. (2001). *Linear order as higher-level decision: Information structure in strategic and tactical generation.* Paper presented at the Annual Meeting of the ACL archive: Proceedings of the 8th European workshop on Natural Language Generation, Toulouse, France.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language 45*, 715-752.

Labov, W. (1991). Three dialects of English. In P. Eckert (Ed.), *New ways of analyzing variation in English* (pp. 1- 45). New York: Academic.

Labov, W. (1994). *Principles of linguistic change*. Oxford, England: Blackwell.

Labov, W. (2001). *Principles of linguistic change: Social factors*. Malden, Massachusetts: Blackwell Publisher Inc.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan Press.

Lafford, B. A. (1995). Getting into, through, and out of a survival situation: A comparison of communicative strategies used by students studying Spanish. In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context* (pp. 97-121). Amsterdam ; Philadephia: John Benjamins.

Lafford, B. A. (2004). The effect of the context of learning on the use of communication strategies by learners of Spanish as a second language. *Studies in Second Language Acquisition, 26*(2), 201-225.

Lafford, B. A. (2006). The effects of study abroad vs. classroom contexts on Spanish SLA: Old assumptions, new insights and future research directions. In C. A. Klee & T. L. Face (Eds.), *Selected Proceedings of the 7th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 1–25). Somerville, MA: Cascadilla Proceedings Project.

Lafford, B. A., & Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language. In R. Salaberry & B. A. Lafford (Eds.), *The art of teaching Spanish: Second language acquisition from research to praxis* (pp. 103–126). Washington, DC: Georgetown University Press.

Lakoff, R. T. (1975). *Language and women's place*. New York: Harper & Row.

Lamarre, C. (2008). The linguistic categorization of deictic direction in Chinese—with reference to Japanese—. In D. Xu (Ed.), *Space in languages of China: Cross-linguistic, synchronic and diachronic perspectives* (pp. 69-97): Springer Science + Business Media B.V.

Lamendella, J. (1977). General principles of neurofunctional organization and their manifestations in primary and non-primary language acquisition. *Language Learning, 27*, 155-196. *Language Learning Aptitude*. Rowley, Mass: Newbury House.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly, 9*, 409-430.

Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning, 26*, 125-134.

Larsen-Freeman, D., & Long, M. (1991). *An introduction to Second Language Acquisition research*. New York: Longman.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Lázaro Ibarrola, A., García Mayo, M. P., & Liceras, J. M. (2001). *Age and the acquisition of English as an L3 by bilingual (Basque-Spanish) children: Is the earlier the better?* Paper presented at the Proceedings of the Second International Conference on Third Language Acquisition and Trilingualism. CD-ROM document.

Leather, J. (2003). Phonological acquisition in multilingualism. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 23-58). Clevedon; Buffalo: Multilingual Matters.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Leow, R. P. (1996). Grammaticality judgment tasks and second-language development. *Linguistics, Language Acquisition, and Language Variation: Current Trends and Future Prospects, 47*, 126–139.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999). Lexical access in speech production. *Behavioral and brain sciences, 22*, 1-75.

Li, C. N. (1975). Synchrony vs. diachrony in language structure. *Language, 51*(4), 873-886.

Li, C. N., & Thompson, S. A. (1976). Subject and topic: A new typology of language. In C. N. Li (Ed.), *Subject and topic* (pp. 457-489). New York: Academic Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Li, D. (1996). *Waiguoren xue Hanyu yufa pianwu fenxi. Error analyses on foreigners' use of grammar features*. Beijing: Beijing Language and Culture University Press

Li, J. (1924). *New Mandarin Chinese grammar*. Beijing: The Commercial Press, 1992 reprinted.

Li, P. (1990). *Aspect and aktionsart in child Mandarin*. Unpublished doctoral dissertation, Leiden University, the Netherlands.

Li, P. (1993). The acquisition of the *zai* and *ba* constructions in Mandarin Chinese. In J. C. P. Liang & R. P. E. Sybesma (Eds.), *From classical 'Fú' to 'Three inches high': Studies on Chinese in honor of Erik Zürcher* (pp. 103-120). Leuven/Apeldoorn: Garant Publishers.

Li, X. [Xiangnong], Zhou, G., & Kong, L. (1990). 2-5 sui ertong yunyong 'ba' ziju qingkuang de chubu kaocha. The preliminary study of the use of the Ba-constructions by 2-5 years olds. *Yuwen Yanjiu. Journal of Language Studies, 37*(4), 43-50.

Li, X. [Xiaoshi] (2010). Sociolinguistic variation in the speech of learners of Chinese as a second language. *Language Learning, 60*(2), 366-408.

