

DEFINING AND ASSESSING CHINESE SYNTACTIC COMPLEXITY

VIA TC-UNITS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE
UNIVERSITY OF HAWAII AT MĀNOA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

EAST ASIAN LANGUAGES AND LITERATURES (CHINESE)

DECEMBER 2016

By

Qiaona Yu

Dissertation Committee:

Song Jiang, Chairperson

John M. Norris

Haidan Wang

Li Julie Jiang

Marta González-Lloret

Copyright 2016 by Qiaona Yu

All Rights Reserved

I humbly dedicate this dissertation to Dr. John M. Norris.

ACKNOWLEDGMENTS

The research reported in this dissertation would not have been possible without the support, participation, and encouragement of many people. First and foremost, I am deeply indebted to my committee. I am most fortunate to have had the continuous support and encouragement of my Chairperson, Dr. Song Jiang, who has been like “Timely Rain” (*Jishi Yu Song Jiang*) throughout my doctoral work. Conversations with him have been inspirational on not only how to be a good researcher but also how to be a good person. I am profoundly grateful to my “Co-Chair”: Dr. John Norris, who has been my academic role model since the first day I walked into his seminar class. I cannot thank him enough for his willingness to join my doctoral committee, even after taking leave, advising me from afar. This dissertation truly would not have been possible without his patience, suggestions, and support. I am very fortunate to have had Dr. Haidan Wang and Dr. Li Julie Jiang on my committee and providing insightful feedback. *Gracias* to Dr. Marta González-Lloret for her insightful comments and suggestions. Her lucidity and vision gave me clear academic direction. I am deeply indebted to Dr. Tao-chung Yao who had been my first Ph.D. advisor and a formal committee member. Years ago without his encouragement and support, I would not have embarked on my journey here at University of Hawaii at Manoa (UHM).

Second, I also deeply appreciate the many individuals who willingly helped in during my dissertation process. Heartfelt thanks go to Dr. Richard Medina, who provided me with technological assistance for the web-based test TW&ST. I could not have developed it on my own. I am very grateful to my friend Dr. Ying Zhou and Dr. Shu-Ling

Wu for kindly allowing the EI test to be adopted as a part of the TW&ST. Special thanks to Dr. Ying Zhou for her meticulous work rating the EI test data with me. *Duoxie* to Stephen Tschudi for his professional job dubbing the video in TW&ST. *Gomabseubnida* to my fellow student Dr. Dongmin Kim for the thought-provoking discussions in our study group. My most sincere mahalo to all my wonderful fellow students, colleagues, and friends for their help in my data collection and writing development: Wendan Li, Jing Wu, Chuan Lin, Di Sun, Daniel Tschudi, Xin Wu, Chengfeng Zhan, Wenlong Wang, Sean Forte, Maurice Hutchinson, Linette Char, and Tiffany Ing. A special thank you to my M.A. advisor Dr. Xiaoqi Li, and Dr. Yang Zhao at Peking University (PKU) for their encouragement and support during my data collection. I would like to also acknowledge the financial support from the China Scholarship Council and UHM-PKU Exchange Program.

Last but not least, I am most indebted to my family for their unyielding love and trust. I am fortunate to have the most understanding and loving mother who gave me unconditional support, chasing my dreams.

ABSTRACT

The triad dimensions of complexity, accuracy, and fluency (CAF) has been widely used for assessing second language performance and development. Unlike accuracy and fluency, the construct of Chinese syntactic complexity has not been comprehensibly conceptualized or operationalized. Moreover, not tailored to the typological differences such as the topic prominence of the Chinese language, measures developed globally were found not as valid for Chinese syntactic complexity assessment as they are for Indo-European languages. Research indicated that the mean length of the T-unit of native Chinese speakers is shorter than that of L2 Chinese speakers (Jin, 2006; Yuan, 2009).

For situations where research findings developed globally are not as applicable when indiscriminately applied to typologically different languages, this dissertation employed the notion of *GlobaLocality* to define and assess Chinese syntactic complexity. First, globally, clause combining was revisited to subsume the topic chain in addition to coordination and subordination. An organic approach was then adopted to investigate complexity via global, clausal, and subclausal levels (Norris & Ortega, 2009). Second, locally, a taxonomy of Topic-Comment units (TC-units) was proposed to examine Chinese syntactic complexity: the number and the nature of a terminable TC-unit's components; and the number and the nature of their constituent relationship. Third, by performing discriminant function analyses on L1 and L2 Chinese speakers' spoken ($N=115$) and written ($N=116$) output elicited from a designed online test, a series of proposed TC-unit based measures were confirmed with high efficiency (61.2%~76.5%) at proficiency group membership classification. Lower-proficiency speakers produced

shorter terminable TC-units consisting of fewer single TC-units, whereas higher-proficiency speakers produced longer terminable TC-units in the form of varied topic chains consisting of more single TC-units. Chinese syntactic complexity development along proficiency increase also displayed a transition from more lengthening to more combining of single TC-units. Fourth, utilizing TC-unit based measures, repeated measures analyses observed more complex language produced in more complex tasks along the resource-directing dimension. Immediate task repetition was observed to lower learners' communication anxiety and improve learners' self-perceived performance. Last, this dissertation provided suggestions on complexity descriptions for proficiency guidelines and on how to develop Chinese syntactic complexity in classroom instruction.

TABLE OF CONTENTS

| | |
|---|------|
| Acknowledgments | iii |
| Abstract | v |
| List of tables | xi |
| List of figures | xiii |
| List of abbreviations and symbols | xiv |
| | |
| Chapter 1: Introduction | 1 |
| 1.1 Motivation of the present study | 1 |
| 1.2 Research Outline | 6 |
| Chapter 2: Defining Chinese syntactic complexity | 12 |
| 2.1 The construct of complexity | 12 |
| 2.1.1 The CAF model in SLA | 12 |
| 2.1.2 Complexity in the triad of CAF | 14 |
| 2.1.3 Conceptualizing and operationalizing syntactic complexity | 17 |
| 2.1.3.1 Defining syntactic complexity | 18 |
| 2.1.3.2 Identifying behavior of syntactic complexity | 21 |
| 2.1.3.3 Task specification for syntactic complexity output | 23 |
| 2.1.3.4 Behavior elicitation for syntactic complexity output | 24 |
| 2.1.3.5 Observation scoring on syntactic complexity output | 25 |
| 2.1.3.6 Data analysis on syntactic complexity output | 27 |
| 2.2 Defining Chinese syntactic complexity | 28 |

| | |
|--|-----|
| 2.2.1 Topic-prominent versus subject-prominent | 30 |
| 2.2.1.1 Topic-comment structure | 30 |
| 2.2.1.2 Coreferential zero | 32 |
| 2.2.1.3 Types of topic chains | 39 |
| 2.2.2 Parataxis-prominence versus hypotaxis-prominence | 50 |
| 2.2.2.1 Parataxis and hypotaxis | 50 |
| 2.2.2.2 Coordination and subordination | 56 |
| 2.2.2.3 Clarifying parataxis, hypotaxis, coordination, and subordination | 77 |
| 2.2.2.4 Covert connectives | 82 |
| 2.2.3 Sentence-oriented versus discourse-oriented | 84 |
| 2.3 Linguistic complexity and task influences | 89 |
| 2.3.1 Cognitive task complexity | 89 |
| 2.3.2 Learner affective variables | 94 |
| 2.3.3 Immediate task repetition for task planning | 95 |
| Chapter 3: Assessing Chinese syntactic complexity | 97 |
| 3.1 T-unit as the unit of analysis | 97 |
| 3.2 Other units of analysis | 100 |
| 3.3 TC-unit as the unit of analysis | 103 |
| 3.4 Elicited Imitation (EI) test | 114 |
| Chapter 4: Research design and methods | 120 |
| 4.1 Research questions | 120 |
| 4.2 Data collection | 120 |
| 4.2.1 Participants | 120 |

| | |
|---|-----|
| 4.2.2 Instruments | 123 |
| 4.2.3 Procedures | 129 |
| 4.3 Scoring and Analysis | 136 |
| 4.3.1 Transcribing | 136 |
| 4.3.2 Measures | 138 |
| 4.3.3 Coding | 139 |
| 4.3.4 Participant classification | 152 |
| Chapter 5: Results and discussion | 155 |
| 5.1 Data screening | 155 |
| 5.2 Results | 162 |
| 5.2.1 Descriptive statistics | 162 |
| 5.2.2 Correlational analysis | 166 |
| 5.2.3 Discriminant function analysis | 171 |
| 5.2.3.1 Speaking tasks | 172 |
| 5.2.3.2 Writing tasks | 183 |
| 5.2.3.3 Interim summary | 198 |
| 5.2.4 Chinese syntactic complexity and cognitive task complexity | 199 |
| 5.2.4.1 Higher language complexity produced in more complex tasks | 199 |
| 5.2.4.2 Immediate task repetition on learners' self-perception | 204 |
| Chapter 6: Discussion and implications | 215 |
| 6.1 Qualitative analysis | 215 |
| 6.1.1 Overuse of complete sentences | 217 |
| 6.1.2 Single TC-unit lengthening and combining | 221 |

| | |
|--|-----|
| 6.1.3 Hypotactic and paratactic syntagma | 232 |
| 6.1.4 Typical topic chain and varied types of topic chain | 240 |
| 6.2 Chinese syntactic complexity in proficiency guidelines | 247 |
| 6.3 Pedagogical implications | 250 |
| 6.3.1 Complexity developmental stages | 250 |
| 6.3.2 Composing Terminable TC-unit | 254 |
| 6.3.3 Teaching TC-unit in class | 257 |
| 6.3.3.1 Task focusing on composing and contrasting | 258 |
| 6.3.3.2 Immediate task repetition for different proficiency levels | 273 |
| 6.3.4 Interim summary | 274 |
| Chapter 7: Conclusions | 278 |
| 7.1 Contribution of the present study | 278 |
| 7.1.1 Findings of this dissertation | 278 |
| 7.1.2 Methodological implication | 287 |
| 7.2 Limitations of the study and suggestions for future studies | 293 |
| Appendix A Ten types of Chinese topic chains | 297 |
| Appendix B TW&ST (English instruction) screenshots | 300 |
| Appendix C TW&ST (Chinese instruction) screenshots | 325 |
| Appendix D Mandarin Chinese EI test | 348 |
| Appendix E EI task scoring rubric | 351 |
| References | 356 |

LIST OF TABLES

| | |
|--|-----|
| Table 1. Categories of clause complexing in English | 57 |
| Table 2. Coordination and subordination in English | 60 |
| Table 3. Categories of Chinese complex sentences | 71 |
| Table 4. Inventory of Chinese syntactic complexity measures | 110 |
| Table 5. Calculation of complete and valid spoken and written data sets | 123 |
| Table 6. Spoken data transcribing guidelines | 137 |
| Table 7. A coding sample for output (59) | 145 |
| Table 8. Calculated scores of output (59) on four measures | 146 |
| Table 9. Correlation coefficient between the two codings for each task on each measure | 147 |
| Table 10. Some special marks for coding | 149 |
| Table 11. Total cases for speaking and writing tasks by proficiency group | 154 |
| Table 12. The results of discriminant analyses on the data with or without adjustment | 157 |
| Table 13. Box's M output across three groups for different tasks | 159 |
| Table 14. Pearson correlation matrix of all the predictors with each other by task | 161 |
| Table 15. The mean values for each measure on five tasks by proficiency level | 164 |
| Table 16. Correlation between EI scores of L2 Chinese speakers and their scores on the four complexity measures by task | 166 |
| Table 17. Classification results for the CS task | 174 |
| Table 18. Classification results for the V1 task | 176 |
| Table 19. Classification results for the V2 task | 178 |
| Table 20. Two functions of each discriminant analysis for speaking tasks | 181 |

| | |
|--|-----|
| Table 21. Classification results for the three speaking tasks with MLTTCU as the only predictor | 182 |
| Table 22. Classification results for the FW task | 184 |
| Table 23. Classification results for the GR task | 186 |
| Table 24. Two functions of each discriminant analysis for the writing tasks | 187 |
| Table 25. Classification results for the writing tasks with MLTTCU and STCU/TTCU as the predictors | 191 |
| Table 26. Two functions of each discriminant analysis for the writing tasks | 192 |
| Table 27. Comparison among (69), (70), and (71) on four Chinese syntactic complexity measures | 221 |
| Table 28. Comparison among (72), (73), and (74) on four Chinese syntactic complexity measures | 225 |
| Table 29. Comparison among (75) and (76) on four Chinese syntactic complexity measures | 231 |
| Table 30. Sentence provided for the GR task in TW&ST and their English translations | 241 |
| Table 31. Question list in the retrospective survey and coding of the answers | 207 |
| Table 32. Participants' retrospective perception on their performance and strategy in V1 and V2 | 208 |
| Table 33. Chinese Topic chain patterns | 297 |

LIST OF FIGURES

| | |
|---|-----|
| Figure 1. A taxonomy of complexity constructs. | 15 |
| Figure 2. The measurement process. | 17 |
| Figure 3. Properties relevant to the cline of clause combining. | 78 |
| Figure 4. Taxonomy of clause complexing. | 81 |
| Figure 5. Illustration of the structure of Telescope Chain (54). | 89 |
| Figure 6. Taxonomy of TC-units. | 108 |
| Figure 7. Screenshot of one slide in the TW&ST. | 133 |
| Figure 8. Mandarin EI score distribution of L2 Chinese speaker participants. | 153 |
| Figure 9. The mean values for each measure on five tasks by proficiency level. | 163 |
| Figure 10. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the CS task. | 173 |
| Figure 11. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the V1 task. | 175 |
| Figure 12. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the V2 task. | 177 |
| Figure 13. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the FW task. | 183 |
| Figure 14. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the GR task. | 185 |
| Figure 15. Predicting proficiency groups by measure MLTTCU and STCU/TTCU: Cases and group centroids for two discriminant functions of the FW task. | 189 |
| Figure 16. Predicting proficiency groups by measure MLTTCU and STCU/TTCU: Cases and group centroids for two discriminant functions of the GR task. | 190 |
| Figure 17. Estimated marginal means of four complexity measures on three tasks by proficiency groups. | 204 |
| Figure 18. Estimated marginal means of four complexity measures on three tasks by proficiency groups. | 213 |

LIST OF ABBREVIATIONS AND SYMBOLS

ACTFL: The American Council on the Teaching of Foreign Languages

CAF: Complexity, Accuracy, and Fluency

CL: classifier

CS: Cartoon strip

CTTCU/ATTCU: Complex Terminable TC-unit/All the Terminable TC-units (both simple and complex)

DE: Chinese modifier marker

EI: Elicited Imitation

FW: Free writing

GR: Guided rewriting

ID: Item Discrimination

ILR: Interagency Language Roundtable

L1: First Languages

L2: Second (or foreign; also third, fourth, and so on) Language

MLTTCU: Mean Length of Terminable TC-unit

MLSTCU: Mean Length of Single TC-unit (both independent and dependent)

NP: A noun phrase or nominal phrase

RQ: Research question

STCU/TTCU: Single TC-units (independent or dependent) per Terminable TC-unit (both simple and complex)

SVO: Subject–Verb–Object

T-unit: Terminable unit

TBLT: Task-Based Language Teaching

TC-unit: Topic-Comment unit

TTCU: Terminal Topic-Comment Unit

TW&ST: Chinese Timed Writing and Speaking Test

V1: The first time video retelling task

V2: The second time (immediately repeated) video retelling task

VP: A verb phrase

WTC: Willingness to Communicate

∅: Coreferential Zero

↕ : indicates a coreferential relation between two topics

↕ : indicates a coreferential relation between non-topic and a topic, each in one of the two consecutive topic-comment structures.

CHAPTER 1

INTRODUCTION

1.1 Motivation of the present study

The triad dimensions of complexity, accuracy, and fluency (CAF) has been widely used for assessing second language performance and development in the field of second language acquisition (SLA). Accuracy indicates the ability to produce target-like and error free language. Fluency shows the degree of automatization in accessing second language capability and is seen as the ability to produce the L2 with native-like rapidity, pausing, hesitation, and reformulation. Complexity reveals the scope of expanding or restructuring second language knowledge and is seen as the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2 (Ellis 2003, 2008; Ellis & Barkhuizen 2005; Lennon 1990, Skehan 1998; Wolfe-Quintero, Inagaki & Kim, 1998). However, besides these working definitions of CAF, there is a dire need for more clarification of the construct of CAF itself and a consistency in its operationalization to warrant the validity and reliability of studies that rely on CAF as a measurement model.

Due to the lack of well-defined constructs and corresponding measures, the interpretability and generalization of studies on syntactic complexity can be questioned. A major threat to validity that occurs during behavior identification is construct underrepresentation (Norris & Ortega, 2003). A failure to completely identify the logically linked behaviors to a theoretical interpretation can cause a partial loss of what the theoretical construct taps into in the empirical operationalized endeavor. Such construct underrepresentation can lead to an incomplete understanding or even a

misunderstanding about the construct. What adds to such a threat is that such links, between the theoretical construct and its empirical operationalized endeavor, can be rigid and inorganic when the differences grounded in the behavioral evidence of different nature were not appropriately taken into account. In other words, the inadequacy in both the quantity and quality of the links between a theoretical interpretation and required behavioral evidence can threaten the validity of construct interpretation. For studies in language complexity, an organic and sustainable approach was advocated to investigate complexity via syntactic varied levels to warrant the adequacy in quantity of the links between a theoretical interpretation and required behavioral evidence (Norris & Ortega, 2009). In addition, such an organic approach should also be extended to warrant the adequacy in quality of such links. In L2 syntactic complexity analysis, the identified behavior for the theoretical definition may also be presumed or follow a tradition without consideration of the particular feature of the target language or the context of investigation. Therefore, the quality of the links between the conceptualization and operationalization of the Chinese syntactic complexity construct cannot be adequate if the syntagmatic mechanism typological difference of different families of languages is overlooked. Other possible forms of clause combining in addition to coordination and subordination may be included according to the typological differences existing in other families of languages.

Compared with accuracy and fluency, complexity development has not yet received sufficient attention in Chinese second language teaching and research. Chinese syntactic complexity development has not been comprehensibly conceptualized or operationalized. In the Chinese nationwide *Standards for Mandarin Chinese Proficiency* (Hanban, 1995),

there are specific descriptive and quantifiable requirements for accuracy and fluency with regards to listening, speaking, reading, and writing for each proficiency level, however, requirements regarding complexity development were overlooked. Perhaps in part this is due to the lack of a clear understanding regarding the nature of complexity development in Chinese or other typologically different languages. Thus, while current complexity measures were developed in a global fashion and presumed to apply cross-linguistically, their definition and operationalization was based on Indo-European languages primarily; little attention has been paid to tailoring our understandings of the construct to the typological differences of other languages. In Chinese, it may be the case that syntax functions distinctly, and that features like topic prominence of the Chinese language provide the basis for syntactic complexing. Therefore measures originally developed for Indo-European languages are not as valid as indices of Chinese syntactic complexity. For instance, it was found that the mean length of the T-unit of native Chinese speakers is shorter than that of L2 speakers (Jin, 2006; Yuan, 2009).