Li, Y., & Deng, X. (2005). "Ba" ziju yufa xiangmu de xuanqu yu paixu yanjiu. Study on selecting and sequencing of "ba" sentence patterns. *Yuyan jiaoxue yu yanjiu. Language Teaching and Linguistic Studies, 3*.

Li, Y.-H. A. (2001). The ba construction. http://www.usc.edu/dept/LAS/eacl/chinling/audreyli.htm.

Liceras, J. M. (1983). *Markedness, contrastive analysis and the acquisition of Spanish syntax by English speakers*. Unpublished doctoral dissertation, University of Toronto.

Lightbown, P. M. (1985). Great expectations: Second-language acquisition research and classroom teaching. *Applied Linguistics, 6*(2), 173.

Lin, H. (2001). *A grammar of Mandarin Chinese*. Muenchen LINCOM Europa.

Lin, Z. (2001). *Hanguo xuesheng xide "ba" ziju qingkuang de kaocha yu pianwu fenxi. The acquistionof the ba construction by Korean students and error analysis.* Unpublished M.A. thesis, Beijing Language and Culture University, Beijing.

Liu, F.-h. (1997). An aspectual analysis of BA. *Journal of East Asian Linguistics, 6*(1), 51-99.

Liu, F.-h. (2007). Word order variation and ba sentences in Chinese. *Studies in Language*, *31*(3), 649-682.

Liu, S., & Wang, Y. (2003). 'ba'ziju lianxi shejizhong de yujing wenti. Analysis on the contexts in exercises designed for Ba-structure. *Chinese Language Learning*(4), 59-66.

Liu, X. (2003). *Xin shiyong Hanyu keben. New practical Chinese reader*. Beijing: Beijing language and Culture University Audio-Visual Publishing House.

Liu, Y. [Yinglin] (Ed.). (1996). *Hanyu shuiping dengji biaozhun yu yufa dengji dagang. The standard of Chinese proficiency levels and outline of grammar proficiency levels* Beijing: Higher Education Press

Liu, Y. [Yuehua], Pan, W., & Gu, W. (2001). *Shiyong xiandai Hanyu yufa. Practical Modern Chinese Grammar* (X. Gao, Trans. 1 ed.). Beijing: The Commercial Press.

Liu, Y. [Yuehua], Yao, T.-c., Shi, Y., & Bi, N.-p. (2005). *Integrated Chinese: Level 2 Textbook: Traditional and Simplified Character Edition* (2nd ed.). Boston: Cheng & Tsui.

LoCoco, V. (1976). A comparison of three methods for the collection of L2 data: free composition, translation and picture description. *Working papers on Bilinguialism, 8*, 59-86.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol; New York; Ontario: Multilingual matters.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*(3), 251-285.

Long, M. H. (1993). Second language acquisition as a function of age: research findings and methodological issues. In K. Hyltenstam & A. Viberg (Eds.), *Progression & regression in language: Sociocultural, neuropsychological, & linguistic perspectives* (pp. 196-221). Cambridge: Cambridge University Press.

Long, M. H. (2005). Problems with supposed counter-evidence to the Critical Period Hypothesis. *International Review of Applied Linguistics in Language Teaching (IRAL), 43*(4), 287-317.

Long, M. H., & Sato, C. (1984). Methodological issues in interlanguage studies: An interactionist perspective. In A. Davies, C. Criper & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 253-279). Edinburgh: Edinburgh University Press.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp. 123-167). Clevedon, Avon: Multilingual Matters.

Lu, J., & Guo, R. (1998). Hanyu yufa yanjiu suo mianlin de tiaozhan. Challenges faced by research into Chinese Grammar. *Chinese Teaching in the World 46*(4), 3-21.

Lü, S. (1948). *Ba Zi Yongfa Yanjiu. Studies on the use of the "ba"* Beijing The Commercial Press, 1984 reprinted.

Lü, W. (1994). Baziju de yuyi leixing. The semantic categories of the ba sentence. *Hanyu xuexi. Chinese Learning*(4), 26-28.

Mackey, A., & Gass, S. M. (2005). *Second language research :Methodology and design*. Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition, 40*(1-2), 121-157.

MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of child language, 12*(2), 271-295.

Major, R. C. (2004). Gender and stylistic variation in second language phonology. *Language Variation and Change, 16*(3), 169-188.

Mandell, P. B. (1999). On the reliability of grammaticality judgement tests in second language acquisition research. *Second Language Research, 15*(1), 73-99.

Mansouri, F. (2005). Agreement morphology in Arabic as a second language: Typological features and their processing implications. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* (pp. 117-154). Melbourn: John Benjamin Publishing Co.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language and Hearing Research, 50*(4), 940.

Marriott, H. (1995). The acquisition of politeness patterns by exchange students in Japan. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 197-224). Amsterdam; Philadelphia: John Benjamins.

McArthur, E. K. (1993). *Language characteristics and schooling in the United States, a changing picture: 1979 and 1989*. Washington, DC: Office of Educational Research and Improvement, National Center for Educational Statistics.

McArthur, E. K., & Siegel, P. M. (1983). Developments in the measurement of English language proficiency *Proceedings of the American Statistical Association (Social Statistics Section)* (pp. 373-378).

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*(3), 395-423.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*(3), 381-401.

McLaughlin, B. (1978). The monitor model: Some methodological considerations. *Language Learning, 28*(2), 309-332.

Meara, P. (1994). The year abroad and its effects. *Language Learning Journal, 10*(1), 32-38.

Mei, L. (2005). Riben xuexizhe xide putonghua juanshe shengmu de yuyin bianyi yanjiu. Variability in pronouncing Chinese blade-palatal initials by Japanese SLA learners. *Shijie Hanyu Jiaoxue. Chinese Teaching in the World 71*(1), 97-105.

Milton, J., & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *ITL Review of Applied Linguistics, 107/108*, 17-34.

Möhle, D., & Raupach, M. (1983). *Planen in der Fremdsprach [Planning in the foreign language]*. Bern: Peter Lang.

Montrul, S. (2001). First-language-constrained variability in the second-language acquisition of argument-structure-changing morphology with causative verbs. *Second Language Research, 17*(2), 144.

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.

Mougeon, R., & Rehner, K. (2001). Acquisition of sociolinguistic variants by French immersion students: The case of restrictive expressions, and more. *The Modern Language Journal, 85*(3), 398-415.

Munnich, E., Flynn, S., & Martohardjono, G. (1994). Elicited imitation and grammaticality judgment tasks: What they measure and how they relate to each other. In S. M. Gass, E. Tarone & A. D. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 227-243). Hillsdale, NJ: Lawrence Erlbaum Associates.

Muñoz, C. (2003). Variation in oral skills development and age of onset. In M. P. García Mayo & M. L. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 161-181). Clevedon; Buffalo: Multilingual Matters.

Murphy-Lejeune, E. (2002). *Student mobility and narrative in Europe: the new strangers*. London; New York: Routledge.

Nadasdi, T. (2010). Agreeing to disagree: Variable subject-verb agreement in immersion French. *Canadian Journal of Applied Linguistics (CJAL)/Revue canadienne de linguistique appliquée (RCLA), 4*(1), 87.

Naiman, N. (1974). The use of elicited imitation in second language acquisition research. *Working Papers on Bilingualism, 2*, 1-37.

Nisbet, D., Tindall, E., & Arroyo, A. (2005). Language learning strategies and English proficiency of Chinese university students. *Foreign Language Annals, 38*(1), 100-107.

Norton, B., & Pavlenko, A. (2004). Addressing gender in the ESL/EFL classroom. *TESOL Quarterly, 38*(3), 504-514.

Nyikos, M. (1990). Sex related differences in adult language learning: Socialization and memory factors. *The Modern Language Journal*(3), 273-287.

O'Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition, 29*(4), 557-581.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.

Odlin, T. (2003). Cross-linguistic influence. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 436-486). Malden, Oxford, Carlton: Blackwell.

Odlin, T. (2005). Crosslinguistic influence and conceptual transfer: What are the concepts? *Annual Review of Applied Linguistics, 25*, 3-25.

Opper, S., Teichler, U., & Carlson, J. (1990). *The impact of study abroad programmes on students and graduates*. London: Jessica Kingsley Publishers.

Oyama, S. (1976). A sensitive period for the acqusition of a nonnative phonological system. *Journal of Psycholinguistic Research, 5*, 261-283.

Oyama, S. (1978). The sensitive period and comprehension of speech. *Working Papers on Bilingualism, 16*, 1-17.

Oyama, S. (1979). The concept of the sensitive period in developmental studies. *Merrill-Palmer Quarterly, 25*, 83-102.

Pallant, J. F. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows (Version 15)* (3rd ed.). Crows Nest, N.S.W.: Allen & Unwin.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam; Philadelphia: John Benjamins Publishing.

Patkowski, M. S. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning, 30*(2), 449-468.