In response, one potential solution of applying topic chain as the unit of analysis for Chinese syntactic complexity was originally proposed by Jin (2006). Jin (2006) for the first time applied a *Terminal Topic-comment Unit* (of which Jin abbreviated as TTCU) in Chinese syntactic complexity assessment. While Jin's visionary proposal pointed to the direction of a potential breakthrough, there still is a long journey ahead to define and assess Chinese syntactic complexity.

First, there is a lack of an accessible definition on the segmentation of the unit of analysis for Chinese syntactic complexity. While most of the currently available syntactic complexity measures are segmented by sentence, the Chinese sentence boundary is

arbitrary spelling thus not reliable in data coding. For the same Chinese text, different native Chinese speakers may provide very different punctuation marks which demarcate the sentences in Chinese (Tsao, 1990). In this sense, the validity and reliability of any Chinese syntactic complexity measures such as a T-unit that depend on the Chinese sentence boundary could be in question. When analyzing Chinese syntactic complexity, the very few currently available studies in Chinese complexity noted the existing punctuation marks by the author as the sentence boundary indicator by default and bypassed the problem of subjectivity of sentence segmentation (Jin, 2006; Jiang, 2013) or did not specify sentence-level segmentation (Yuan, 2009). For TTCU proposed by Jin (2006), there was not a clear segmentation criterion, however, the boundaries of the exemplified TTCUs provided were all overlapped with the sentence boundaries indicated by punctuation marks. Without clearly defining the beginning and end points of such topic chain based unit, its application is limited and debatable.

Second, a clarified and comprehensive conceptualization of Chinese syntactic complexity analyzed in the unit of topic chain is lacking. The TTCU in Jin (2006) was not consistently applied to analyze all the written output, instead, only a limited amount of topic chains was identified as TTCUs out of all the written output collected. Such partial application of topic chains in data analysis was due to the lack of clarification on the number and the nature of topic chain's components as well as the number and the nature of their constituent relationship. The relationship between the topic chains and non-chain topic-comment structures needs to be coordinated, that way such topic-prominent unit can be applied to consistently analyze all the Chinese output instead of the partial output.

Third, a study that systemically investigates on the validity and reliability of measures based on topic chain in assessing Chinese syntactic complexity has not been performed. A series of one-at-a-time t-tests was conducted in Jin (2006) between each L2 Chinese speaker group and the native Chinese speaker group on each Chinese syntactic measures. In her study, L2 Chinese speaker groups of varied Chinese language proficiency (Group Intermediate, Group Intermediate-High, and Group High) were compared with Group Native speakers on each Chinese syntactic complexity measure (See Section 3.3). There was no comparison done among the three L2 Chinese speaker groups to provide more insight on the validity of TTCU-based measures.

Lastly, more variety of task designs in terms of task type and cognitive complexity shall be included in order to elicit substantial Chinese spoken and written output for Chinese syntactic complexity analysis. Two guided rewriting tasks were used to elicit written Chinese output in Jin (2006). The participants were given instructions to rewrite two passages for better sentence structures. They were also told they “may manipulate the sentences, change the order of words, and omit words, but try not to leave out any of the information” (Jin, 2006, p.139-140). One passage was presented in the form of six VPs with no punctuation marks. These six VPs were each presented on individual sequential screen on the computer. The second passage was more extended. It was a semantically coherent but formally incohesive passage, with sentences presented in groups on sequenced computer screens. To ensure every participant comprehend the provided Chinese texts, the VPs and sentences in both guided rewriting tasks were presented with both pictorial cues and English translation. A free writing task in the form of a letter writing was applied in Jiang (2013), in which the genre of letter writing was

conformed to inviting a friend to a dinner party. However, different instructions and requirements, such as different must-include information and different length of the letter (measured as the number of characters), were required for participants of different Chinese proficiency levels. Provided with different instructions and requirements which vary the complexity and difficulty of task, the comparability of the task products can still be questioned. In addition to the task types, in order to elicit language output triggered by a varied cognitive task complexity, variables operated in the resource-dispersing dimension and resource-directing dimension should also be taken into account for task design.

This dissertation, based on these previous research, aims to provide comprehensive insights into conceptualizing and operationalizing Chinese syntactic complexity in terms of topic chain. Topic chain is taken as one way of clause complexing in addition to coordination and subordination. This dissertation proposes a taxonomy of Topic-Comment Units (TC-units) to examine Chinese syntactic complexity as the number and the nature of a terminable TC-unit's components as well as the number and the nature of their constituent relationship. Based on the taxonomy of TC-units, a series of TC-unit based measures can then be proposed and checked in a carefully developed measurement approach. Utilizing such validated measures, the interaction between task cognitive complexity and Chinese syntactic complexity can then be tapped into.

1.2 Research Outline

This dissertation consists of six main parts. It will begin by clarifying the conceptualized definition of syntactic complexity. Extending past its most widely agreed

working definition and taking into account its multifaceted traits, this dissertation examines syntactic complexity as summarized by Bulté and Housen (2012, p. 22) — that is, complexity is conceptualized as the number and the nature of the discrete components that the entity consists of, and the number and the nature of the relationships between the constituent components. Via a contrastive analysis between English and Chinese on syntagmatic mechanisms, this dissertation reviews Chinese typological features in terms of topic-prominent, parataxis-prominent, and discourse-oriented dimensions, in contrast to the subject-prominent, hypotaxis-prominent, and sentence-oriented features of English. Instead of analyzing the Chinese syntactic structure in terms of coordination and subordination, as widely used in Indo-European languages, this dissertation, for the first time in this line of research, proposes to include topic chain into the taxonomy of clause complexing (as shown in Figure 4 in Section 2.2.2.3) and to use it as the primary unit for Chinese syntactic complexity analysis.

Second, conceptualizing topic chain as the primary clause complexing mechanism for Chinese syntactic complexity, this dissertation proposes a taxonomy of Topic-Comment units (TC-unit) for Chinese syntactic complexity analysis, with the constituent components and the relationships between the units illustrated in Figure 6 in Section 3.3. In such a taxonomy of TC-units, a *terminable TC-unit* is the umbrella unit of analysis. A *terminable TC-unit* refers to a topic chain or a single independent topic-comment structure of which the topic was not repeated in the preceding or subsequent topic-comment structure. The terminable TC-units subsume simple terminable TC-units and complex terminable TC-units. A *simple terminable TC-unit* refers to a terminable TC-unit consisting of only one independent single topic-comment structure. A *complex*

terminable TC-unit refers to a terminable TC-unit takes the forms of a topic chain consisting of two or more dependent single TC-units. The dependent single TC-units in a complex terminable TC-unit are connected via coreferential zeros to form various types of topic chains. A *coreferential zero* refers to an element that does not have any phonological content and is unpronounced but corefers to the topic mentioned in preceding or subsequent clauses. Whenever a topic is not repeated in the form of a *coreferential zero*, a new terminable TC-unit is then activated. By applying coreferential zero instead of relying on the intonation marks in Chinese language output segmentation, this dissertation explores a possibility for more reliable analysis of Chinese syntactic complexity in the form of both written and spoken language output.

With a terminable TC-unit comprehensibly defined and operated as the unit of analysis on Chinese syntactic complexity, this dissertation proposes a series of indices based on the terminable TC-unit to measure Chinese syntactic complexity at various syntactic levels using an organic and sustainable approach (See Table 4 in Section 3.3). Such TC-unit based Chinese syntactic complexity measures include: a) global complexity measures: Mean length of terminable TC-unit (MLTTCU), Complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU), and Ratio of different types of terminable TC-unit; b) clausal complexity measures: Mean length of single TC-unit (both independent and dependent) (MLSTCU), Single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU); c) subclausal/phrasal complexity measure Dependents per head; and d) specific form complexity measure Frequency of a specific form, etc.

Third, to investigate the validity of the above-proposed TC-unit based Chinese complexity measures, this dissertation employs a designed *Chinese Timed Writing and Speaking Test* (TW&ST) to elicit both written and spoken output from L1 and L2 Chinese speakers. In taking the TW&ST, all participants complete three speaking tasks and two writing tasks. The speaking tasks are: a comic strip description (CS) task, a video story retelling task (V1), and an immediate, repeated video story retelling task (V2). The writing tasks consist of a free writing task (FW) and a guided re-writing (GR) task. Additionally, L2 Chinese speaker participants complete a Mandarin elicited imitation (EI) test (Zhou & Wu, 2009) in order to attain an assessment of their global Chinese proficiency level. Four of the TC-unit based measures proposed in this dissertation (as shown in Table 4 in Section 3.3) are applied in order to code the elicited complete spoken ($N=115$) and written ($N=116$) data set. These four measures are: ① mean length of terminable TC-unit (MLTTCU), ② complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU), ③ mean length of single TC-unit (both independent and dependent) (MLSTCU), and ④ single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU). To further investigate the validity of the proposed measures, this dissertation then conducts discriminant function analyses, correlating the participants' Chinese proficiency level with their syntactic complexity level. Comparing and utilizing all four measures or different combinations of the four measures as predictors, this dissertation shows that with high efficiency the MLTTCU itself can be chosen as one of the most valid measures for spoken Chinese syntactic complexity, and for written Chinese, applying both

MLTTCU and STCU/TTCU measures as predictor variables generates more accurate group membership classification.

Fourth, this dissertation will further look into how Chinese syntactic complexity is developed along with global proficiency development, as well as how a TC-unit per se develops. The corresponding growth of Chinese syntactic complexity along with an increase in Chinese proficiency confirms and amends the three stages of Chinese complexity development as outlined and described by Jin (2006) (translated as): *threshold, growth, and leap*. In both speaking and writing tasks, generally, participants of higher Chinese proficiency produce longer terminable TC-units with more dependent single TC-units, while participants of lower Chinese proficiency produce shorter terminable TC-units with less dependent single TC-units. Checking the quantitative statistics against qualitative developmental features, the increase in the Chinese syntactic complexity development along proficiency increase showed a transitional reliance on single TC-unit lengthening to single TC-unit combining, which is confirmed with the different patterns of correlation at global and clausal complexity level between participants' Chinese proficiency scores, and their scores on both the length and ratio syntactic complexity measures. With further depiction of Chinese syntactic complexity development, this dissertation also provides suggested descriptive requirements for language proficiency guidelines in terms of syntactic complexity. In addition, utilizing the proposed TC-unit based Chinese syntactic complexity measures, the interaction between presumed cognitive task complexity and Chinese syntactic complexity will also be explored by conducting repeated measures analyses. Along the resource-directing dimension of cognitive task complexity, a repeated measures analysis shows that higher

language complexity was produced in the tasks of higher cognitive complexity. Along the resource-dispersing dimension of cognitive task complexity, the findings suggest that immediate task repetition can lower learners' communication anxiety and increase positive self-perception of their own performance regardless if the actual language complexity did not show any clear increase along with task repetition.

Fifth, by providing a clearer picture of *what* Chinese syntactic complexity is and *how* it is developed, this dissertation provides pedagogical implications from both a macro as well as a micro perspective for developing Chinese syntactic complexity in Chinese as a second language teaching and learning. At a macro level, it is hoped that the complexity dimension via TC-units may be introduced into Chinese language learning and assessing with corresponding teaching emphases integrated alongside the three Chinese complexity development stages. At a micro level, this dissertation also suggests both individual topic chain composing steps as well as classroom teaching cycle design with sample task designs. Some samples of classroom teaching task designs provided include: (i) extending, (ii) sequencing, (iii) combining, (iv) inserting, (v) chain forming, (vi) punctuation marking, (vii) conjunction converting, and (viii) translating.

Last, this dissertation concludes by considering the contributions of the present study as well as the limitation of the study, and by making suggestions for future research into syntactic complexity measurement and instructed development in non-Indo-European languages.

CHAPTER 2

DEFINING CHINESE SYNTACTIC COMPLEXITY

2.1 The construct of complexity

2.1.1 *The CAF model in SLA*

In the field of SLA, second language development and performance has been assessed via different dimensions. In the 1980's, as the emphasis of second language learning and teaching shifted from fostering knowledge about the language to more communicative competence, high levels of and dependence on accuracy could no longer meet all expectations. Fluency was thus included to form a dichotomy with accuracy, where a distinction between the two was initially proposed to separate foci of L2 classroom teaching (Brumfit, 1984). After debate over foci in terms of purpose, activity design, material development, and feedback type and timing, among others, a consensus was reached that accuracy and fluency are both important goals to pursue in communicative language teaching (Brown, 2001). In addition to accuracy and fluency for assessing L2 language performance and development, another trend of including grammatical complexity in addition to accuracy arose and was shown to be reliable in L1 acquisition research (Larsen-Freeman, 1978). With these two dichotomies presented, a triad model of accuracy, complexity and fluency was first introduced, when addressing the pedagogic goals for task-based approaches, to lead learners not only “to the capacity to be an effective communicative problem solver but also to longer-term linguistic development” (Skehan, 1996, p. 21). The most widely agreed understanding about CAF in SLA is its working definition. Accuracy compares interlanguage with target language

norms and is seen as the ability to produce target-like and error free language. Fluency shows the degree of automatization in accessing second language capability and is seen as the ability to produce the L2 with native-like rapidity, pausing, hesitation, and reformulation. Complexity reveals the scope of expanding or restructuring second language knowledge and is seen as the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2 (Ellis 2003, 2008; Ellis & Barkhuizen 2005; Lennon 1990, Skehan 1998; Wolfe-Quintero, Inagaki & Kim, 1998).

These three dimensions have been operationalized in diversified indices for assessing L2 language performance. For the measurement of grammatical and lexical complexity alone, an inventory of forty different indices was observed in a survey by Bulté and Housen (2012) on forty empirical L2 studies on task-based language learning published between 1995 and 2008, not to mention more indices should accuracy and fluency measurement be included. In addition to such abundant yet inconsistent practice at the operational level, CAF also lacks clarity and depth in its definition and component identification at the theoretical level. Housen, Kuiken and Vedder (2012) raised a concern regarding the validity and reliability of studies applying CAF measures without explicitly defining the construct they are measuring.

Many L2 studies that investigate CAF either do not explicitly define what they mean by these terms, or when they do, they do so in rather general and vague terms (e.g. ‘fluency refers to the ease with which learners produce the L2’) or in terms of concrete psychometric instruments and quantitative metrics (e.g. ‘complexity refers to the extent to which the learners use syntactic embedding and subordinate clauses, relative to the

total number of clause produced'). As a result, the terms 'complexity', 'accuracy' and 'fluency' are often used with different meaning across studies (and sometimes also within studies). This limits the interpretation and comparability of CAF findings and may also explain why the CAF literature has produced many inconsistent findings (Housen & Kuiken, 2009; Norris & Ortega, 2009; Robinson, Cadierno, & Shirai, 2009). (p. 3)

Conducting studies without carefully defining and operationalizing the construct can cause delusive inconsistency in research findings and limited contribution to the accumulation of knowledge. For example, the testing of two rival models of task complexity as they affect language performance has drawn a lot attention in task-based language teaching studies. The Limited Attentional Capacity Model (Skehan, 1998) and the Multiple Resources Attentional Model (Robinson, 2001, 2005) conflict in whether L2 language complexity goes up along an increase of cognitive task complexity at the price of lower L2 accuracy, or whether L2 language complexity and accuracy actually go up simultaneously. As Housen et al. (2012, p. 6) concluded, the empirical evidence available so far does not equivocally support either model in part because of the lack of conceptual and operational clarity of the dependent variables of CAF. Therefore, there is a dire need for more clarification of the construct of CAF itself and a consistency in its operationalization to warrant research validity and reliability that rely on CAF as a measuring model.

2.1.2 Complexity in the triad of CAF

Complexity has been operationalized with diversified indices for L2 performance and proficiency assessment. However, as the youngest dimension included to the triad of CAF, it is the most underexplored. “As befits the term, *complexity* is the most complex, ambiguous and least understood dimension of the CAF triad” (Housen & Kuiken, 2009, p. 464). The construct of complexity is problematic because of its polysemous nature (Pallotti, 2009). It is used as a dimension of L2 language performance. The same word “complexity” is also used in task design as a criterion for cognitive sequencing. Cognitive factors can be manipulated to decrease or increase the task complexity. As is shown in the Figure 1, Bulté and Housen (2012, p. 23) differentiated two types of “complexity”: relative complexity and absolute complexity. Relative complexity was defined in relation to language users as difficulty or cognitive complexity, where “a language feature or system of features is seen as complex if it is somehow costly or taxing for language users

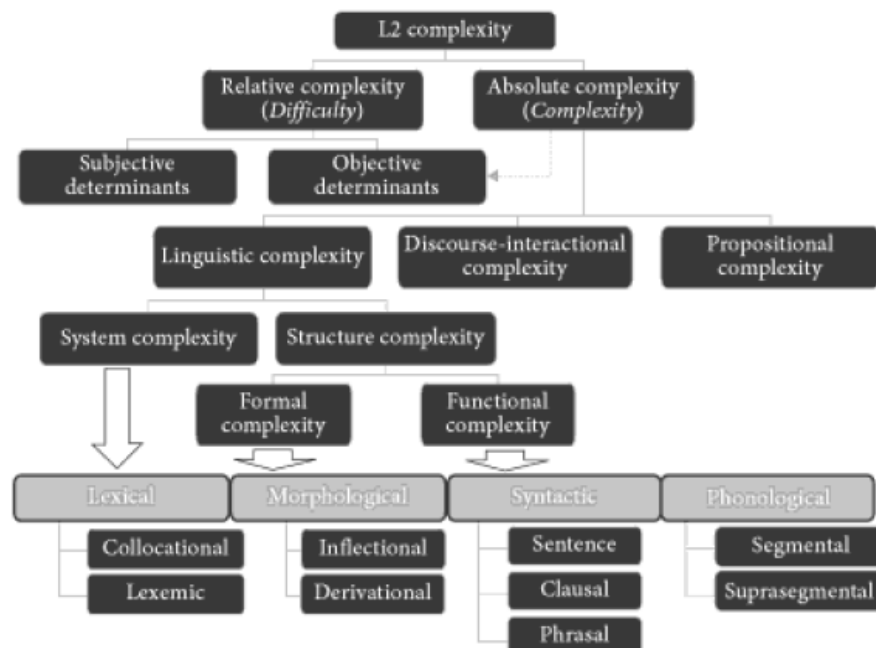


Figure 1. A taxonomy of complexity constructs (from Bulté & Housen, 2012, p. 23).

and learners, particularly in terms of the mental effort or resources that they have to invest in processing or internalizing the feature(s)". Absolute complexity was given as the dimension in the triad of CAF. Absolute complexity itself, however, is a multidimensional, multilayer, and multifaceted construct. Various perspectives and levels of the language are required to analyze complexity. Discourse-interactional complexity is mainly about learners' dialogic discourse in terms of the number and type of turn changes. Propositional complexity refers to the density of information encoded in a certain language act. Linguistic complexity subsumes phonological, morphological, lexical, and syntactic complexity. This dissertation will focus on syntactic complexity concerned at phrasal, clausal, and sentential levels.