Patkowski, M. S. (1994). The critical age hypothesis and interlanguage phonology. In M. Yavas (Ed.), *First and second language phonology* (pp. 205–221): Singular Publishing Group.

Pavlenko, A., & Piller, I. (2001). New directions in the study of multilingualism, second language learning, and gender. In A. Pavlenko (Ed.), *Multilingualism, second language learning, and gender* (pp. 17–52). Berlin: Walter de Gruyter.

Pellegrino Aveni, V. A. (2005). *Study abroad and second language use: Constructing the self*. Cambridge; New York; Port Melbourne; Madrid; Cape Town: Cambridge University Press.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.

Pica, T. (1983). Adult acquisition of English as a second language under different conditions of exposure. *Language Learning, 33*(4), 465-497.

Pica, T., Halliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition, 11*, 63-90.

Pienemann, M. (1987). Determining the influence of instruction on L2 speech processing. *Australian Review of Applied Linguistics, 10*(2), 83-113.

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics, 10*(1), 52-79.

Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam; Philadelphia: John Benjamins.

Pienemann, M. (2005). *Cross-linguistic aspects of processability theory*. Amsterdam; Philadelphia: J. Benjamins Pub. Co.

Pienemann, M., Di Biase, B., & Kawagichi, S. (2005). Extending processability theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* Amsterdam; Philadelphia: J. Benjamins Pub. Co.

Plough, I., & Gass, S. M. (1993). Interlocutor and task familiarity: Effects on interactional structure. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 35-56). Clevedon, England: Multilingual Matters.

Polanyi, L. (1995). Language learning and living abroad. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 271–291). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Preston, D. R. (1989). *Sociolinguistics and second language acquisition*. Oxford, UK; New York, NY, USA: Blackwell.

Preston, D. R. (1996). Variationist perspectives on second languae acquisition. In R. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 1-45). Amsterdam/ Philadelphia: John Benjamins B.V.

Qian, X. (1997). Riben Liuxuesheng hanyu quxiang buyu xide shunxu. The Acquisition order of the Chinese directional complements by Japanese students. *Shijie Hanyu Jiaoxue. Chinese Teaching in the World*(1).

Regan, V. (1995). The acquisition of sociolinguistic native speech norms. In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 245-267). Amsterdam; Philadelphia: John Benjamins.

Regan, V. (1996). Variation in French interlanguage: A longitudinal study of sociolinguistic competence. In R. Bayley & D. R. Preston (Eds.), *Second Language Acquisition and Linguistic Variation* (pp. 177-201). Amsterdam, The Netherlands: John Benjamins.

Regan, V. (1998). Sociolinguistics and language learning in a study abroad context. *Frontiers, 4*, 61-90.

Regan, V., Howard, M., & Lemée, I. (2009). *The acquisition of sociolinguistic competence in a study abroad context*. Bristol; Buffalo: Multilingual Matters.

Rehner, K. (2002). *The development of aspects of linguistic and discourse competence by advanced second language learners of French.* Unpublished doctoral dissertation, University of Toronto, Canada, Toronto.

Rehner, K., Mougeon, R., & Nadasdi, T. (2003). The learning of sociolinguistic variation by advanced FSL learners. *Studies in Second Language Acquisition, 25*(01), 127-156.

Ren, X., & Xu, J. (2005). *Boya Hanyu: Chuji Qibupian (1). Boya Chinese: Elementary Chinese - 1 Getting started*. Beijing: Peking University Press.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studes in Second Language Acquisition 18*(1), 27-67.

Robinson, P. (2001a). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design: A tradic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press.

Robinson, P. (2001c). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2002). Learning conditions, aptitude complexes and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning*. Amsterdam: John Benjamins.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition, 27*(02), 235-268.

Robinson, P., & Ellis, N. (2008). *Handbook of cognitive linguistics and second language acquisition*. New York / London Routledge

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Rodriguez, S. (2001). *The perception of requests in Spanish by instructed learners of Spanish in the second and foreign-language contexts: A longitudinal study of acquisition patterns.* Unpublished doctoral dissertation, Indiana University, Bloomington.

Romaine, S. (2003). Variation. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 409-435). Malden, Oxford, Carlton: Blackwell Publishing.

Rosansky, E. J. (1976). Methods and morphemes in second language acquisition research. *Language Learning, 26*(2), 409-425.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*, 532-547.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Mahwah, NJ: Erlbaum.

Sabourin, L., Stowe, L., & de Haan, G. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research, 22*(1), 1-29.

Salaberry, R. (1998). The development of aspectual distinctions in classroom L2 French. *Canadian Modern Language Review, 54*(4), 508-542.