Overall, there have been diversified practices at the operational level yet not much literature carefully defining the construct of complexity at the conceptual level prior to its operationalization. In the pursuit to muster disciplinary consensus regarding the design, interpretation, and report of assessments within SLA, Norris and Ortega (2012, p. 574) summarized the concerns of assessment method choice making in SLA: "(a) *what* gets assessed, or the L2 knowledge constructs researchers want to know about; (b) *how* to assess, or the ways of eliciting and analyzing phenomena related to these constructs; and (c) *who* gets assessed and why, or the clearly specified learners and populations that researchers investigate and the explicitly considered purposes for assessing them." In the following sections, this paper will try to address the concerns about "*what*" and "*how*" in L2 syntactic complexity assessment, and discuss "*who*" in part in terms of measuring Chinese syntactic complexity for adult L2 Chinese speakers.

2.1.3 Conceptualizing and operationalizing syntactic complexity

Two primary challenges in L2 learner knowledge assessment were called to attention in Norris and Ortega (2012, p. 574): “(a) defining the specific L2 knowledge constructs of interest from their theoretical perspective, and (b) proceduralizing data collection through assessment such that interpretable light is shed upon them.” Following the framework for understanding and executing measurement in SLA research as illustrated in Figure 2 (Norris & Ortega, 2003), this dissertation will try to address these two challenges in syntactic complexity assessment via two parts, conceptualization and proceduralization, of which *construct definition, behavior identification, task specification, behavior elicitation, observation scoring, and data analysis* are six inherent cycling categories.

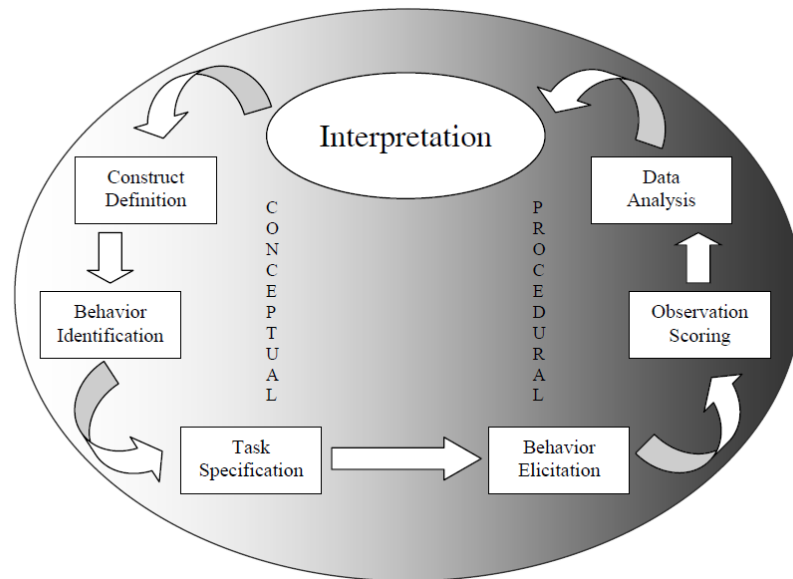


Figure 2. The measurement process (from Norris & Ortega, 2003, p. 720).

2.1.3.1 Defining syntactic complexity

Second language acquisition starts from initial noticing and awareness of form-meaning connections. Input turns into intake, and is then pushed out as interlanguage output. Initial chunked output is sided along with expanding, restructuring, and automating, whereupon L2 development reaches the stage of fluent, accurate, and complex language performance. As a triad model detecting L2 language proficiency and performance, complexity, accuracy, and fluency are interdependent and interacted constructs. The construct of complexity is thus conceptually defined as part of the whole picture as CAF captures the stages of L2 development. Housen et al. (2012) brought about the following review:

Theoretically, these three dimensions have been claimed to imply the major stages of change in the underlying L2 system: (i) internalization of new L2 elements (or greater complexity, as more elaborate and more sophisticated L2 knowledge systems are developed); (ii) modification of L2 knowledge (as learners restructure and fine-tune their L2 knowledge, including the deviant or non-targetlike aspects of their interlanguage (IL) so that they become not only more complex but also more accurate L2 users); (iii) consolidation and proceduralisation of L2 knowledge (i.e. higher fluency, through routinisation, lexicalization, and automatising of L2 elements leading to great performance control over the L2 system; De Graaff & Housen, 2009; Skehen, 1998, 2003). (p. 3)

Syntactic complexity is partly related to grammatical diversity that refers to the elaboration, size, range, and variation of L2 elements. Internalization of the diversified,

elaborated, and varied L2 elements is related to stage (i) above, the internalization of new L2 elements, as the breadth of the syntactic complexity construct. Complexity as a multidimensional construct itself, however, has at least both dimensions of breadth and depth. In addition to breadth, the depth of syntactic complexity covers how to assemble and restructure these L2 elements in the form of denser and coherent L2 form. This is to say, syntactic complexity depth refers to the embeddedness and compositionality of grammatical L2 structures. “The clause complex ... represents the dynamic potential of the system — the ability to ‘choreograph’ very long and intricate patterns of semantic movement while maintaining a continuous flow of discourse that is coherent without being constructional” (Halliday, 1985, p. 202).

From a cognitive-interactionist SLA perspective, complexity is viewed as the interplay between declarative knowledge, or explicit knowledge, and procedural knowledge, or implicit knowledge (Towell & Hawkins, 1994; Wolfe-Quintero et al., 1998). Declarative and explicit knowledge is about what something is while procedural and implicit knowledge is about how to do something. Explicit L2 knowledge is conscious, declarative, actively controlled, verbalized, and learnable, whereas implicit L2 knowledge is intuitive, procedural, automatically accessible, behavioral but not verbalizable, and developmentally constrained (Ellis, 2004; Ellis et al., 2009). Declarative knowledge enables a L2 learner to describe a rule and complete a language knowledge test by applying it. Procedural knowledge, on the other hand, enables the L2 learner to actually apply that rule in real language use. A native speaker can have high levels of procedural knowledge in terms of speaking a perfect L1 but has no declarative knowledge in terms of knowing about the grammar. On the other hand, a L2

speaker may have full declarative knowledge of the lexis, syntax, and discourse of the target language but not be able to apply such knowledge in language use as procedural knowledge. Declarative knowledge is related to the expansion of L2 knowledge by internalizing L2 elements like lexis, expressions, grammar, and rules for restructuring L2 elements. Procedural knowledge helps the learners directly using these L2 elements in breadth and applying the acquired composing rules in depth in real language use. In this way L2 speakers internalize knowledge of syntactic complexity both in terms of breadth and depth.

This dissertation adopts the conceptualized definition of syntactic complexity concluded by Bulté and Housen (2012): “at the most basic level, complexity refers to a property or quality of a phenomenon or entity in terms of (a) the number and the nature of the discrete components that the entity consists of, and (b) the number and the nature of the relationships between the constituent components” (p. 22). Such two folds correspond to the breadth and depth of syntactic complexity. The first fold can be traced back to its etymological Latin origin. According to the *New Oxford North American Dictionary*, complexity was defined as “consisting of many different and connected parts”. The number and the nature of the discrete L2 components is about breadth of syntactic complexity, which is about grammatical diversity that refers to the elaboration, size, range, and variation of L2 elements. However, not only are the “parts” of L2 different, but they are also connected in different ways. In the second fold, the number and the nature of the relationships between the discrete L2 components corresponds to the depth of syntactic complexity, which is about embeddedness and compositionality of grammatical L2 structures. Therefore, complexity is first about discrete L2 components

and the mechanisms that connect these components. Second, complexity entails both quality and quantity features of language components and their composing mechanisms. However, given varied typological features, the quality and quantity features of discrete language components as well as the relationship of embeddedness and compositionality of such language components may vary greatly. Therefore, one complexity measure works for this group of languages might not be acutely applicable for another group of languages. In order to increase L2 syntactic complexity, learners need to not only expand their declarative knowledge by internalizing L2 elements like lexis, expressions, grammar, and rules for restructuring L2 elements, but also elevate procedural knowledge to directly use these L2 elements in breadth and apply the acquired composing rules in L2 communication.

2.1.3.2 Identifying behavior of syntactic complexity

Provided the construct definition, as shown in Figure 2, particular behavior or constellations of behaviors with the qualities or variations is searched in order to show sufficient information for a complete construct interpretation. From different syntactic levels of a target language and the corresponding constituent nature at these levels, however, the behavior of this construct in L2 performance can be identified from two perspectives, corresponding to the construct's breadth and depth dimensions. At the breath dimension, the behavior identified refers to grammatical expansion: lengthwise global/sentential complexity, clausal complexity, and phrasal complexity, among others. At the depth dimension, the behavior identified refers to grammatical structuring: head modifying, clause embedding, and clause combining, among others. As the construct of

syntactic complexity is conceptualized in terms of the number and the nature of the discrete components as well as their constituent relationship, language output of different component constituents in different relationships are searched for analysis. Such language output may be collected from L2 speakers of different proficiency levels on different topics in varied contexts. It can be written and/or spoken output of different languages with different linguistic features.

A major threat to validity that occurs during behavior identification is construct underrepresentation. “Construct underrepresentation occurs when the complex link between a theoretical interpretation and required behavioral evidence is inadequately understood and/or conveyed into practice.” (Norris & Ortega, 2009, p. 729) What we are measuring may not necessarily sufficiently reflect what the theoretical construct targets. Therefore, what the construct can tap into might be partially lost in the empirical operationalized endeavor due to a failure to completely identify the logically linked behaviors. Such under-investigation can cause incomplete understanding or even misunderstanding about the construct. What adds to such a threat is that such a link can be rigid or inorganic when the differences grounded in the behavioral evidence of a different nature were not appropriately taken into account. In other words, not only the inadequacy in quantity but also the quality of the links between a theoretical interpretation and required behavioral evidence threatens the validity of construct interpretation. In L2 syntactic complexity analysis, the behavior identified for the theoretical definition may also be presumed or it may follow a tradition without consideration for the particular feature of the target language or the context of investigation. In this dissertation, when defining and measuring Chinese syntactic

complexity the grounding differences lie in the different language typological features. Therefore, in this dissertation, employing the notion of *GlobaLocality*, a top-down approach that starts from clarifying the theoretical construct of syntactic complexity is mixed with a bottom-up approach that takes a particular feature of the investigated language.

2.1.3.3 Task specification for syntactic complexity output

With the target behavior identified as spoken and written discourse production, tasks that can generalize spoken and written output are desirable. Those most commonly used in existing literature can be mainly categorized into tasks of free writing and speaking, and structured writing and speaking. Free writing and speaking tasks collect segments of any written or spoken output by the participant without any control on the topic, context, or time. To better control output, Hunt (1970, 1977) started applying rewriting to elicit written output from students of different ages. Such method of rewriting was more controlled by providing students with a passage written in extremely short sentences and asking them to rewrite the passage in a better way. Another way of eliciting structured writing was to provide a series of semantically coherent but formally incohesive sentences and ask the students to revise the sentences to be better structured (Jin, 2006). A third way of eliciting structured speaking and writing was to provide students with non-language prompts such as comic strips and to ask them to describe or narrate the non-language prompts (Kormos, 2011). When syntactic complexity is tapped as an independent variable, the tasks eliciting discourse production should be of the same design for different data sources from any subject group. The same task or same series of

tasks should be applied to participants of different Chinese proficiency levels. In this way, the topic and complexity of tasks, as well as the genre of the targeted language output are all under control. Such an output is then comparable via proficiency groups since the output is elicited under the same condition except participants' varied Chinese proficiency levels. Different syntactic complexity measures can then be applied to such an output to show how different language complexity is produced by participants of varied proficiency levels. Applying different tasks tailored to L2 speakers' corresponding proficiency levels will contrastively result in incomparability between the outputs of speakers of different proficiency levels. Therefore, the difficulty and complexity of such a series of tasks should be well manipulated to cover the variability of the speakers of different proficiency levels. In other words, the tasks should be at least entriable for the lower-level speakers as well as leaving sufficient space for the higher-level speakers to produce more complex language output. Too low difficulty and complexity of a task leads to a possible ceiling effect for higher proficiency speakers, while too high difficulty and complexity of a task leads to a possible floor effect for lower proficiency speakers.

2.1.3.4 Behavior elicitation for syntactic complexity output

When applying particular tasks in the elicitation of targeted behaviors, other variables that may affect the observed or recorded behaviors should be carefully controlled and accounted for. When conducting tasks to elicit targeted behaviors, those tasks actually administrated may not turn out to be exactly as planned due to the condition of having to conduct a task. That is to say, task-as-process may turn out to be not the same as task-as-workplan. For instance, engaging in pre-task planning has been

observed to help adult L2 speakers across contexts and languages achieve higher levels of fluency and linguistic complexity during the actual task performance (Crookes, 1989; Foster & Skehan, 1996; Yuan & Ellis, 2003; Ortega, 1999, 2005). Therefore, controlling the pre-task planning time as well as the task performing time across the proficiency groups can help diminish other factors, other than participants' language proficiency, that might contribute to language complexity performance. When collecting output from L2 speakers, participants' biographic information including their age, gender, language background, learning experience and relevant factors should be recorded. For instance, while it is common for studies to use students' class standing and institutional status as a grouping variable, heritage speaking background of the target language that can cause a profile difference in a learner's L2 performance compared with other foreign language learners of the same educational level (Kondo-Brown, 2005). A global proficiency test of high reliability can be used as the grouping variable instead of L2 speakers' class standing and institutional status.

2.1.3.5 Observation scoring on syntactic complexity output

With a construct not carefully defined, subsequent scoring in the observation may become mismatched. Based on an in-depth and synthesizing review of current practices and discussion of the constructs within syntactic complexity, Norris and Ortega (2009) advocated an organic and sustainable approach to investigate CAF in SLA. Three levels of complexity are subsumed: global complexity, clausal complexity, and subclausal complexity. Global complexity can be measured via the mean length of a potential multiple-clausal unit of production. Clausal complexity detects the number of

subordination or coordination in a unit of upper-clausal level. Subclausal complexity via phrasal elaboration is measured by mean clausal length. In addition, it is suggested that the variety, sophistication, and acquisitional timing of the form also be taken into consideration. With such a framework providing a guideline for measuring syntactic complexity at different levels, a comprehensive investigation of complexity can be achieved via different types of indices.

In the field of L2 complexity measurement research, length-based and ratio-based indices are the two most commonly used types of measures. Length-based indices refer to words or morphemes including mean length of utterance, mean length of T-unit, mean length of c-unit, mean length of AS-unit, mean length of clause, and others (See Section 3.2 for more discussion). Length-based indices are considered generic indices since they simultaneously tap into different layers of syntactic structure and different sources of complexity – phrasal, clausal and sentential (Norris & Ortega, 2009). Ratio-based indices assign different weights to different syntactic structures as putatively different degrees of complexity. In SLA, available ratio-based indices include coordinated clauses/total clauses, total clauses/T-unit, total clauses/c-unit, total clauses/AS-unit, subordinate clauses/total clauses, relative clauses/T-unit, syntactic arguments/clause, and dependents/(noun, verb) phrase, among others. Yet, when applying length-based or ratio-based indices measures in scoring, reliability is the major concern. Norris and Ortega (2003) pointed out that reliability and error in measurement scoring is “at best infrequently considered and only inconsistently reported” (p. 745). Considering the gap between what researchers want to argue and what the measures are actually measuring, further articulation of the measurement application and reliability is needed. The

accumulation of trustworthy knowledge can hardly be achieved if no reliability is reported.

One potential problem for such measures is that they simultaneously tap into several subcomponents and subdomains of complexity. For example, measuring the length of T-unit is aiming at global complexity. However, length-based type indices are measured in terms of a word or morpheme, which simultaneously includes phrasal complexity. It's not that one measure has to purely detect one particular level. However, "...it is important that we motivate our complexity measures by stating what particular type, component or sub-construct, of complexity they represent and, in the case of hybrid and generic measures, by explaining how the different measures for one conglomerate complexity construct interact" (Bulté & Housen, 2012, p. 36).

2.1.3.6 Data analysis on syntactic complexity output

At the next stage, data analysis, individual scores are summarized, categorized, and compared for statistical analysis. The selection of statistical approaches should be selected based on the research questions and the methods of the study. An appropriate statistical approach shall be sufficiently sensitive as it leads to results that best reveal the nature of the target construct.

The general hypothesis underlying complexity development is that the higher one's proficiency level is, the more complex language one is able to produce. Therefore, one way of validating complexity measures is to correlate measure-based results with the proficiency level of the target language speakers. However, complexity as a multi-dimensional construct may not linearly correlate with proficiency level or time for

acquisition, and there may also be interaction among the subsystems of complexity (Pallotti, 2009; Larsen-Freeman, 2006; Verspoor, Lowie & Van Dijk, 2008). To achieve a more comprehensive investigation of a multidimensional construct, an organic and sustainable approach should also adopt more qualitative analysis on the longitudinal development of subsystems of complexity in line with L2 development.

2.2 Defining Chinese syntactic complexity

From a contrastive perspective, Chinese features many typological differences from Indo-European languages, English, for example, in terms of phonology, morphology, and syntax. Phonologically, Chinese is understood as a tonal language that employs pitch to distinguish meaning while English on the other hand does not have such tones. At the morphological level, world languages are categorized into four main types according to the different ways that morphemes compose words: isolating type, fusional type, agglutinating type, and polysynthetic type, with non-distinct types classified as mixed (O'Grady & Dobrovolsky, 1989). For languages of isolating type, a word consists of a single morpheme with no affixes. Chinese belongs to the isolating type that has no inflections, using instead word order and functional words as grammatical devices. Such lack of inflections causes difficulty and ambiguity when applying formal rules in analyzing Chinese syntax such as identifying part of speech. For fusional type languages, their word affixes and the word base to which they are attached are fused together in pronunciation as a result of phonological processes or change and therefore are not easily separated from one another. Most Indo-European languages like English are categorized as fusional language. In addition, there is generally a fusion of meanings that is

represented by the affixes in such languages. Agglutinating languages are languages in which affixes can easily be separated from the stems to which they are attached and in which each affix generally conveys only one meaning, like Korean and Japanese.