Sanders, R., & Yao, N. (2009). *Fundamental Spoken Chinese*. Honolulu: University of Hawai'i Press.

Schachter, J. (1989). Testing a proposed universal. In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 73-88). Melbourne, Australia: Cambridge University Press.

Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research, 6*(2), 93-124.

Schmidt, M. (1980). Coordinate structures and language universals in interlanguage. *Language Learning, 30*(2), 397-416.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge.

Schwartz, B., & Sprouse, R. (1994). Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. Schwartz (Eds.), *Language acquisition studies in generative grammar* (pp. 317-368). Amsterdam, Philadelphia: John Benjamins.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the Full Transfer/Full Access Model. *Second Language Research, 12*, 40-72.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House.

second language learning. In K. Diller (Ed.), *Individual Differences and Universals in*

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*(2), 173-199.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10*, 209-231.

Selinker, L., & Douglas, D. (1985). Wrestling with 'context' in interlanguage theory. *Applied Linguistics, 6*(2), 190-204.

Shameem, N. (1998). Validating self-reported proficiency by testing performance in an immigrant community: The Wellington Indo-Fijians. *Language Testing, 15*(1), 86-108.

Shannon, J. F. (1995). *Variability and the interlanguage production of the English definite article.* Unpublished doctoral dissertation, Ohio State University.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*, 118-132.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition 15*(2), 165-179.

Sharwood Smith, M., & Kellerman, E. (1986). Crosslinguistic influence in second language acquisition: An introduction. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 1-9). Oxford: Pergamon Institute of English.

Shen, J. (2002). Ruhe chuzhi 'chuzhishi'- shilun baziju de zhuguanxing. Can the disposal construction be disposed of? - On the subjectivity of Ba construction in Mandarin Chinese. *Zhongguo Yuwen. Chinese Language, 290*(5), 387-478.

Shi, C. (2010). Cong jushi qun kan baziju ji xiangguan jushi de yufa yiyi. A study of the grammatical meanings of the ba-construction and other relevant constructions from the perspective of construction group. *Shijie hanyu jiaoxue. Chinese Teaching in the World, 24*(3), 291-309.

Shi, D. (1999). Baziju he beiziju yanjiu. A study into the ba construction and the bei construction. In L. Xu (Ed.), *Gongxing yu gexing - Hanyu yuyan xue zhong de zhengyi. Commonality and individuality- Contentions in Chinese linguistics* (pp. 111-138). Beijing: Beijing yuyan wenhua daxue chubanshe. Beijing language and cultural university press

Shi, J. (1998). Waiguo liuxuesheng 22 lei xiandai hanyu jushi de xide shunxi yanjiu. The acquisition order of 22 sentence structures in Modern Chinese. *Shijie hanyu jiaoxue. Chinese Teaching in the World, 46*(4), 77- 98.

Shi, J. (2006). Guonei hanyu di er yuyan xide yanjiu ershi nian. A review and prospects of Second Language Acquisition research of Chinese. *Yuyan jiaoxue yu yanjiu. Language Teaching and Linguistic Studies*(1), 15-26.

Siegel, J. (2003). Social Context. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 178-223). Malden, Mass: Blackwell.

Singleton, D. M. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching (IRAL), 43*(4), 17.

Singleton, D. M., & Ryan, L. (2004). *Language acquisition: The age factor* (2nd ed.). Buffalo, N.Y.: Multilingual Matters Ltd.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93-120.

Slavoff, G., & Johnson, J. (1995). The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition, 17*(01), 1-16.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 1114-1128.

Song, Y. (1979). Chuzhishi xinjie. New explanations of the 'disposal' construction. *Tianjin shifan xueyuan xuebao. News paper of Tianjin Normal Institute* (3).

Song, Y. (1981). Guanyu 'ba' ziju de liang ge wenti. Two issues regarding the 'ba' construction *Yuwen Yanjiu. Chinese Studies*(2).

Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics, 6*, 239-254.

Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non native grammars of Italian. *Second Language Research, 9*(1), 22-47.

Sorace, A. (1996). The use of acceptability judgments in second language acquisition research. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 373-409). San Diego: Academic Press.

Stevens, J. J. (2001). *The acquisition of L2 Spanish pronunciation in a study abroad context.* Unpublished Ph.D. thesis, University of Southern California, United States -- California.

Sun, C. (1995). Transitivity, the ba construction and its history. *Journal of Chinese Linguistics, 23*(1), 159-195.

Sun, C. (1996). *Word-order change and grammaticalization in the history of Chinese*. Stanford, Calif.: Stanford University Press.