Polysynthetic languages are languages in which several stem forms may be combined (along with affixes) into a single word. Such a word is usually a verb with its associated nouns “built-in” or incorporated, so that verb alone expresses what seems to us about a whole sentence. Many native languages of North America have polysynthetic structures like Inuktitut, Cree, and Sarcee. Many languages are listed as mixed type since they do not exclusively belong to any of the aforementioned categories.

In terms of syntax, English is categorized as a SVO language based on a base order of subject (S)–verb (V)–object (O). For the most part, Chinese is also categorized as a SVO language, yet there is still debate due to its many variations from the prototypical SVO order. Many syntactic typological differences exist between these two languages. For instance, Chinese is a meaning-driven language that is coded directly corresponding to meaning, while at the other end, English is a form-driven language that is coded indirectly corresponding to meaning via form control (Pan, 2002). The Principles of Temporal Sequence and Temporal Scope (Tai, 1985, 1993, 2002) is one example of Chinese language mirrors a sequential order and scope of events. Second, English is also a subject-prominent language. An English sentence is formed with a structure of subject-predicate. An English sentence has one subject which is generally taken in the form of a noun/NP, and one predicate of which verb/VP is the essential part. In contrast, the subject-predicate structure does not perfectly fit to Chinese syntax analysis due to its lack of inflection. Mechanically applying the subject-predicate

structure in Chinese syntactic analysis has resulted in problems such as causing a large amount of trans-classed words, misplacing grammatical sentences as incomplete sentences, creating ambiguity in the central verb in serial verb constructions, and forming particular Chinese sentence patterns, etc. Such mechanical application has been criticized as the Indo-European lens in Chinese linguistics study (Zhu, 1985, 1994). Therefore, given many challenges caused by not applying the approach of *Globallocality*, Chinese linguistics studies is in dire need of researches from the standpoint of its own characteristics instead of seeing through the lens of Indo-European language studies (Zhu, 1985; Xu, 1991, 1997; Lu & Guo, 1998).

2.2.1 Topic-prominent versus subject-prominent

2.2.1.1 Topic-comment structure

Different languages feature different coding principles and strategies. Xu (1991, 1997) suggested that the Indo-European languages can be categorized as grammatical languages whereas Chinese belongs to a category of semantic language. While Indo-European languages like English are form-driven languages that are coded indirectly corresponding to meaning via form control, Chinese was found to be a meaning-driven language that is coded directly corresponding to meaning (Pan, 2002). A subject-predicate structure works well for an analysis of form-driven languages like English. However, mechanically applying a subject-predicate structure corresponding to parts of speech in Chinese syntactic analysis can result in a number of problems, such as causing a large amount of trans-classed words, misplacing grammatical sentences as incomplete sentences, creating ambiguity of the central verb in serial verb constructions, and leading

to redundant special sentence patterns. Rather, Chinese is found as a topic-prominent language. For Chinese, “the subject is literally the subject matter to talk about, and the predicate is what the speaker comments on when a subject is presented to be talked about” (Chao, 1968, p. 70). What makes a clause participant topical is not its grammatical status of subject or object, but rather, its thematic importance, recurrence, or continuity in discourse (Givón, 1992, p. 202). According to the prominence of the notions of topic and subject in the construction of sentences, Li and Thompson (1976) classified four categories after surveying thirty languages: (a) topic-prominent, e.g., Chinese, Lahu, and Lisu; (b) subject-prominent, e.g., English and most Indo-European languages, Dyirbal and Indonesian; (c) neither topic-prominent nor subject prominent, e.g. Tagalog and Ilocano; and (d) both subject- and topic-prominent, e.g., Japanese and Korean.

A subject-predicate form reflects the structure of an English sentence. This is to say, each English sentence is composed of one subject that is generally realized as a noun and one predicate, of which a verb is often the essential part. However, in Chinese this is not always the case. In addition to nouns and nominal phrases, verbs, adjectives, verbal phrases, and other such clauses can fulfill the subject position; and in addition to verbs and verbal phrases, nouns, nominal phrases, and adjectives can fulfill the predicate position. Different from subject-prominent languages like English, for Chinese language topic-comment structure is argued to be a better fit (Chao, 1968; Li & Thompson, 1976; Huang, 1984; Chu, 1998). The relationship of the topic and comment is “aboutness”.

Consider the following example sentence:

(1) *Jīntiān xīngqīwǔ*.¹

今天星期五。

[Today Friday.]

Today is Friday.

Sentence (1) represents a frequently used structure in Chinese, which is composed by only two nouns, *jīntiān* (今天, today) and *xīngqīwǔ* (星期五, Friday), and no verb. *Xīngqīwǔ* (星期五, Friday) is a comment about the topic *jīntiān* (今天, today). If we use a subject-predicate framework to analyze such a sentence, we have to assume that there is an omitted verb *shì* (是, is). However, the usage of *shì* (是, is) as in the assumed sentence *Jīntiān shì xīngqī wǔ* (今天是星期五。Today is Friday) only occurs in a comparative sentence when clarifying the day of the week, as in *Jīntiān shì xīngqī wǔ, búshì xīngqī liù* (今天是星期五, 不是星期六。It is Friday today but not Saturday).

2.2.1.2 Coreferential zero

When successive topic-comment structures are about the same topic, the topic may be overtly stated only once, with repeated mentions of the same topic in subsequent topic-comment structures realized as null phonological form, thus resulting in a topic

¹ Without specific indication, all the example sentences of native Chinese speakers in this dissertation are cited from the Peking University Center for Chinese Linguistics (CCL) Modern Chinese corpus: http://ccl.pku.edu.cn:8080/ccl_corpus/index.jsp?dir=xiandai Example sentences of L2 Chinese speakers are from the data collected by a Chinese language test completed by English speaking Chinese L2 learners. All the English example sentences in this paper are cited from “Corpus of Contemporary American English” which includes 450 million words, 1990-2012. <http://corpus.byu.edu/coca/>

chain. Sentence (2) below was an example in Xu (1991) illustrated a very typical topic chain:

(2) *Nàgǒu huáng máo, hēi yǎnquān, cháng shēncái, xì gāo tuǐ, tèbié de xiōngměng, yào yǎo zhù rén, bùjiàn diǎnr xuèxīng wèir, jué bù piēzuǐ.*

那狗_i黄毛, \emptyset_i 黑眼圈, \emptyset_i 长身材, \emptyset_i 细高腿, \emptyset_i 特别地凶猛, \emptyset_i 要咬住人, \emptyset_i 不见点儿血腥味儿, \emptyset_i 绝不撇嘴。

[That dog_i yellow haired, \emptyset_i black eye socket, \emptyset_i long body, \emptyset_i thin tall leg, \emptyset_i particularly DE ferocious, \emptyset_i once bite-achieved people, \emptyset_i no see little smell of blood, \emptyset_i never opens mouth.]

That yellow haired dog with black eye sockets is tall and has long legs. It is particularly DE ferocious. Once it bites someone, it will never let go until it draws blood. (p. 264)

In English, the subject of every uncoordinated clause should be stated to achieve subject-predicate consistency. When successive clauses or sentences share one same subject, the full form of the subject, pronoun, or demonstrative still has to be placed as the subject within each sentence, or conjunctions must be applied to connect the clause with the subject and the other clauses sharing the same subject. In Chinese, however, this is frequently not the case. In topic chain (2), *nàgǒu* (那狗, that dog) in the first clause functions as the topic. While the semantic subjects of the subsequent clauses above in (2) are all the same *nàgǒu* (那狗, that dog), there is no phonological form of this word in said clauses. These empty phonological shells connect the independent topic-comment structures into a topic chain via coreference to the same topic, realized not in audible

phonological form but in a deeper semantic stream. Such connection of the topic-comment structures is as well meaning-driven and coded directly corresponding to meaning. The order of these clauses follows the Principle of Temporal Sequence (Tai, 1985), which mirrors a sequential order of events, like a stream of consciousness. As such, we can understand a compound sentence in Chinese to be defined as a topic chain, where a sequence of clauses shares a single topic (Tsao, 1979; Li & Thompson, 1981; Tsao, 1990; Chu, 1998; Liu, 2004).

Zero anaphors were applied to treat reduced parts in Chinese conversation as null anaphoric forms that depend on antecedents for reference (Chen, 1989). These zero anaphors are phonetically null NPs by this definition. However, as discussed in Shi (1993), such reduction in Chinese is not limited to nouns, as other classes of word and sentence components can also be reduced and become zero anaphors within the context provided. As shown below (Shi, 1993), the intended reading for (3a) is what (3b) means, but both the verb *juān* (捐, to donate) and part of its object *màizi* (麦子, wheat) were reduced and became zero anaphors in the conversation. Without the context provided, the meaning of (3a) turns to be, I weight one-hundred-fifty catties, which diverges from the intended reading for (3a) as given in (3b). As Chinese is a meaning-driven language that is coded directly corresponding to meaning, in line with the occurrence of null phonological form at connecting topic-comment structures, the occurrence of zero anaphor is as well meaning-driven. Pu (1997) remarked that the higher the degree of thematic coherence, the higher the degree of referential accessibility, and therefore the more likely the occurrence of a zero anaphor.

(3) a. *Wǒ yìbǎi wǔshí jīn.*

我一百五十斤。

[I one hundred fifty catty (half-kilo).]

I weight one-hundred-fifty catties.

b. *Wǒ juān yìbǎi wǔshí jīn màizi.*

我捐一百五十斤麦子。

[I donate one-hundred-fifty-catty wheat.]

I donate one hundred and fifty catties (half-kilo) of wheat. (p. 314)

This dissertation hereby proposes *coreferential zero* to refer to an element that does not have any phonological form and is unpronounced but corefers to the topic mentioned in the preceding or subsequent clause(s) in a topic chain. Such clauses in Chinese are treated as single topic-comment structures to be consistent with topic chains. Coreferential zero is not restricted in conversations. Different from zero anaphor, such coreferential zero can be anaphoric or cataphoric reference. Coreferential zero does not have to be limited as nouns or nominal phrases, as they are used in complex topic chains to repeat any same topic with a phonetically null form.

In order to form a topic chain consisting of two or more topic-comment structures, successive topic-comment structures that share the same topic have to be connected with coreferential zero. Repeating the same topic via coreferential zero keeps the referent in active memory in a dense and coherent semantic flow in the form of a complex topic chain. If several successive topic-comment structures share the same topic however the topic is not repeated in the form of coreferential zero but the full forms, a pronoun, or a demonstrative, it is then not a topic chain but several sequenced independent topic-

comment structures. Whenever repetition of the same topic takes the form of its full form, a pronoun, or a demonstrative, a new topic-comment structure or topic chain is then activated. Whenever a different topic is introduced in its full form, a pronoun, or a demonstrative, or cataphorically repeated in coreferential zero, a new topic-comment structure or topic chain is as well activated. In English, lexical pronouns are mostly used to carry on active referents while zeros are only used with syntactic constraints (Tomlin, 1987, 1989). In Chinese, contrastively, zeros are extensively coded to keep up with the active referents while pronouns are reserved for coding minor discontinuities in discourse (Pu, 1997; Tomlin & Pu, 1991). Without being separated by phonologically repeated topics, connection via coreferential zero as a matter of fact reduces phonological distances between the successive comments, and thus leads to a higher density of semantic flow in the form of a complex topic chain.

The following two examples are provided to show the difference between a single topic-comment structure and a topic chain. The two examples (4) and (5) are transcription of the spoken output by two English speaking L2 Chinese speakers describing the same picture, (4) by a lower-proficiency speaker and (5) by a higher-proficiency speaker. The syntactic structure difference of (4) and (5) illustrates the difference in composing Chinese syntactic complexity. Though one same topic *tā* (她, she) is shared by the two sentences in (4), this is not a topic chain since the two topics are both in the form of a pronoun but not repeated in coreferential zero. The two sentences in (4) are thus less well-knit but more equal in isolation, therefore are considered two consecutive yet independent topic-comment structures, but not a terminable topic chain. Contrastively, the advanced L2 Chinese speaker efficiently employed coreferential zero to connect the

single topic-comment structures in (5). After the first mention, the topic *tā* (她, she) was repeated via coreferential zero seven times in the subsequent topic-comment structures. At the sentential or clausal level connection in Chinese, the way seen in (4) is commonly applied by lower level L2 Chinese learners of English speaking. With L1 transfer, full SVO structure of English is applied to every single Chinese sentence with each subject in the form of a full noun or nominal phrase, pronoun, or demonstrative. To connect successive topic-comment structures into topic chains requires appropriate topic-controlled deletion strategies. To establish a clear picture where coreferential zero should be placed in topic chains, categorization of topic chain patterns will be further discussed below in 2.2.1.3.

(4) *Tāxǐ gǒu. Tācā gǒu de tóu.*

她_i洗狗。她_j擦狗的头。

[She_i wash dog. She_j wipe dog DE head.]

She washed the dog. She wiped the dog's head.

(5) *Nǚde guì xiàláì, kāishǐ gěi xiǎo gǒu xǐzǎo, yòng shuāzi bǎ tā shuā de hěn xìxīn.*

Ránhòu, xǐ wán zǎo hòu, yòng máojīn bǎ xiǎo gǒu cā gān, ránhòu hái gěi tā shū máo.

Wánliǎo zhīhòu, yòu hǎoxiàng ránhòu duì zìjǐ hěn mǎnyì, gōngzuò zuò de hěn hǎo.

女的_i跪下来, \emptyset_i 开始给小狗洗澡, \emptyset_i 用刷子把它刷得很细心。然后, \emptyset_i 洗完澡后, \emptyset_i 用毛巾把小狗擦干, \emptyset_i 然后还给它梳毛。完了之后, \emptyset_i 又好像然后对自己很满意, \emptyset_i 工作做得很好。

[Woman_i kneel down-come, \emptyset_i start for doggie shower, \emptyset_i use brush PREP-it brush DE very careful. Then, \emptyset_i wash-completed shower after, \emptyset_i use towel PREP-doggie

wipe dry, \emptyset_i then also PREP-it comb hair. Finish-PRT later, \emptyset_i also seems then PREP-self very satisfied, \emptyset_i job done DE very well.]

The woman kneeled down to start washing the doggie by carefully brushing it with a brush. Then, after the shower, she dried the doggie with towel. Further, she even combed the doggie's hair. After all this, she seemed very satisfied with her own work. She thought she did a very good job.

A topic chain can go beyond the boundary of one sentence, as seen in (5) below. A topic chain does not necessarily end with the period mark. The notion of a topic may extend its semantic domain to more than one sentence. Correferential zero does not have to be in line with the sentence boundary either. On the other hand, a topic chain can also be shorter than one sentence, in which case there can be more than one independent topic-comment structures or topic chain existing within one sentence. Example (6) is a sentence produced in a free writing task by a L1 Chinese speaker. It is one sentence as marked by the period; however, there are four independent topic-comment structures in this one sentence. Since none of the four topics was repeated as coreferential zero in previously or subsequently, these four topic-comment structures are successive yet independent and do not form a topic chain. Within one topic chain, all the successive topic-comment structures are dependent.

(6) *Wǒ_i de fùqīn shì yì míng huǒchē diàodù, tā zài nánjīng gōngzuò, wǒ hǎn shǎo qù tā de bànɡōnɡshì, yīnwèi bàba de dānwèishíxíng de shì bàn jūnshìhuà guǎnlǐ.*

我的父亲_i是一名火车调度，他_j在南京工作，我_k很少去他的办公室，因为爸爸的单位_m实行的是半军事化管理。

[I DE father_i is a train dispatcher, he_j at Nanjing work, I_k rarely go to his office, because father's working unit_m apply DE was semi-militarized managed.]



My father was a train dispatcher who worked in Nanjing. I rarely went to his office, because his unit was semi-militarized managed.

2.2.1.3 Types of topic chains

The strategies a topic chain employed to connect the same topic successive topic-comment structures still require further thorough investigation. Of these strategies, this dissertation focuses on coreferential zero, covert conjunctions, and topic chain types.

A coreferential zero commonly occurs sentence initially or clause initially, however, these are not the only positions it can take. A coreferential zero can be in the middle or at the end of a sentence as well as a clause. The topic of full phonological form appears at the first topic-comment structure, while the shared topic in the form of coreferential zero can follow in the subsequent topic-comment structure; or the full phonological form can be placed at or after the second clause while the coreferential zero is placed at the first topic-comment structure(s). To better capture the semantic relation as well as further understand how the coreferential zero works among topic-comment structures in a topic chain, Li (2005) categorized ten patterns of topic chains (See Appendix A). In addition to the typical topic chain and cataphoric topic chain with one shared subject, also included were topic chains with overt double topics, covert double topics. In addition to topic chains where the full form topic and its coreferential zero both plays thematic roles as agent-agent, also included were topic chains where the full form topic and its coreferential zero play the thematic roles of patient-theme, patient-patient,

proposed patient, and presented patient, etc. Besides the above patterns, it is interesting to know that a montage topic chain pattern is also subsumed. According to a corpus-based study by Li (2005), the selection of the ten patterns of topic chains for Chinese L1 speakers in writing depends mainly on the style of writing and the subject matter of the texts. Yet, there are not many studies that have looked at the difficulty or acquisition sequence of these patterns for L2 Chinese speakers.

Below, the ten topic chain patterns proposed by Li (2005) are illustrated with examples. Thematic roles of the topics are marked in the illustrations as follows: agent is defined as the actor of an action verb; patient refers to the recipient of the action or the grammatical object in a clause; and theme identifies an entity undergoing a change of state or location and is usually described by a non-action verb or predicate. The double arrow sign “” indicates a coreferential relation between two topics, each in one of the two consecutive topic-comment structures. “” indicates a coreferential relation between non-topic and a topic, each in one of the two consecutive topic-comment structures. These ten topic chain patterns are not exclusive but can also occur in various combinations of two or more topic chain patterns.

Though Li's (2005) ten topic chain patterns provide great insights on how topic chains are composed, she did not propose such ten topic chain patterns for the research on Chinese syntactic complexity. Therefore, when exemplify and analyze Li's topic chain categorization, this dissertation casts more lights on the application of such topic chain patterns in the research of Chinese syntactic complexity conceptualization and operationalization. Some necessary modification is applied accordingly, such as choosing the topic in a topic chain when discussing Topic Chain Pattern 3: Patient-Theme/Agent

Topic. Furthermore, as we proposed in 2.2.1.2, a topic chain can go beyond or stay within the boundary of one punctuated sentence. Li (2005) did not discuss if a sentence and a topic chain share the same boundary, or in its written form, the same punctuation marks. Therefore, in addition to the typical examples consists of two topic-comment structures as in the illustrations, this dissertation included some extended examples that consists more than two topic-comment structures for each of the ten topic chains patterns. In review, a topic chain can be part of or go beyond a sentence as marked by punctuation marks. Some extended examples of topic chain such as (10) shows an inconsistent boundary with a sentence boundary and confirms that a topic chain and a sentence do not necessarily share the same boundary.

The two successive topic-comment structures in (7) share the same topic *māma* (妈妈, mother) which plays the thematic role of agent in both topic-comment structures. In the first topic-comment structure, the topic appears in its full phonological form. The same topic is then repeated in the form of coreferential zero at the beginning of the subsequent topic-comment structure which connects the second topic-comment structure with the preceding one, in which way it becomes a topic chain. As it is shown in (8), a topic chain can include more than two successive topic-comment structures, as long as a mutual topic is shared and repeated in the form of coreferential zero.

| | |
|--------------------------------------|---|
| Topic Chain Pattern 1: Typical Topic | T1 (Agent/Theme) — C1 ↓↑ T2 (∅, Agent/Theme) — C2 |
|--------------------------------------|---|

(7) *Māma zhīzhe xiǎohái de fángjiān, jiào érzi qù zhǎo bàba.*

妈妈_i指着小孩的房间, ∅_i叫儿子去找爸爸。

[Mom_i point-ING kid DE room, ∅_i ask son go find dad.]

Mom pointed at the kid's room, and asked the son to find dad.

(8) *Xiǎolín zuótiān zǎoshang bā diǎn zhōng qǐchuáng, qǐle chuáng yǐhòu qiánwǎng jiàoshì, zhǔnbèi bā diǎn bàn shàngkè.*

小林_i昨天早上八点钟起床, \emptyset_i 起了床以后前往教室, \emptyset_i 准备八点半上课。

[Little Lin_i yesterday morning 8:00am get up, \emptyset_i get up PRT then head to classroom, \emptyset_i prepare 8:30am take class.]

Little Lin got up at 8:00am yesterday, after which he headed to the classroom, preparing to go to class at 8:30.

The mutual topic of the topic-comment structures in topic chain (9) and (10) also plays a thematic role of agent at the beginning of each clause. However, the full form topic is placed at the second topic-comment structure, while the same topic shared in the first topic-comment structure takes the form of a coreferential zero. (10) is an extended version of such pattern of topic chain with a cataphoric topic. It is interesting that the full form topic does not even show up in the first sentence. It is after the full stop of the sentence and in the beginning of the second sentence that the shared topic of two sentences appears. All the four topic-comment structures in the first sentence share the same topic which all take the form of coreferential zero. This confirms that a topic can extend its semantic domain beyond sentence boundary, and that a topic chain can consist of several sentences. In this sense, a topic is more of a discourse notion if compared with the strict definition of a sentence.

| | |
|---|---|
| Topic Chain Pattern 2: Cataphoric Topic | T1 (\emptyset , Agent) — C1 ↓↑ T2 (Agent) — C2 |
|---|---|

(9) *Chī wánliǎo wǔfàn, wáng tàitài kāishǐ kàn bàozhǐ.*

\emptyset_i 吃完了午饭，王太太_i开始看报纸。

[\emptyset_i Eat-finish PRT lunch, Mrs. Wang_i start reading newspaper.]

After the lunch, Mrs. Wang started reading newspaper.

(10) *(Rén na, nǐ shì duōme róngyì shòu qíngshì de bǎibù, duōme róngyì wàngì guòqù ya!) Zài tā jiā chī wán fàn, huí dào “jiā”, yòu cóng huǒfáng dǎle yí fēn bàizi miàn mómo, yě chīle xiàqù. Wǒ cái zhīdào shénme shì “bǎo”.*

(人哪，你是多么容易受情势的摆布，多么容易忘记过去呀！) \emptyset_i 在她家吃完饭， \emptyset_i 回到“家”， \emptyset_i 又从伙房打了一份稗子面馍馍， \emptyset_i 也吃了下去。我_i才知道什么是“饱”！

[(Human being-PRT, you are how easily get situation DE pushed around, how easily forget the past-PRT) \emptyset_i at her home eat-finish meal, \emptyset_i return home, \emptyset_i then from cafeteria bought one-person portion barnyard millet flour bread, \emptyset_i also eat-PRT down-go. I_i finally know what is full!]

(It is so likely that the human beings are pushed around by the situation and forget the past!) After had meal at her place, I returned home. I also bought one-person portion barnyard millet flour bread from the cafeteria and finished it. Until this point I finally felt full!

In (11), the topic in the preceding topic-comment structure is *wǒ* (我, I). The post-verb object patient in the preceding topic-comment structure, *tā* (他, he), is also the topic in the subsequent topic-comment structure in the form of a coreferential zero. In line with

Li's (2005) analysis, there are then two different topics identified in this topic chain (11), which goes against the definition that a topic chain consists of connected topic-comment structures sharing the same topic. Viewing this topic chain as a whole, however, this dissertation argues that it is the zero coreferred topic, *tā* (他, he), that is the topic of the two dependent topic-comment structures in one topic chain, which also links the two topic-comment structures. Once two independent topic-comment structures are composed into one topic chain, they become dependent topic-comment structures and share one same topic with the whole topic chain. As Givón (1992) pointed out, that what makes a clause participant topical is not its grammatical status of subject or object, but rather, its thematic importance, recurrence, or continuity in discourse (p. 202). Such recurrence and continuity of the topic takes the form of coreferential zero. Coreferential zero therefore can be taken as one marker that identifies a topic chain.

| | |
|--|--|
| Topic Chain Pattern 3: Patient-Theme/Agent Topic | T1 (Agent) — C1 [... V NP(Patient)] ↓ T2 (∅, Theme) — C2 |
|--|--|

(11) *Wǒ qù kànkan tā, shuō bu dìng zài xuéxí ne.*

我去看看他_i, ∅_i说不定在学习呢。

[I go see-see him_i, ∅_i say-not-sure at study PRT.]

I will go check him. Perhaps he is studying.

Similar to the Topic Chain Pattern 3 of Patient-Theme Topic, in the example (12) for Topic Chain Pattern 4 of Patient-Patient Topic, when viewed separately, Li (2005) identified different topics *tā* (他, he) and *wǒ* (我, I) respectively for the two topic-comment structures. However, this dissertation, again, analysed *yì zhāng míngpiàn* (一张

名片, one namecard) as the topic of the whole topic chain as well as the topic shared by the two separate but dependent topic-comment structures since *yì zhāng míngpiàn* (一张名片, one namecard) is repeated in the form of coreferential zero. What is different from Topic Chain Pattern 3 is that the zero coreferred topic *yì zhāng míngpiàn* (一张名片, one namecard) in the subsequent topic-comment structure also plays the thematic role of patient.

| | |
|---|-----------------------------------|
| Topic Chain Pattern 4 : Patient-Patient | T1 (Agent) — C1 [... NP(Patient)] |
| Topic | ↓↑ T2 (∅, Patient) — C2 |

(12) *Tā dì gěi wǒ yì zhāng míngpiàn, wǒ piǎole yìyǎn.*

他递给我一张名片_i, 我瞟了∅_i一眼。

[He hand over to me one piece name card_i, I glance-PRT ∅_i one eye.]

He gave me a name card, at which I took a glance.

In (13), *ménpiào* (门票, admission) is the topic of the preceding topic-comment structure and it plays the thematic role of theme. While in the subsequent topic-comment structure, *ménpiào* (门票, admission) plays the thematic role of patient and takes the form of coreferential zero.

| | |
|--|----------------------------|
| Topic Chain Pattern 5: Theme-Patient Topic | T1 (Theme) — C1 |
| | ↓↑ T2 (∅, Patient) — C2 |

(13) *Ménpiào hěn guì, dàn yíshòu 'érkōng.*

门票_i很贵, 但∅_i一售而空。

[Gate-ticket_i very expensive, but ∅_i once sell then empty.]

The admission tickets were expensive, but sold out in no time.

Like in Topic Chain Pattern 4 of Patient-Patient Topic, the topic of Topic Chain Pattern 6 of Preposed Patient Topic, both its full form and its coreferential zero play the thematic role of patient in each topic-comment structure. According to Li (2005), Topic Chain Pattern 6 of Preposed Patient Topic is different from the Topic Chain Pattern 4 of Patient-Patient Topic. In Topic Chain Pattern 4 of Patient-Patient Topic, Li (2005) considered *tā* (他, he) and *wǒ* (我, I) in (12) as different topics and *yì zhāng míngpiàn* (一张名片, one namecard) as the topic for the whole topic chain. In Topic Chain Pattern 6, Li (2005) considered both the full form and coreferential zero of the topic, like *fúlǚ* (俘虏, captive) in (14), is the topic of both the preceding and subsequent topic-comment structures, as well as the whole topic chain. However, this dissertation, as argued above, does not discriminate the topic of the whole chain from the topics of separate topic-comment structures. This dissertation views the topic of the whole chain as the same one topic of its constituent dependent topic-comment structures.

| | |
|---|----------------------------|
| Topic Chain Pattern 6: Preposed Patient | T1 (Patient) — C1 |
| Topic | ↓↑ T2 (∅, Patient) — C2 |

(14) *Fúlǚ dài lái le méiyǒu? Kuài yāshàng lái!*

俘虏_i带来了没有? ∅_i快押上来!

[Captive(s)_i bring-PRT not? ∅_i Hurry escort up come!]

Did you bring the captive(s)? Escort them here right now!

Topic Chain Pattern 7 has to do with the use of the existential/presentative constructions in Chinese. In an independent existential/presentative construction, the topic is either a locative or time expression, and a new entity is introduced in the post-verb position. For example in (15) the indefinite noun phrase *yìrén* (一人, one person) in the preceding topic-comment structure as well as the presentative construction is coded as a coreferential zero being a definite topic in the subsequent topic-comment structure.

| | |
|----------------------------------|--|
| Topic Chain Pattern 7: Presented | T1 (Location/Time) — C1 [...NP] |
| Topic | $\downarrow \uparrow$ T2 (\emptyset) — C2 |

(15) *Mǎlù duìmiàn zǒu lái yìrén, hēlìn tāmen lìjí zǒu kāi.*

马路对面走来一人_i, \emptyset_i 喝令他们立即走开。

[Road opposite walk-come a person_i, \emptyset_i shout an order them right away walk away.]

One person came from across the road and shouted aloud an order for them to walk away right away.

When a series of existential/presentative constructions are connected via coreferential zero, it constitutes the Topic Chain Pattern 8 of Montage Topics. Presented a series of existential/presentative constructions, listeners or readers can even envision a picture of what they hear or read, moving their focal attention from one picture to another, or one part of the picture to another, as the topic-comment structures lead the visualized semantic flow. This is done in the same fashion as montage techniques in motion pictures (Li, 2005). Upon reading (16), readers can envision a picture of a photon spaceship. Following the connected topic-comment structures presenting different parts of the photon spaceship, readers' eyes or focal attention will first be laid on the front (*qiánmiàn*,

前面), and then led to the back (*hòumiàn*, 后面), and last to the part in between (*zhōngjiān*, 中间). The coreferentials zero in the second and third topic-comment structure help connect and form this montage topic chain.

| | |
|--------------------------------|--|
| Topic Chain Pattern 8: Montage | T1 (Locative) — C1 [...NP] |
| Topics | $\downarrow \uparrow$ T2 (\emptyset +position word, Locative) — C2 |

(16) *Guāngzǐ fēichuán qiánmiàn shì yǔháng yuán de zuòcāng hòumiàn shì guāngzǐ huǒjiàn fādòngjī; zhōngjiān yǒu yìcéng hěn hòu de bǎohù píng.*

光子飞船_i前面是宇航员的座舱, \emptyset_i 后面是光子火箭发动机; \emptyset_i 中间有一层很厚的保护屏。

[Photon spaceship_i front is astronauts DE capsule cabin, \emptyset_i back is photon rocket engine; \emptyset_i Between have one layer very thick DE protective screen.]

The astronauts' capsule cabin is in the front, and the rocket engine is at the back. Also there is a thick protective screen in between the two.

Double nominative construction is another Chinese particular construction. Li (2005) argued that in the Topic Chain Pattern 9 of Overt Double Topics, in addition to the general topic in a topic-comment structure, a secondary topic follows right inside the comment part of the general topic. In (17), *zhànshìmen* (战士们, soldiers) is the general topic while *yǒude* (有的, some) is the secondary topic of the first topic-comment structure. The general topic takes the form of coreferential zero in the subsequent two topic-comment structures, which connects these three topic-comment structures into a topic chain. Since secondary topic coexist in its overt full form with the general topic

within one topic-comment structure, this pattern of topic chain was named Topic Chain of Overt Double Topics. However, this dissertation argues that the general topic *zhànshìmen* (战士们, soldiers) is as well the topic of all the dependent topic-comment structures which were composed into one topic chain.

| | |
|-------------------------------------|-------------------------------|
| Topic Chain Pattern 9: Overt Double | T1 — C1 [T' — C'] |
| Topics | ↓↑ T2 (∅) — C2 [T'' — C''] |

(17) *Zhànshìmen yǒude zài jǐnjí yùnsòng dànàoyào, yǒude zhànshì zài tái sòng*

shāngyuán, hái yǒude zài pīnmìng shā dí.

战士们_i有的在紧急运送弹药, ∅_i有的在抬送伤员, ∅_i还有的在拼命杀敌。

[Soldiers_i some at urgency ferry ammunition, ∅_i some at carry send the wounded, also ∅_i some at risk one's life kill enemy.]

Some of the soldiers are urgently ferrying ammunition, some are carrying the wounded, and some of them are risking their lives in battle.

As befits the name, Li (2005) argued in Topic Chain Pattern 10 of Covert Double Topics, in the topic-comment structure that consists of the general topic, the secondary topic is covert or there is no secondary topic. Instead of overtly coexisting with the general topic within one topic-comment structure, the secondary topic appears in a different topic-comment structure of the same topic chain. As in (18), *tā* (她, she) is the general topic of the topic chain. *Tā* (她, she) takes its full form in the first topic-comment structure with no secondary topic. This general topic is then repeated in the form of a coreferential zero in the subsequent topic-comment structure to form a topic chain. In the subsequent topic-comment structure, there exists a secondary topic, *jǐbèi* (脊背, back).

Again, this dissertation argues that the topic gets repeated in the form of coreferential zero, as *tā* (她, she) here in (18), is the topic of whole topic chain as well as its constituent dependent topic-comment structures.

| | |
|---------------------------------------|-----------------------------|
| Topic Chain Pattern 10: Covert Double | T1 — C1 |
| Topics | ↓↑ T2 (∅) — C2 [T' — C'] |

(18) *Tā zhàn zài nàlǐ, jǐbèi kào zhe qiáng.*

她_i 站在那里, ∅_i 脊背靠着墙。

[She_i stand at there, ∅_i back lean against the wall.]

She stands there, leaning back against the wall.

2.2.2 Parataxis-prominence versus hypotaxis-prominence

2.2.2.1 Parataxis and hypotaxis

Besides the contrast of topic-comment versus subject-predicate regarding clause/sentence structure, another contrast exists between Chinese and English, regarding clause constituent in terms of embedding and combining, i.e., parataxis-prominence versus hypotaxis-prominence. Chinese is generally considered as a parataxis-prominent language while English is hypotaxis-prominent (Lian, 1993; Wang, 1957; Shen, 1988; Wang, 1990, 1992). Parataxis and hypotaxis are two types of fundamental syntagmatic relation. Parataxis is defined by *Webster's Third New International Dictionary* as “the coordinate ranging of clauses, phrases, or words one after another without coordinating connectives as in *He laughed; she cried*; the placing of a subordinate clause beside a main clause without a subordinating connective as in *I believe it is true* and *There is a man wants to see you*”. With parataxis, language units such as, words, phrases, or clauses

are connected without covert connectives. Hypotaxis, in contrast to parataxis, according to the *American Heritage Dictionary*, is “the dependent or subordinate construction or relationship of clauses with connectives, for example, *I shall despair if you don't come*”. With hypotaxis, overt lexical, grammatical, and inflectional devices connect language units. In a broad sense, parataxis and hypotaxis apply to the connection between words, phrases and clauses. This dissertation applies a narrow definition and focuses on the syntactic connection between clauses or sentences. Parataxis and hypotaxis take two opposite extremes of a swinging pendulum in the dimension of clause-combining mechanisms. In different languages, rather than exclusively applying parataxis or hypotaxis, the more frequent case is for both of these two mechanisms to be employed yet with varied partial adoption.

With parataxis, language units such as words, phrases, or clauses are connected in a semantic stream without overt coordinating connectives. A Chinese example from Thompson and Longacre (1985) is as follows:

(19) *Tā méi niànshū, tā dǎqiú le.*

他没念书，他打球了。

[he NEG study book, he hit ball ASP.]

Instead of studying, he played ball. (p. 175)

With no overt connectives as in English, such juxtaposition of two complete subject-predicate structures over a single intonation contour is very commonly used in Chinese to form a compound sentence. Another example is given in (20):

(20) *Běidà wǒ dāi de duǎn, Yāndà dāi de zhǎng.*

北大我呆得短，燕大呆得长。

[Peking University I stayed DE short, Yanjing Univerisity stayed DE long.]

I stayed in Peking Univerisity for a short time, while stayed in Yanjing University for a long time.

Here in (20), two paralleled structures “place + verb + time” are juxtaposed with no overt conjunctions but a rising tone instead. Via such parataxis a semantic comparison is presented on the duration of my stay at two universities.

One well-know example of parataxis is from classical Latin, a sentence reportedly written by Julius Caesar in 47 BC as a comment on his short war with Pharnaces II of Pontus in the city of Zela (currently known as Zile, in Turkey). While there was no overt connective used to connect the three clauses, the three autonomous clauses were connected by an overarching intonation contour as one sentence:

(21) Veni, vidi, vici.

I came, I saw, I conquered. (*c.* 46, Suetonius, Jul. 37)

The parataxis examples (19), (20), and (21) above are composed by two or three formally, semantically, and functionally equal units. In some registers of English, parataxis is additionally used to connect a matrix clause and a subordinate clause or an embedded clause. For example, in (22) (Hopper & Traugott, 2003), the relative clause “just walked out the store” is embedded without a relative pronoun to the main clause and functions as the modifier of the head noun “the guy.”