Sun, C. (2006). *Chinese: A linguistic introduction*. Cambridge: Cambridge University Press.

Sun, C., & Givón, T. (1985). On the so-called SOV word order in Mandarin Chinese: A quantified text study and its implications. *Language, 61*(2), 329-351

Sybesma, R. P. E. (1999). *The Mandarin VP*. Dordrecht, [Netherlands] ; Boston: Kluwer Academic Publishers.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Tai, J. H.-Y. (1985). Temporal sequence and word order in Chinese. In J. Haiman (Ed.), *Iconicity in Syntax* (pp. 49-72). Amsterdam John Benjamins.

Talmy, L. (2000a). *Toward a cognitive semantics, Volume 1: Concept structuring systems*. Cambridge, MA: MIT Press.

Talmy, L. (2000b). *Towards a cognitive semantics II: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Tanaka, K., & Ellis, R. (2003). Study abroad, language proficiency, and learner beliefs about language learning. *JALT JOURNAL, 25*(1), 63-85.

Tarone, E. (1979). Interlanguage as chameleon. *Language Learning, 29*(1), 181-191.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-82.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4*(2), 142-164.

Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning, 35*(3), 373-403.

Tarone, E. (1987). Methodologies for studying variability in second language acquisition. In R. Ellis (Ed.), *Second language acquisition in context* (pp. 35-46). Englewood Cliffs: Prentice-Hall Inc.

Tarone, E. (1988). *Variation in interlanguage*: London: Edward Arnold.

Tarone, E. (1989). Accounting for style-shifting in interlanguage. In S. M. Gass (Ed.), *Variation in second language acquisition volumn II: Psycholinguistic issues* (pp. 13-21). Clevedon; Philadelphia: Multilingual Matters Ltd.

Tarone, E. (2002). Frequency effects, noticing, and creativity *Studies in Second Language Acquisition, 24*(02), 287-296.

Tarone, E. (2007). Sociolinguistic approaches to second language acquisition research---1997-2007. *The Modern Language Journal, 91* (Focus Issue), 837-848.

Tarone, E., & Liu, G.-q. (1995). Situational context, variation, and second language acquisition theory. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 107-124). Oxford; New York: Oxford University Press.

Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language learning, 38*(1), 21-44.

Taylor, J. R. (1995). *Linguistic categorization: Prototypes in linguistic theory* (2 ed.). Oxford: Clarendon.

Taylor, J. R. (1998). Syntactic constructions as prototype categories. In M.Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 177-202). Mahwah, NJ:Erlbaum.

Taylor, J. R. (2003). *Linguistic categorization*. Oxford; New York: Oxford University Press.

Teng, S.-H. (1975). *A semantic study of transitivity relations in Chinese*. Birkeley: University of California Press.

Thomas, A. (2004). Phonetic norm versus usage in advanced French as a second language. *International Review of Applied Linguistics in Language Teaching (IRAL), 42*(4), 365-382.

Thorne, B., & Henley, N. (1975). Language and Sex: Difference and Dominance.

Tomasello, M. (2003). *Constructing a language: A usage based view of language acquisition*. Cambridge MA: Harvard University Press.

Torres, J. P. (2003). *A cognitive approach to the acquisition of clitics in L2 Spanish: Insights from study abroad and classroom learners.* Unpublished Ph.D. thesis, Cornell University, Ithaca, NY.

Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition, 15*(2), 181-204.

Tsao, F.-f. (1987). A topic-comment approach to the *ba* construction. *Journal of Chinese Linguistics, 15*(1), 1-54.

Tsao, F.-f. (1990). *Sentence and clause structure in Chinese: A functional perspective*. Taipei, Taiwan, Republic of China: Student Book Co.

Tse, J., Tang, T., Shie, Y., & Li, Y. (1991). *Chinese children's language acquisition and development. Unpublished technical report*. Taipei, Taiwan: National Taiwan Normal University.

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*(03), 287-301.

Victori , M., & Tragant, E. (2003). Learner strategies: A cross-sectional and longitudinal study of primary and high-school EFL learners. In M. P. Garc á Mayo & M. L. Garc á Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 182-209). Clevedon; Buffalo: Multilingual Matters.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Learning, Memory, 18*(2), 328-344.

Walsh, T., & Diller, K. (1981). Neurolinguistic considerations on the optimal age for

Wang, H. (1981). *'Ba' ziju he 'Bei' ziju. The ba construction and the bei construction* Shanghai: Shanghai Jiaoyu Chubanshe. Shanghai Education Press

Wang, H. (1985). 'Ba' ziju zhong 'ba' de binyu. The object of the ba in the ba construction. *Zhongguo yuwen. Chinese linguistics, 184*(1), 48.