(22) That guy [0] just walked out the store reminds me of the photo in the post-office window. (p. 180)

Another Chinese example of clause complexing without overt conjunction can be seen in (23). The first topic-comment structure *tīngxìn le tā de huà* (听信了他的话, *believed his word*) is semantically the cause of the subsequent topic-comment structure *Xǔ Lì chéngle Liú Xiǎoxióng de “qiānyuē yǎnyuán”* (许丽成了刘小雄的“签约演员”, Li Xu became Xiaoxiong Liu’s contracted actress). The sentence can be marked with the cause and effect conjunction pair *yīnwèi...suǒyǐ...* (因为……所以……, because...so...), and thus exhibit a cause subordination “[*Yīnwèi*] *Tīngxìn le tā de huà*, [*Suǒyǐ*] *Xǔ Lì chéngle Liú Xiǎoxióng de ‘qiānyuē yǎnyuán’*.” (因为)听信了他的话, (所以)许丽成了刘小雄的“签约演员”。 Because she believed his word, Li Xu became Xiaoxiong Liu’s contracted actress.) However, with no overtly formally marked connectives, the two clauses in (23) can as well be taken as two chronological continuous situations. As Lian (1993) suggested, English sentences were more precise and Chinese sentences were more concise. Brackets can also be added with a different conjunction *yúshì* (于是, thereupon) and exhibit a coordination of continuity for the coexisting sequential relationship between the two actions: “*Tīngxìn le tā de huà*, (*yúshì*) *Xǔ Lì chéngle Liú Xiǎoxióng de ‘qiānyuē yǎnyuán’*.” (听信了他的话, (于是)许丽成了刘小雄的“签约演员”。 Believed his word, and then Li Xu became Xiaoxiong Liu’s contracted actress.) With both scenarios of connectives insertion grammatical and reasonable, whichever semantic relation was implied or expressed thus depended on the context. Therefore, these two clauses in (23) not only can be coordinated as they exhibit coexisting sequential relationship, but also

can be subordinated as they exhibit cause and effect relationship. It is therefore not explicit and arbitrary if these two clauses are coordinated or subordinated. As such, it is typically not a good idea to add on overt connectives to make explicit the syntagmatic relation between Chinese clauses that are connected paratactically.

(23) *Tīngxìn le tāde huà, Xǔ Lì chéngle Liú Xiǎoxióng de “qiānyuē yǎnyuán”.*

听信了他的话，许丽成了刘小雄的“签约演员”。

[Listen believe ASP he DE word, Li Xu became ASP Xiaoxiong Liu DE
“contracted actress”.]

Deceived by his word, Li Xu became Xiaoxiong Liu’s contracted actress.

Sentence (23) above, however, is not the only case of such ambiguous clausal relation under the condition that no overt conjunctions are marked. In sentence (24), the English translation provided in Li (2005) presumed the clause combining as subordination of supposition: “(*Yàoshì*) *Tā bìngle, jiù yīnggāi qù kàn yīshēng.* ((要是) 他病了，就应该去看医生。 If he is sick, he should go to see the doctor.) (p. 2)” However, the sentence can also be a subordination of cause: “(*Jìrán*) *Tā bìngle, jiù yīnggāi qù kàn yīshēng.* ((既然) 他病了，就应该去看医生。 Since he is sick, he should go to see the doctor.)” The syntagmatic relations of such paratactically connected clauses can be ambiguous to define or categorize into one subordination type. Such sentences consisting of paratactically combined clauses may have to be put into the context or discourse to tell of the semantic relationship in order to identify the syntagmatic relation between the clauses of the sentence.

(24) *Tā bìngle, jiù yīnggāi qù kàn yīshēng.*

他病了，就应该去看医生。

[He sick PRT, then should go see doctor.]

If he is sick, he should go to see the doctor. (p. 2)

In contrast to parataxis, hypotaxis applies overt connectives to form the dependent or subordinate construction or relationship of clauses. Sentence (25) exhibits coordination between the conjunctive “and” connecting the two paralleled nucleus clauses “They lived in a prosperous suburb of Hartford, Connecticut,” and “bought a Victorian beach house on the shore.” Sentence (26) exhibits a subordination of condition with *zhǐyǒu...cái...* (只有……才……, unless...then...) overtly marking the subordinate condition clause *zhǐyǒu héping gòngchǔ* (只有和平共处, *unless living in peace*) and the matrix result clause *shìjiè cáinéng wěndìng* (世界才能稳定, the world then can be stable).

(25) They lived in a prosperous suburb of Hartford, Connecticut, and bought a Victorian beach house on the shore.

(26) *Zhǐyǒu héping gòngchǔ, shìjiè cáinéng wěndìng.*

只有和平共处，世界才能稳定。

[Unless peace coexist, world then stable.]

The world cannot be stable without our living in peace.

In considerations of the above analyses, parataxis and hypotaxis can be seen to take opposite swings of a pendulum in terms of the dimension of clause-combining mechanisms. In different languages, rather than exclusively applying parataxis or

hypotaxis, it is often the case that both of these two mechanisms are employed yet with different partial adoption. From a contrastive perspective, one of the most prominent differences between Chinese and English at the syntactic level falls on the use of parataxis and hypotaxis. Chinese is generally considered as a parataxis-prominent language with English as a hypotaxis-prominent language (Lian, 1993; Wang, 1957; Shen, 1988; Wang, 1990, 1992). With sufficient contextual clues provided, overt conjunctions with no phonologically content do not interrupt or cut off the continuity between clausal units therefore serve to exhibit a higher density of information.

2.2.2.2 Coordination and subordination

Coordination and subordination are the most widely used terminologies for clause combining analysis. Coordination and subordination can also be defined in a broad sense and a narrow sense. In a broad sense, coordination and subordination can refer to ways of combining words, phrases, and clauses into more complex forms. While in a narrow sense, as well as parataxis and hypotaxis, coordination and subordination can be restricted to the sentential level. Coordination and subordination discussed in this dissertation are in the narrow sense. A complex sentence is referring to a coordinate or subordinate sentence that consists of more than one clause. Coordination uses coordinating conjunctions, conjunctive adverbs (with appropriate punctuation), and punctuation, among others, to combine short independent clauses into a single sentence. Subordination uses mechanisms like subordinating conjunctions or relative pronouns to attach one or more independent clauses to another independent clause thus creating subordinated clauses to this independent matrix clause. “In coordination the units are

constituents at the same level of constituent structure, whereas in subordination they form a hierarchy, the subordinate unit being a constituent of the superordinate unit” (Quirk, Greenbaum, Leech, & Svartvik, 1985, p. 919). Clauses in coordination are of equal semantic value and formal position, whereas in subordination the main sentence is given precedence over other subordinated sentences.

As Table 1 below illustrates, in traditional grammar subordination subsumes three different types: complement clauses, relative clauses, and adverbial clauses. Complement clauses function as arguments of a predicate in a superordinate clause, as in examples of (27) and (28). Relative clauses function as attributes of a noun phrase as in (29) and (30). Adverbial clauses take the position of modifier of the associated matrix clause or verb phrase, as shown in (31) and (32). Coordination includes coordinated clauses that function as equal parts of a sentence, as in (33).

Table 1

Categories of clause complexing in English

| | Example sentence | Subordination types |
|------|---|-----------------------------|
| (27) | Peter promised <u>that he would come</u> . | Finite complement clause |
| (28) | Sue wants <u>Peter to leave</u> . | Nonfinite complement clause |
| (29) | Sally bought the bike <u>that was on sale</u> . | Finite relative clause |
| (30) | Is that the driver <u>causing the accidents</u> ? | Nonfinite relative clause |
| (31) | He arrived <u>when Mary was just about to leave</u> . | Finite adverbial clause |
| (32) | She left the door open <u>to hear the baby</u> . | Nonfinite adverbial clause |
| (33) | He tried hard, <u>but he failed</u> . | Coordinate clause |

(from Diessel, 2004, p. 1)

Matthiessen and Thompson (1988) analogized rhetorical organization in text to clause combining. They differentiated embedding from clause combining: “If subordinate clause is taken to mean a clause that functions as subordinated to another grammatical unit, this fails to make the distinction between embedding and clause combining” (p.286). Complement clauses and relative clauses that subordinated to part of the matrix clause were classified as *clause embedding*. While adverbial clauses that subordinated to the matrix clause were called hypotactic enhancing clauses, and such were considered as *clause combining*. In addition, from the perspective of acquisition of these complex sentences, Diessel (2004) distinguished two different developmental pathways for acquiring the above mentioned four types of complex sentences. Complement and relative clauses emerged from simple sentences that were gradually expanded to multiple-clause structures. Adverbial and coordinate clauses developed by integrating two independent sentences into a specific biclausal unit. In concert with such differentiation of acquisitional pathways, complement clause and relative clause were separated from subordination. Such differentiation of acquisitional pathways confirmed the differentiation of between complement clause and relative clause as *clause embedding* and adverbial clauses as *clause combining/subordination*.

Based on the categorization of coordination and subordination, Hunt (1965) defined T-unit as the “shortest grammatically allowable sentences into which (writing can be split) or minimally terminable unit”, and operated it as “one main clause with all subordinate clauses attached to it” (p.20). Hunt (1970) identified three main indices to measure the level of text productions’ syntactic complexity: (a) Median Length of Minimal Terminable Syntactic Unit (T-Unit), (b) Clause-Length, and (c) Syntactic

Complexity Index (SCI). The length of T-unit and clause were calculated by the total amount of words used in one T-unit or one clause. Syntactic Complexity Index (SCI) is the ratio of the number of subordinated sentences for every T-unit.

With the categorization of coordination and subordination, an English coordinated sentence consists of two or more clauses joined either by a semicolon or by a comma working in concert with a coordinating conjunction; a subordinate sentence attaches to a main clause a dependant clause with either a subordinating conjunction or a relative pronoun. As aforementioned, relative and complement clauses are categorized as clause embedding, while adverbial clauses are taken as subordination. Subordination can be semantically sub-categorized according to varying semantic relations between the subordinate and main clause: time, reason or cause, purpose or result, condition, contrast, choice, and place or location. These different semantic relations can be overtly demonstrated through the use of different conjunctions or relative pronouns.

As listed in Table 2 below, in English the coordination, subordination, and embedding sentence categorization can be exemplified with typical connective mechanisms. For coordination, there is paratactically connected coordination of which clauses are connected without overt conjunctions. Hypotactically connected coordinating sentences may apply conjunctives like *for*, *and*, *nor*, *but*, *or*, *yet*, and *so* (Curzan & Adam, 2009). *And* presents non-contrasting item(s) or idea(s), e.g., They gamble, and they smoke. *For* presents a reason, e.g., He is gambling with his health, for he has been smoking far too long. *Nor* presents a non-contrasting negative idea, e.g., They do not gamble nor do they smoke. *But* presents a contrast or exception, e.g., They gamble, but they don't smoke. *Or* presents an alternative item or idea, e.g., Every day they gamble or

they smoke. *Yet* presents a contrast or exception, e.g., They gamble, yet they don't smoke. *So* presents a consequence, e.g., He gambled well last night so he smoked a cigar to celebrate. Besides these single conjunctions, there are also correlative conjunctions working in pairs to conjoin clauses. There are six different pairs of correlative conjunctions: *either...or...*, e.g., You either do your work or prepare for a trip to the office; *not only...but (also)...*, e.g., Not only is he handsome, but he is also brilliant; *neither...nor...* (or increasingly *neither...or...*), e.g., Neither the basketball team nor the football team is doing well; *both...and...*, e.g., Both the cross country team and the swimming team are doing well; *whether...or...*, e.g., Whether you stay or you go, it's your decision; and *just as...so...*, e.g., Just as Aussies love Aussie rules football, so many Canadians love ice hockey.).

Table 2

Coordination and subordination in English

| Syntactic structures | English Connective mechanisms of complexity |
|----------------------|--|
| Coordination | Semicolon, comma + coordinating conjunctions like <i>and, but, for, or, nor, yet, so; either...or..., not only...but (also)...</i> , <i>neither...nor...</i> (or increasingly <i>neither...or...</i>), <i>both...and...</i> , <i>whether...or...</i> , <i>just as...so.</i> |

| | | | |
|---------------|-------------------|---|---|
| Subordination | Adverbial clause | Time | <i>After, before, once, since, until, when, whenever, while</i> |
| | | Reason or cause | <i>As, because, since</i> |
| | | Purpose or result | <i>In order that, so that, that</i> |
| | | Condition | <i>If even, if, provided that, unless</i> |
| | | Contrast | <i>Although, even though, though, whereas, while</i> |
| | | Choice | <i>Than, whether</i> |
| | | Place or location | <i>Where, wherever</i> |
| | | Manner | <i>As, like, the way</i> |
| Embedding | Complement clause | <i>Who, whom, what, which, whoever, whomever, whatever, when, where, how, why</i> | |
| | Relative clause | <i>That, which, who, whom, whose</i> | |

These English syntactic structures listed in Table 2, however, are not correspondingly applied in Chinese linguistics analysis and Chinese second language acquisition studies.

Instead of being embedded subsequent to the head in the form of a clause led by conjunctions like *that, which, who, whom, or whose*, the Chinese equivalent of English

relative clause appears before the head noun marked by the particle *de* (的, modifier marker) (Li et al, 1989, p. 216). The Chinese equivalent of English relative clauses take the form of pre-head modifiers and function as attributes of a noun or NP. Below (29) and (30) are previously listed example sentences in Table 1 with relative clauses. (29a) and (30a) are respectively the Chinese translation of (29) and (30). After translated into Chinese, the Chinese equivalents of both English finite and nonfinite relative clauses take the form of attributes ended with *de* (的, modifier marker) preceding the head noun. In (29a), *nà liàng zài cùxiāo de* (那辆在促销的, that was on sale) is the attribute for the head noun *zìxíngchē* (自行车, bike). In (30a), *zàochéng shìgù de* (造成事故的, causing the accidents) is the attribute for the head noun *sījī* (司机, driver).

(29) Sally bought the bike that was on sale. (finite relative clause)

(29a) *Sàilì mǎile nà liàng zài cùxiāo de zìxíngchē.*

塞丽买了那辆在促销的自行车。

[Sally bought LE that CL on sale DE bike.]

(30) Is that the driver causing the accidents? (nonfinite relative clause)

(30a) *Nà shì zàochéng shìgù de sījī ma?*

那是造成事故的司机吗?

[That is causing accident DE driver?]

As for the Chinese equivalent of the English complement clause, instead of being embedded subsequent to the head in the form of a clause led by conjunctions like *who*, *whom*, *what*, *which*, *whoever*, *whomever*, *whatever*, *when*, *where*, *how*, or *why*, it can appear as an argument of the previous verb, or form a serial verb construction with the

previous verb. (27a) and (28a) are the Chinese equivalents of the two complement clause examples (27) and (28) listed previously in Table 1. In (27a) which is the translation of an English finite complement clause (27), the Chinese equivalent of the English complement clause *tā huì lái* (他会来, he will come) functions as the argument of the predicate *dāyìng* (答应, to promise). In (28a), the Chinese equivalent of the English nonfinite complement clause can also be analyzed as the argument of verbal phrase *xiǎngyào* (想要, want to). However, (27a) and (28a) can be as well analyzed as the serial verb construction “NP1+VP1+NP2+VP2.”

(27) Peter promised that he would come. (finite complement clause)

(27a) *Pítè dāyìng tā huì lái.*

皮特答应他会来。

[Peter promised he will come.]

(28) Sue wants Peter to leave. (nonfinite complement clause)

(28a) *Sū xiǎng yào Pítè líkāi.*

苏想要皮特离开。

[Sue want Peter leave.]

The Chinese equivalent of an English adverbial clause also takes a very different form. (31a) and (32a) are the Chinese equivalents of two adverbial clause examples (31) and (32) listed in Table 1. The Chinese equivalent of the English finite adverbial clause in (31a), *Mǎlì gāng yào zǒu de shíhòu* (玛丽刚要走的时候, when Mary was just about to leave), forms an adverbial modifier of the clause *tā lái le* (他来了, he came). The Chinese

equivalent of the English non-finite adverbial clause in (32a), *lái tīng bǎobao de dòngjìng* (来听宝宝的动静, to hear the baby) forms a serial verb construction with the previous verbal phrase *kāizhe mén* (开着门, to keep the door open).

(31) He arrived when Mary was just about to leave.

(31a) *Mǎlì gāng yào zǒu de shíhòu, tā lái le.*

玛丽刚要走的时候, 他来了。

[Mary just want leave DE time, he come LE.]

(32) She left the door open to hear the baby.

(32a) *Tā kāizhe mén lái tīng bǎobao de dòngjìng.*

她开着门来听宝宝的动静。

[She open ZHE door to listen baby DE movement.]

Since the Chinese equivalents of the English relative, complement, and adverbial clause take very different forms, analysis on such structures shall correspondingly take a different stand but not mechanically apply the analysis for English structures. Alongside the distinction between clause embedding and clause combining as the clause complexing mechanism, adding attributive or adverbial modifiers can be analyzed as clause embedding at subclausal level to lengthen the superordinate unit. As for serial verb structures, whether considered as clause embedding at subclausal level or clausal combining at clausal level also depends on if two verb structures share one same intonation contour or different ones. In this dissertation, when an argument stays within one intonation contour with its preceding predicate, it is analyzed as one unit within a clause. Therefore adding an argument within the same intonation contour of the predicate

is treated as clause embedding at the subclausal level. When a post-verbal argument stays outside of the intonation countour of its preceding predicate, it is analyzed as another separate language unit from the preceding language unit. Therefore adding an argument outside of the same intonation contour is treated as clausal level complexing. For example, in (34) the post-verbal argument *tā dōu hěn téng wǒ yě ài wǒ* (他都很疼我也愛我, he has always doted on me as well as love me) is analyzed as a part of the language unit of *wǒ zhīdào* (我知道, I know) since the argument stays within the same intonation contour of the predicate *zhīdào* (知道, know). There is no punctuation pause or stop between *zhīdào* and *tā dōu hěn téng wǒ yě ài wǒ*. Attaching a post-verbal argument that shares the same intonation contour with the predicate is analyzed as clause embedding at subclausal level. In (35), the post-verbal argument *wǒ hé wǒ de gēge chūqù wán* (我和我的哥哥出去玩, me and my elder brother go out having fun) stays outside the intonation contour of *wǒ jì de yǒuyí cì*, it is then analyzed as a separate language unit. The relationship between *wǒ jì de yǒuyí cì* and *wǒ hé wǒ de gēge chūqù wán* is then analyzed as clause combining in terms of clause complexing.

(34) *Wǒ zhīdào tā dōu hěn téng {téng} wǒ yě ài wǒ.*

我知道他都很騰{疼}²我也愛我。

[I know he always very dote on me also love me.]

I know that he has always doted on me as well as loved me.

(35) *Wǒ jì de {jìdé} yǒuyí cǐ {cì}, wǒ hé wǒ de gēge chūqù wán.*

² Throughout this paper, for the spoken output transcription, characters provided in “{ }” are the correct version for a typographical error made by the TW&ST participants.

我計的{記得}有一此{次}, 我和我的哥哥出去玩。

[I remember has once, I and I DE elder brother go out playing.]

I remember there is once that me and my elder brother go out having fun.

Same for the serial verb constructions, when such two-verb construction shares one same intonation countour, they are treated as clause embedding at subclausal level. When such two-verb constructions are under different intonation countours, they are treated as clause combining at a clausal level. (36) and (37) are both sentences with a serial verb construction. (36) is analyzed as one clause under one intonation contour, therefore the serial verb construction in (36) is considered a subclausal level complexity syntagma. (37) is analyzed as two language units under two separate intonation contours. More examples analyzed in terms of TC-units are provided in Table 10 in Section 4.3.3.

(36) *Māma jiào nǐ qù chīfàn.*

妈妈叫你去吃饭。

[Mother ask you go eat food.]

Mother asks you to go to eat.

(37) *Māma jiào nǐ, qù chīfàn.*

妈妈叫你, 去吃饭。

[Mother ask you, go eat food.]

Mother asks you to go to eat.

To fit the concepts of coordination and subordination into their Chinese shoes of different typological features, different mechanisms and strategies had to be amended. In

addition, whereas both semantic value and formal position of the clauses are taken into account when categorize coordination or subordination in English, for Chinese which lacks of inflection, semantic relation became the only classification criteria. According to the semantic relation between separate clauses, Chinese complex sentences are subcategorized into coordinate complex sentences, in which the separate sentences are semantically equal, and subordinate complex sentences, in which one separate clause is the semantically main clause while the other is semantically subordinate. Besides coordinate complex sentences and subordinate complex sentences, multi-complex sentences are those composed by separate clauses which are not all at the same level but instead have one complex sentence embedded in another.

One of the most popular linguistics textbook at higher education level in China, *Xiandai Hanyu (Modern Chinese, Beijingdaxue Zhongwenxi Xiandaihanyu Jiaoyanshi, 2004)*, introduced a widely accepted categorization of coordination and subordination in Chinese. According to this textbook, the semantic and logic relationships between separate clauses in a coordinate complex sentence include parallelism, continuity, progressiveness, and choice. Complex sentences of coordination connect separate clauses of paralleled events, situations or equal aspects of the same object. The semantic and logic relationships between the separate clauses in a subordinate complex sentence include conversion, supposition, condition, cause and effect, purpose, according variation, and time. In a subordinate complex sentence, one separate clause is semantically more important than the other(s). The more semantically important clause is called the principal clause while the less semantically important clause(s) is called the subordinate clause.

When illustrate different types of Chinese coordinations and subordinations, the commonly used conjunction is often listed to give typical examples of each type. The commonly used conjunctions to connect coordinate clauses of parallelism are *yě...yě...* (也……也……, (also)...also...), *jì...yě...* (既……也……, both...and...), *yòu...yòu...* (*yòu...*) (又……又……(又……), also...also...(also...), as well as), and *yímiàn...yímiàn...* (一面……一面……, while). Complex sentences of continuity connect clauses referring to a series of continued events. The commonly used conjunctions here are *yòu* (又, and), *jiù* (就, as soon as), *ránhòu* (然后, then), and *hòulái* (后来, later). Coordinate sentences of progressiveness connect clauses of which one is semantically progressive based on the other. The commonly used conjunctions here are *hái* (还, also, even), *érqiě* (而且, and that), *jìn'ér* (进而, and then), *hékuàng* (何况, the rather that), *kuàngqiě* (况且, in addition), *nǎizhì* (乃至, and even), *shènzhì* (甚至, even), and *bùdàn...érqiě...* (不但……而且……, not only... but also...). Coordination of choice connects separate clauses that show two or several mutually exclusive choices. The commonly used conjunctions to connect these clauses are *huòzhě* (或者, or), *háishì* (还是, or), *nìngkě* (宁可, preferably), *yǔqí* (与其, rather than), *bùshì...jiùshì...* (不是……就是……, if not... then...), and *yàome...yàome...* (要么……要么……, either...or...).

The semantic and logic relationships between the clauses of Chinese subordinations include conversion, supposition, condition, cause and effect, purpose, according variation, and time. In contrast to the coordination where the clauses share equal semantic weight, in a subordinate complex sentence, one separate clause is semantically more important than the other(s). The more semantically important clause is

called the principal or matrix clause while the less semantically important clause(s) is called the subordinate clause. Complex sentences of conversion connect clauses of opposite meaning. The commonly used conjunctions to connect these clauses are *dànshì* (但是, but), *kěshì* (可是, but), *ér* (而, yet), *rán'ér* (然而, whereas), and *què* (却, yet). Complex sentences of supposition connect subordinate clauses of hypothesis and principle clauses of result. The commonly used conjunctions here are *jiǎshǐ* (假使, be it that), *jiǎrú* (假如, provided that), *rúguǒ* (如果, if), and *yàoshi* (要是, suppose). Complex sentences of condition connect subordinate clauses of condition or premise and principle clause of result. Subordinations of condition subsume three kinds of condition: sufficient condition, necessary condition, and noncondition. In complex sentences of sufficient condition, as long as what the subordinate clause talking about is realized, what the principle clause refers to can be guaranteed to be realized as the result. The commonly used conjunctions to connect the separate clauses of sufficient condition are *zhǐyào* (只要, as long as) and *dànshǐ* (但使, as long as). In complex sentences of necessary condition, the result that the principle clause refers to cannot be fulfilled unless the condition told in the subordinate clause is realized. The commonly used conjunctions to connect the separate clauses of necessary condition are *zhǐyǒu* (只有, no but) and *chúfēi*, (除非, unless). In complex sentences of noncondition, the result that the principle clause refers to will not change no matter what condition shown in the subordinate clause is applied. The commonly used conjunctions to connect the separate clauses of noncondition are *wúlùn...yě...* (无论……也……, no matter), *bùguǎn...dōu...* (不管……都……, in spite of), *búlùn* (不论, no matter), and *rènpíng* (任凭, no matter). Continuing

on, complex sentences of cause and effect, as the term suggests, connect the subordinate clause of cause and the principle clause of effect. The conjunctions commonly used in the subordinate clause of cause are *yīnwèi* (因为, because) and *yóuyú* (由于, since), and the conjunctions commonly used in the principle of effect are *yīncǐ* (因此, therefore) and *yīn'ér* (因而, thus). Complex sentences of purpose connect the principle clause of action and the subordinate clause of purpose. The action is taken to realize the goal of achieving or avoiding something. The commonly used conjunctions here are *wèile* (为了, for), *yìbiàn* (以便, so that), *miǎnde* (免得, lest), *shěng dé* (省得, so as to avoid), and *yǐmiǎn* (以免, in order to avoid). Complex sentences of according variation show a semantically varied relation between the subordinate clause and the principle clause. The likeliness of goal realization, the direction of development, or the state of being described in the principle clause varies according to the change described in the subordinate clause. The sentence pattern *yuè... yuè...* (越……, 越……, the more..., the more...) and interrogative pronouns like *nǎlǐ* (哪里, where) and *shéi* (谁, who, whoever) are commonly used in complex sentences of according variation. Complex sentences of time connect an event in a principle clause to a subordinate clause, which provides a time reference for that event. As exemplified, most of these complex sentence types have their own commonly used conjunctions or correlative conjunctions. However, as mentioned earlier these overt conjunctions are not indispensable, in which case we argue covert conjunctions are applied. More discussion about covert conjunctions will be included in the following discussion regarding the topic-prominent trait of Chinese. Table 3 below summarizes with examples the types of complex sentences observed in Chinese.

Table 3

Categories of Chinese complex sentences

| Categories | Semantic relation | Conjunctions | Example |
|-------------------------------|-------------------|---|--|
| Coordinate complex sentences | Parallelism | Overt conjunctions; (yě)...yě..., (也……)也……; jì...yě..., 既……也……; yòu...yòu... (yòu...), 又……又…… (又……); yímiàn...yímiàn..., 一面……一面……; etc. | (38) <i>Tā yímiàn gěi bìngrén zhìbìng, yímiàn yuèdú dàliàng yīshū.</i> 他一面给病人治病，一面阅读大量医书。 [He while for patients cure disease, while read many medical book.] He sees the patient and read many medical books. |
| | Continuity | Overt conjunctions; yòu, 又; jiù, 就; ránhòu, 然后; hòulái, 后来; etc. | (39) <i>Xiān chīfàn, hòu mǎidān.</i> 先吃饭，后买单。 [First eat-meal, later buy the bill.] Eat first and then pay. |
| | Progressiveness | Overt conjunctions; hái, 还; érqǐě, 而且; jìn'ér, 进而; hékuàng, 何况; kuàngqǐě, 况且; nǎizhì, 乃至; shènzhì, 甚至; búdàn...érqǐě..., 不但……而且……; etc. | (40) <i>Zhè zhǒng fādòngjī bùdàn fēixíng sùdù kuài, érqǐě jiéyuē ránliào.</i> 这种发动机不但飞行速度快，而且节约燃料。 [This kind engine not only fly speed fast, but also economize fuel.] This engine not only flies fast but also economizes fuel. |
| | Choice | Overt conjunctions; huòzhě, 或者; hái, 还是; nìngkě, 宁可; yǔqí, 与其; búshì...jiùshì..., 不是……就是……; yào...yào..., 要么……要么……; etc. | (41) <i>Yàome xīshēng, yàome qǔdé shènglì.</i> 要么牺牲，要么取得胜利。 [Either sacrifice our lives, either get victory.] (We will) Either sacrifice our lives or win the victory. |
| Subordinate complex sentences | Conversion | Overt conjunctions; suīrán, 虽然; dànshì, 但是; kěshì, 可是; ér, 而; | (42) <i>Wǒ xiǎng qù zhǎo tā, kěshì tā bù lǐ wǒ.</i> 我想去找她，可是她不理我。 [I want go find her, but she not take notice of me.] I want to go find her, but she brushes me off. |

| | | | |
|------------------|---|--|---|
| | <i>rán'ér</i> , 然而; <i>què</i> , 却; <i>suīrán</i> ... | | |
| | <i>dànshì</i> ..., 虽然……但是……; etc. | | |
| Supposition | Overt conjunctions; <i>jiǎshǐ</i> , 假使; <i>jiǎrú</i> , 假如; <i>rúguǒ</i> , 如果 <i>yàoshi</i> , 要是; etc. | | (43) <i>Rúguǒ hǎochī, huānyíng gòumǎi!</i> 如果好吃, 欢迎购买! [If delicious, welcome purchase!] If it is delicious, you are welcome to buy it! |
| | Sufficient condition | Overt conjunctions; <i>zhǐyào</i> , 只要; <i>dànshǐ</i> , 但使; etc. | (44) <i>Zhǐyào nǐ tóngyì, wǒ jiù qù.</i> 只要你同意, 我就去。 [As long as you agree, I just go.] I will go as long as you agree. |
| Condition | Necessary condition | Overt conjunctions; <i>zhǐyǒu</i> , 只有; <i>chúfēi</i> , 除非; etc. | (45) <i>Zhǐyǒu héping gòngchǔ, shìjiè cáinéng wěndìng.</i> 只有和平共处, 世界才能稳定。 [Unless peace coexist, world then stable.] The world cannot be stable without our living in peace. |
| | Noncondition | Overt conjunctions; <i>wúlùn...yě</i> ..., 无论……也……; <i>bùguǎn...dōu</i> ..., 不管……都……; <i>búlùn</i> , 不论; <i>rènpíng</i> , 任凭; etc. | (46) <i>Háizi jīnhòu gēn wǒmen shēnghuó, bùguǎn nǐ tóngyì bù tóngyì!</i> 孩子今后跟我们生活, 不管你同意不同意! [The child from now on with us live, no matter you agree or not agree!] The child will live with us from now on, no matter you agree or not! |
| Cause and effect | Overt conjunctions; <i>yīnwèi</i> , 因为; <i>yóuyú</i> , 由于; <i>yīncǐ</i> , 因此; <i>yīn'ér</i> , 因而; <i>yīn wèi</i> ... <i>suǒyǐ</i> ..., 因为……所以……; <i>zhī suǒyǐ</i> ... <i>shì yīnwèi</i> ..., 之所以……是因为……, etc. | | (47) <i>Yīnwèi tā shèngxíng yú Táng dài, suǒyǐ chēng tángsāncǎi.</i> 因为它盛行于唐代, 所以称唐三彩。 [Because it was popular in Tang Dynasty, so called Tang Tri-colored.] It is called Tang Tri-colored since it was popular in Tang Dynasty. |

| | | |
|---------------------|--|--|
| Purpose | Overt conjunctions; <i>wèile</i> , 为了; <i>yǐbiàn</i> , 以便; <i>miǎnde</i> , 免得; <i>shěngdé</i> , 省得; <i>yǐmiǎn</i> , 以免; etc. | (48) <i>Duō wèn jǐ jù, miǎnde zǒu cuò lù.</i> 多问几句, 免得走错路。 [More ask several sentence, avoid walk wrong way.] Ask more so as to avoid choosing the wrong way. |
| According variation | Overt conjunctions; <i>yuè... yuè...</i> , 越……越……; <i>nǎlǐ</i> , 哪里; <i>shéi</i> , 谁; etc. | (49) <i>Pá de yuè gāo, shuāi de yuè téng.</i> 爬得越高, 摔得越疼。 [Climb DE higher, fall down DE more painful.] The higher up, the greater the fall. |
| Time | Covert conjunctions | (50) <i>Xiě wán zuòyè, tiān kuài liàngle.</i> 江涛从贾老师那里走出来, 天快黑了。 [Tao Jiang from Teacher Jia there walk outside, the sky soon dark-PRT.] When Tao Jiang walked out from Teacher Jia's, it was getting dark. |

As aforementioned, when introduce and apply the concepts of coordination and subordination from English to Chinese clause combining analysis, the categorizations were however revised and some subcategorization were replaced or amended according to the different typological features of Chinese.

First, coordination and subordination classification for English based on the semantic relation as well as formal position of the clauses, while Chinese coordination and subordination is distinguished by mostly the semantic relationship between the clauses. English compound sentence consists of two or more independent clauses joined either by a semicolon or by a comma working in concert with a coordinating conjunction are defined as coordination. English subordination is one or more clauses dependently subordinate to a main clause by subordinating conjunctions or relative pronouns. In contrast to the coordination and subordination distinguish in terms of formal coordination differences, Chinese coordination and subordination is distinguished by mostly the semantic relationship between the clauses. Coordinations include clauses of equal semantic weight, while subordinations include clauses of different semantical importance.

As a meaning-driven language, Chinese is coded in direct correspondence to meaning; English, at the other end, is a form-driven language (Pan, 2002). As for Chinese, the speakers' conceptualization of the external world is directly shown in the language form. English speakers' conceptualization of the external world, on the other hand, is not directly reflected in the English linguistic form, but is seen via another layer of grammatical form added when English is morphologicalized. Given such language coding differences, the categorization of syntactic structure differs as well. In Chinese,

semantics plays a crucial role at analyzing the relation between combining clauses. Semantic and logical relations instead of formally position count the most in categorizing Chinese complex sentences which include two or more clauses. Though those commonly used conjunctions are often used to help identify the relation between the clauses in one complex sentence, overt conjunctions are not mandatory. It is perfectly natural for complex sentences to not have any overt conjunction. The semantic relation between the clauses, as provided by a clear context, is used to identify the coordination or subordination structure of a complex sentence.

One example of such mismatched classification basis of coordination and subordination between English and Chinese lies in the categorization of clauses of conversion. For instance, *but* is used to connect two or more English coordinated clauses which are treated as formally equal clauses, whereas *although*, *even though*, *though*, *whereas*, and *while* are used to connect adverbial clauses in subordinations of contrast. By contrast, Chinese does not differentiate clauses combined with *but* from clauses combined with *although*, *even though*, *though*, *whereas*, and *while*. While *but* in English indicates coordination of two or more formally equal clauses, its equivalent in Chinese, *dànshì* (但是, but), is analyzed to connect one main clause and one subordinate clause. In Chinese, the clause got negated, possibly marked with *suīrán* (虽然, *although*), is considered information background therefore the subordinated clause, whereas the clause with *dànshì* (但是, but) which raised conversion is considered as the foreground and holds more semantic weight therefore the principle clause. Another instance of the mismatched classification between English and Chinese lies in the distinguishment between *so* and *because*, *since* in English, in contrasting to the same treatment of *suǒyǐ*

(所以, so), *yīnwèi* (因为, because), and *yóuyú* (由于, since) in Chinese. In English, *so* is used to connect two or more English coordinated clauses which are treated as formally equal clauses, while *because*, and *since* are used to connect adverbial clauses in subordinations of reason or cause. In contrast, Chinese does not differentiate clauses combined with *so* from clauses combined with *because*, and *since*. The equivalents of *so*, *because*, and *since* in Chinese, *suǒyǐ* (所以, so), *yīnwèi* (因为, because), and *yóuyú* (由于, since), are all analyzed to connect one main clause and subordinations of cause and effect in Chinese.

Second, given such different classification basis in distinguishing coordinations and subordinations, the subcategorization of coordination and subordinations between two languages is therefore not corresponding either. For example, as shown in Table 2, English subordination subsumes adverbial clauses of time, reason or cause, purpose or result, condition, contrast, choice, place or location, and manner. Chinese subordination subsumes clauses in the relationship of conversion, supposition, condition, cause and effect, purpose, according variation, and time. Clauses of choice relationship are categorized as subordination in English, however, they are categorized as coordination for Chinese. Clauses of time in English are categorized as subordination. However, these type of subordinated clauses introduced with conjunctions of “*After, before, once, since, until, when, whenever, while*” are mostly categorized as coordination of continuity for Chinese. These differences between English and Chinese are related the different coding principles and strategies of these two types of languages.

It has been long criticized that modern Chinese linguistics research has been reviewing Chinese language through the lens of western linguistics, without considering

the unique traits of Chinese. It is advocated that, owing to the fundamental difference in coding systems between Chinese and Indo-European languages, a linguistic theory should be developed for Chinese's own sake. More sememic-oriented syntax research, like the Principles of Temporal Sequence and Temporal Scope (Tai, 1985, 1993, 2002), are gaining recognitions in the field.

2.2.2.3 Clarifying parataxis, hypotaxis, coordination, and subordination

In addition to the different classifications of the same domain that have been discussed above, there are different interpretations for the same terms. Though the same terms are used, there is a good deal of variation in terminology, interpretation, and implication for clause combining analysis. For parataxis, hypotaxis, coordination and subordination, there is overlapping among their definitions and uses. The terminological lack of clarity also reflects a mixture of definition and application of the different ways of combining clauses. Such mixture of definition and application causes confusion that needs to be carefully clarified before discussion.