Wang, H. (1995). *Duiwai Hanyu Jiaoxue Yufa Dagang 'Syllabuses for Pedagogical Grammar of Chinese as a Second Language'*. Beijing: Beijing Language Institute Press.

Wang, L. (1943). *Zhongguo xiandai yufa. Modern Chinese Grammar*. Beijing: The Commercial Press, 1985 reprinted.

Wang, Y., Jiao, Q., & Pang, Y. (1987). *Hanyu dongci jieguobuyu dapei cidian. Dictionary of collocation of Chinese verbs and resultative complement* (Di 1 ban. ed.). Beijing: Beijing yuyan xueyuan chubanshe (Beijing Language Institute Press).

Weinreich, U. (1953). *Languages in contact*: The Hague: Mouton.

White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness, 11*(4), 259-290.

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research, 12*(3), 233.

White, L., Valenzuela, E., Kozlowska–Macgregor, M., & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics, 25*(01), 105-133.

Whitman, R., & Jackson, K. (1972). The unpredictability of contrastive analysis *Language Learning, 22*, 29-41.

Willis, F. M., Doble, G., Sankarayya, U., & Smithers, A. (1977). *Residence abroad and the student of modern languages: A preliminary survey*. Bradford: University of Bradford, Modern Languages Centre.

Wolfram, W. (1985). Variability in tense marking: A case for the obvious. *Language Learning, 25*(2), 229-253.

Wu, Z. (2001). *Empirical study of the Chinese ba construction as a second language*. Unpublished M.A. thesis, National Taiwan Normal University.

Xiong, W. (1996). Liuxuesheng 'ba' zi jiegou de biaoxian fenxi. Performance analysis of the ba construction by L2 learners. *Chinese Teaching in the World, 35*(1), 80-87.

Xu, D. (2006). *Typological change in Chinese syntax*. Oxford ; New York: Oxford University Press.

Xu, J., & Ren, X. (Eds.). (2004). *Boya Hanyu: Chuji Qibupian II. Boya Chinese: Elementary - Getting started II*. Beijing: Peking University Press.

Yamada, J., Takatsuka, S., Kotabe, N., & Kuruse, J. (1980). On the optimum age for teaching foreign vocabulary to children. *International Review of Applied Linguistics in Language Teaching (IRAL), 28*, 245-247.

Yang, S. (1998a). Cong qingzhuang leixing lai kan 'ba'ziju (shang). The aspectal typology of the ba construction(I). *Hanyu xuexi. Chinese Language Learning*(2), 10-13.

Yang, S. (1998b). Cong qingzhuang leixing lai kan 'ba'ziju (xia). The aspectal typology of the ba construction(II). *Hanyu xuexi. Chinese Language Learning*(3), 10-12.

Yao, N., Lee, M., & Sanders, R. M. (2009). *Fundamental Written Chinese*. Honolulu: University of Hawai'i Press.

Yao, T.-c., Liu, Y., Bi, N.-p., Chen, Y.-F., Ge, L., Hayden, J. J., et al. (2005). *Integrated Chinese, Level 1, Part 2: Textbook, Simplified Characters*. Boston: Chung & Tsui Company.

Ye, X. (2004). Baziju de zhishixing jieshi. Causative interpretation of the Ba-sentence. . *Shijie hanyu jiaoxue. Chinese Teaching in the World* (2), 25-39.

Yew, E. K. K. (1995). *Investigation of the critical period and second language acquisition: Johnson and Newport refuted.* Unpublished honours thesis, York University.

Young, R. (1989). *Variation in interlanguage morphology: (s) Plural-marking in the speech of Chinese learners of English.* Unpublished doctoral disseration, University of Pennsylvania, United States -- Pennsylvania.

Young, R. (1991). *Variation in interlanguage morphology*. New York: Peter Lang.

Young, R. (1996). Form-function relations in articles in English interlangauge. In R. J. Bayley & D. R. Preston (Eds.), *Language acquisition and linguistic variation* (pp. 135-175). Amsterdam ; Philadelphia: J. Benjamins.

Young, R., & Bayley, R. J. (1996). VARBRUL Analysis for Second Language Acquisition Research. In R. J. Bayley & D. R. Preston (Eds.), *Second language acquisition and linguistic variation* (pp. 253-306). Amsterdam; Philadelphia: John Benjamins.

Yu, W. (2000). A report of an investigation of CSL learners using of 'ba' sentence. *Chinese Language Learning*(5), 49-54.

Yuan, B. (1995). Variability and systematicity in the performance of the four Chinese tones by English SLA learners of Chinese. *Shijish Hanyu Jiaoxue. Chinese Teaching in the World, 31*(1), 47-62.

Yuan, B. (2004). Negation in French–Chinese, German–Chinese and English–Chinese interlanguages. *Transactions of the Philological Society, 102*(2), 169-197.

Yuan, B. (2006). *Variability at different interfaces in the development of English speakers' L2 Chinese Wh-questions.* Paper presented at the the 8th Generative Approach to Second Language Acquisition Conference (GASLA 6), Somerville, MA.

Yuan, B. (2007). Behaviours of *wh*-words in English speakers' L2 Chinese *wh*-questions: Evidence of no variability, temporary variability and persistent variability in L2 grammars. *Bilingualism: Language and Cognition, 10*(3), 277-298

Yuan, B. (2010). Domain-wide or variable-dependent vulnerability of the semantics--syntax interface in L2 acquisition? Evidence from wh-words used as existential polarity words in L2 Chinese grammars. *Second Language Research, 26*(2), 219.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1-27.

Yule, G., & Macdonald, D. (1990). Resolving referential conflicts in L2 interaction: The effect of proficiency and interactive role. *Language Learning, 40*(4), 539-556.

Zeldes, A. (2009). Applying morphological productivity measures to syntactic constructions: German comparatives and the *je --- desto* construction. from http://docs.google.com/viewer?a=v&q=cache:lS-quJLaYEEJ:www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiter-innen/amir/pdf/DocDay2009+productivity+of+morphological+words+based+on+the+same+criteria:+unencountered,+regular+and+transparent+(e.g.+Zeldes+2009)&hl=zh-CN&pid=bl&srcid=ADGEESj1vNIrB0nkv-VDMJQgPdDfm-DKjE9kKw3hxXrjuBcQXiJV1s_fkdEIBmQFikGo74Z8AQPE4h3h9xKp5316llqS-

LRwjM7Pbz0SFYH05VQAmbCmb2QKroGOuzrlZYN3cdq9H-PJ&sig=AHIEtbRfnOQywLIlF3ZalGl7oiqquTEvQw

Zhang, B. (2000). Lun 'ba' ziju de jushi yuyi. On the semantic function of the 'ba' construction *Yuyan yanjiu. Language Studies 38*(1), 29-39.

Zhang, S. (2002). *Second language acquisition of the ba-construction in contemporary Mandarin Chinese.* Unpublished M.A. thesis, University of Southern California, United States -- California.

Zhang, W. (2001). "Ba"ziju de weiyi tushi. The displacement schema of the *ba* sentence. *Yuyan jiaoxue yu yanjiu. Language Teaching and Linguistic Studies* (3), 1-10.

Zhang, Y. (2001). *Second language acquisition of Chinese grammatical morphemes: A processability perspective.* Unpublished Ph.D. thesis, The Australian National University, Canberra.

Zhang, Y. (2002). A processing approach to the L2 acquisition of Chinese grammatical morpheme. In B. Di Biase (Ed.), *Developing a second language: Acquisition, processing and pedagogy of Arabic, Chinese, English, Italian*. Melbourne: Language Australia.

Zhang, Y. (2004). Processing constraints, categorical analysis and the second language acquisition of the Chinese adjective suffix -de (ADJ). *Language Learning, 54*(3), 437-468.

Zhang, Y. (2005). Processing and formal instruction in the L2 acquisition of five Chinese grammatical morphemes. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* (pp. 155-177). Amsterdam; New York: John Benjamins.

Zhao, S., Liu, S., & Hu, X. [Xiang] (1997). Danju juxing tongji yu fenxi. An analysis of the frequency of simple sentence structures. *Yuyan jiaoxue yu yanjiu. Language Teaching and Linguistic Studies*(2), 61-73.

Zhou, H. (2006). Lun 'ba' ziju de zhengxiang zhishixing yu jushi tezheng. On the positive causativity and sentence feature of the *ba* construction. *Journal of Yunmeng, 27*(1), 132-135.

Zobl, H. (1995). Converging evidence for the "acquisition-learning" distinction. *Applied Linguistics, 16*, 35-56.

Zobl, H., & Liceras, J. (1994). Functional categories and acquisition orders. *Language Learning, 44*(1), 159-180.

Zou, K. (1995). *The syntax of the Chinese BA-constructions and verb compounds: A morphosyntactic analysis.* Unpublished Ph.D., University of Southern California, United States -- California.

273