Systemic functional grammar works distinguish embedding from clause combining or clause complexing. Embeddings function within a clause as a constituent like a complement, an adverbial, a modifier of the head noun, or a restrictive relative clause. Systemic functional grammar distinguishes between parataxis and hypotaxis in terms of interdependence (Halliday & Hasan, 1976; Halliday, McIntosh, & Strevens, 1964; Halliday & Matthiessen, 2004). In such classification, parataxis includes coordination, apposition, and quoting, while hypotaxis included non-restrictive subordination with relative clauses, clauses of reported speech, and head-dependent

subordination which was also called enhancing hypotaxis. Matthiessen and Thompson (1988) clarified that complement clause and relative clause belong to clause embedding instead of clause combining. Refusing to use the terms coordination and subordination, they employed parataxis, hypotaxis, and embedding to define three clause-combining types. By differentiating nucleus-satellite and nucleus-nucleus relations between clauses, they suggested hypotaxis and coordination are two degrees of clause combining.

Hopper and Traugott (2003) put parataxis and hypotaxis with subordination in a cline of clause-combining constructions as three “cluster points” showing the ascending degree of dependence of two clauses. Clauses in a parataxis relation were more independent and not embedded. As it is shown in Figure 3, clauses in a hypotaxis relation are interdependent and not embedded, where clauses in a subordination relation are the most dependent with one embedded in another. Within the continuum of clause combining and integration, at the end of parataxis the combined clauses both take the role of nucleus, while at the end of subordination, one clause will be the margin of the other nucleus clause. Furthermore, the end of parataxis holds minimal integration and maximal

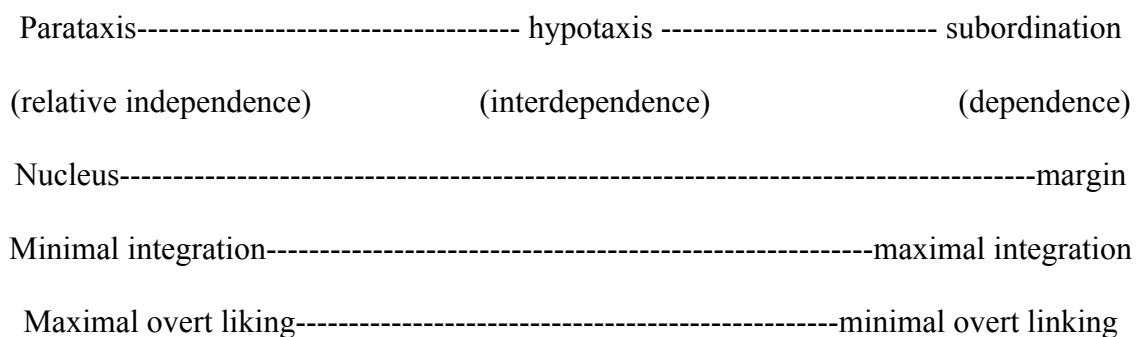


Figure 3. Properties relevant to the cline of clause combining.

(from Hopper & Traugott, 2003, p.179)

overt linking, while contrastively, the end of subordination holds maximal integration and minimal overt linking.

The current definition and use of parataxis, hypotaxis, coordination, subordination and embedding are mixed and calls for differentiation in the following aspects. First, clarification is needed between embedding and subordination. Parataxis, hypotaxis, and embedding were employed in Matthiessen and Thompson (1988) to define three clause-combining types. In the cline of parataxis, hypotaxis, and subordination in Hopper and Traugott (2003), subordination is placed as equivalent to embedding. However, subordination in their definition does not all fall in embedding since subordination includes non-restrictive subordination and restrictive subordination that is not embedded in another clause. To identify subordination with embedding leads to more mixture and confusion of different terminologies. Second, parataxis and hypotaxis are not equivalents of coordination and subordination. Clauses in coordination can take the form of parataxis with unmarked conjunctions as well as hypotaxis with marked conjunctions. When giving types of coordination, Li, Cheng, Foster, Ho, Hou, and Yip (1989) presented coordination with coordinating conjunctions and coordinating disjunctions, each of which is subdivided into coordination with unmarked connectors and marked connectors. In addition, it's not only coordination that can be presented in the form of parataxis, subordination can also be paratactically connected with covert conjunctions. Paratactic subordination is overlooked because the English prototypical subordinations connected with overt connectors have a very ample and precise conjunction, relative pronoun, and adverb system. In Chinese, however, it is actually sometimes preferred with parataxis in a form of covert conjunction clause combining (Li, 2005). Such covert conjunction can

take the form of rising intonation, pause (indicated by comma), interjection particles, and other such strategies.

From a cross-linguistic perspective, Matthiessen and Thompson (1988) noticed an interesting consequence under the claim that clause combining has evolved as a grammaticalization of the rhetorical units in discourse defined by rhetorical relations. Following this I, different languages with different discourse-based characteristics shall therefore be able to differ as well in clause combining mechanisms. In other words, Chinese as a topic-prominent and parataxis-prominent language does not have to mechanically apply the clause combining analyzing system invented for English; instead, a tailor-made analysis might bring different insight.

Based on the clarification of parataxis, hypotaxis, coordination, subordination, and embedding, as well as the logic clause combining has evolved as a grammaticalization of the rhetorical units in discourse, this dissertation for the first time presents a taxonomy of clause complexing in consideration of the different clause combining mechanism for languages of different typological features. As shown below in Figure 4, in this taxonomy *clause complexing* subsumes both clause embedding and clause combining. *Clause embedding* refers to those clauses which function as subordination to part of the matrix clause. *Clause combining* concludes both coordination and subordinations in which those clauses function as subordination to the matrix clause. For any of clause embedding, subordination and coordination, clauses can be either paratactically or hypotactically connected. However, different languages may enjoy different mechanisms for clause combining. Both coordination and subordination are clause-combining units for subject-prominent languages, which are not necessarily

appropriate for topic-prominent languages such as Chinese. As is shown in Figure 4, this dissertation adopts the topic chain as another form of clause combining, and hereby proposes the taxonomy of clause complexing in consideration of the different clause combining mechanism for languages of different typological features including topic-prominent languages like Chinese. Under clause combining, there includes, but not limited to, subordination, coordination, and topic chain. How the subservient unit of topic chain compose a topic chain is considered clause combining, and clause embedding takes place within each subservient unit of topic chain. In addition to the topic chain, other possible forms of clause combining may be included according to the typological difference of other languages.

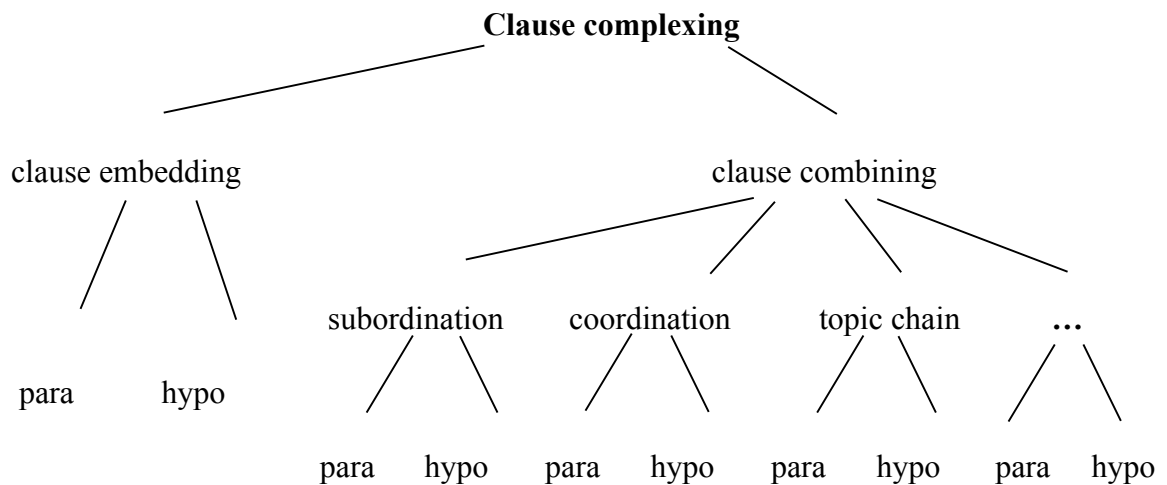


Figure 4. Taxonomy of clause complexing.

Note. “Para” here is the abbreviation of paratactically; “hypo” here is the abbreviation of hypotactically; “...” stands for other possible forms of clause combining.

Since the notion of topic is a discourse-based one, topic chain is argued to be a discourse unit as Chu (1998) called topic-comment units a “discourse sentence” or “SENTENCE” (all capital letters). However, due to the controversy of the sentence as a syntactic unit for Chinese as well as to the arbitrariness of sentence boundary, this dissertation presents the topic chain as a sentential level unit for Chinese and topic-comment structures that compose topic chains as a clausal level unit, in order to be consistent and comparable with coordination and subordination in English. For the purpose of Chinese syntactic complexity analysis, global complexity refers to the overall use of the topic chain which can be measured via the mean length, proportion, and compositionality of such unit of production. Clausal complexity detects the number and the nature of the composing unit of topic chain and their constituent relationship in a unit of upper-clausal level. Subclausal complexity via phrasal elaboration is measured by mean clausal length and nature of such elaboration. However, the clausal and subclausal units might be composed into their upper-level unit in a very different way from Indo-European language.

2.2.2.4 Covert connectives in Chinese

Since Chinese is a parataxis-prominent language, overt conjunctions are not mandatory for complex sentences as they are for coordination or subordination in English. Clauses can be in juxtaposition without explicitly indicating their structural relationships. Grammatical judgment tests were conducted among Chinese native speakers, and as long as context was clear, they preferred to use sentences with covert conjunctions (Li, 2005). Therefore, we can extrapolate that native Chinese speakers

prefer to use covert instead of overt clause subordination when the context provides sufficient semantic information. For example, the following sentences (51) and (52) illustrated in Jin (2006, p. 122) share the same meaning. While the conjunctions *rúguǒ* (如果, if) and *yīncǐ* (因此, therefore) in (52) are overt, those conjunctions in (51) are covert. (51) is understood completely without the additional conjunctions added in (52). In fact, the clauses in (52) sounds closer than those in (51) separated by the conjunctions. Clauses in juxtaposition like in (51) imply hidden logic forms instead of explicitly marking the coordination or subordination. Such preference of covert connectors contributes to the argument that Chinese is a meaning-driven language instead of a form-driven language. It can also be considered reflecting the typical characteristics of a Chinese mode of thinking.

(51) *Rúguǒ chūle wèntí, dānwèi jiāng quánmiàn fùzé jiějué, yīncǐ, gèrén bùbì cāoxīn.*

如果出了问题_i, 单位将全面负责解决 \emptyset_i , 因此, 个人不必操心 \emptyset_i 。

[If comes-out PRT problem_i, the work unit will completely take in charge solve \emptyset_i , therefore, individual no need worry \emptyset_i .]

If there is a problem, the work unit will be completely responsible for it.

Therefore individuals do not have to worry about it.

(52) *Rúguǒ chūle wèntí, dānwèi jiāng quánmiàn fùzé jiějué, yīncǐ, gèrén bùbì cāoxīn.*

出了问题_i, 单位全面负责解决 \emptyset_i , 个人不必操心 \emptyset_i 。

[Come-out PRT problem_i, the work unit will completely take in charge solve \emptyset_i , individual no need worry \emptyset_i .]

If there is a problem, the work unit will be completely responsible for it. Therefore individuals do not have to worry about it.

2.2.3 Sentence-oriented versus discourse-oriented

The sentence boundary in Chinese is arbitrary. The sentence in Chinese is a notion with no viable structural definition and no generally accepted criteria. A sentence is completed when there is a major break needed in the flow of thought. In terms of writing, full stops are used to mark the end of a sentence. In classical Chinese, there were originally no punctuation marks marked by the author in text like English does. Readers add punctuated breaks in the text to separate clauses and sentences to a certain extent based on functional words, which indicate pause and an understanding of the general idea of the whole text. Therefore, there could be different versions of punctuation for the same classical text depends on different understanding or interpretation. In a study conducted by Tsao (1990), 18 Chinese ESL college students were asked to apply punctuation marks in two Chinese written passages and two English passages, where the original punctuation marks have been removed. Interestingly, the results showed that the students, who are all native speakers of Chinese, disagreed considerably both among themselves and with the original author as to the numbers of sentences contained in the Chinese paragraphs. Contrastively, in their punctuation of the two English texts, these Chinese native speakers, who were far from having a native command of English, showed considerably more agreement among themselves and with the original author about sentencehood. A paragraph-based punctuation study by Chu (1998) confirmed a consistent result. Therefore, it is suggested that the sentence in Chinese seems to be an

arbitrary unit. A Chinese sentence is thus probably unreliable when applied as the segmentation unit for Chinese syntactic complexity analysis.

To analyze the Chinese syntactic complexity, in this dissertation, topic-comment structure is used as the unit of analysis instead of the conventional sentences. The notion of topic is a discourse-based one, and it may, and often does, extend its semantic domain to more than one punctuational sentence. Chu (1998) argued a Chinese sentential unit is a “discourse sentence” or “SENTENCE” (all capital letters) which vastly differs from a Western sentence. A discourse-oriented topic chain not only does not necessarily share the same boundary with a conventional sentence. As we discussed in 2.2.1.2, *coreferential zero* is proposed in this dissertation to help identify the starting and end points of a topic chain. A coreferential zero refers to an element that does not have any phonological content and is unpronounced, but corefers to the topic mentioned in previous or subsequent consecutive clause(s). Such a coreferential zero does not have to be a noun or nominal, but is used in a topic chain to repeat the same topic with a phonetically null form. When successive topic-comment structures share the same topic with the topic only overtly stated in its full form once and the rest repeated in the form of coreferential zero, a topic chain thus forms. If the same topic is not repeated as a coreferential zero but with the full form, a pronoun, or demonstrative, instead of continuing the topic chain by keeping the referent in active memory, a new topic-comment structure or topic chain is then activated. If a different topic is introduced, a new topic chain is then activated.

A topic-comment structure or a topic chain may not always share the same demarcating boundary with the conventional sentence. A topic chain can go beyond the

period and be longer than a sentence marked with punctuation marks. A topic chain can as well be only part of a sentence. The aforementioned Example (10) (See section 2.2.1.3) is an example of two complete sentences that form one topic chain. Each topic-comment structure shares the same topic *wǒ* (我, I) which has the phonological shell at the second sentence but takes the form of coreferential zero in all the topic-comment structures in the first punctuational sentence. According to Li (2005)'s categorization, *wǒ* (我, I) is a cataphoric topic for this one topic chain which consists of two sentences.

In addition to the ten types of topic chain proposed by Li (2005), there is a special type of topic chain proposed by Chu and Pan (2006, p. 255-8): telescope chain. In a telescope chain, two or more topic chains are intertwined. Such topic chains are connected like the telescope in which one smaller section can slide inside another somewhat larger section. Two or even more chains can be intertwined in a telescope chain. (53) was an example cited in Chu (2006, p. 255-8), in which the first half of this topic chain with shared topic *tā* (他, he) ranges from (a) – (b) and crosses the sentence boundary. The second half of this chain with shared topic *lǎozhě* (老者, an old man) ranges from (c) – (e) and is only part of the second sentence. Clause (b) was considered the hook which links the two separated topic chain and connects them together into one topic chain. The boundary of these two topic chains mismatched with the boundary of the two sentences. The first period marked the end of the first sentence at the end of clause (a), however, the first topic chain goes beyond the first sentence and ended after clause (b).

(53) ...*tā zhīdào zhǐyǒu zhèyàng cái zúyǐ jiǎnshǎo cūnrén de huàiyí. Zuòle yíhūr, yuàn zhòng chūlái gè lǎozhě, lán bù xiǎoguà changzhe huái, liǎn shàng hěn liàng, ...* (p.255)

.....他_i知道只有这样才足以减少村人的怀疑。∅_i坐了一会儿，院中出来个老者_j，∅_j蓝布小褂敞着怀，∅_j脸上很亮，.....

[... (a) he_i know only this-way can sufficient reduce villagers DE suspicion. (b) ∅_i sit PRT a-while, (c) yard-in come-out an old man_j, (d) ∅_j blue shirt open-in-front, (e) ∅_j face-on very bright.]

...he knew this was the only way to lessen the villagers' suspicions. An old man came out of the impressive house after Hsiang Tzu had been sitting for a while. He wore a blue gown and a short jacket. His shiny face looked friendly....

For syntactic complexity analysis in this dissertation, however, (49) is considered as two separate topic chains regardless of the demarcating boundary of the conventional sentence. In other words, (53) is not categorized as a telescope chain because this dissertation argues that the boundary of a topic chain is not restricted by the sentence boundary. The first topic chain in (49) covers (a) – (b) that share the topic *tā* (他, he), whereas (c) – (e), which share a different topic *lǎozhě* (老者, an old man) form a second topic chain. A topic chain can be part of or go beyond the sentence boundary marked by the punctuation: the first topic chain (a-b) in (49) goes beyond one sentence while the second topic chain (c-e) is only part of one sentence.

Different from (53), (54) below is a good example of the telescope chain according to the operationalized definition of a topic chain in this dissertation. In (54), *tā* (他, he) and *sùshè* (宿舍, dorm) are the topics of two intertwined chains. As we discussed

above, the topic chains do not always share the same demarcating boundary with a conventional sentence, as a topic chain can be only part of a punctuational sentence as well as go beyond the period and extend its boundary outside a punctuational sentence. If the topic is repeated in the form of coreferential zero decides whether the topic-comment structures that share the same topics form a topic chain. As illustrated in Figure 5, the first topic *tā* (他, he) takes its full form once in (a) and is repeated in the form of coreferential zero in (b), (c), and (e). Meanwhile, clause (a) shares a second topic *sùshè* (宿舍, dorm) with (d) and (f). *Sùshè* (宿舍, dorm) takes its full form in (a) and is repeated in form of coreferential zero in (d) and (f), it therefore is picked up to be the topic of a second topic chain. Therefore, one topic chain of a-b-c-e is intertwined with another topic chain of a-d-f. Clause (a) presents both two topics and functions as the link of the two topic chains. Unlike (49) in which there is a clear cutting point between topic chain of a-b and topic chain of c-d-e, in (54) the two topic chains are actually intertwined and cannot be cut at any point between the clauses. (54) therefore can be seen as a telescope chain. The topic chain boundary in (54) also does not fall along the punctuational sentence boundary. Two topic chains in (54) overarched three punctuational sentences that were marked by the two exclamation marks and one period.

(54) *Tā jiù juéding háishì xiǎng bànfǎ xíguàn sùshè ba! Tài chǎo kěyǐ yòng ǎrsāi, fángjiān tài xiǎo méiguānxi, kěyǐ sòng diào yìxiē tā de shū hé yīfú. Méi dìfāng zuò fàn zuì bù chéng wèntí, zài zhōngguó shàng jiē chīfàn kě fāngbiànle!*

他_i就决定还是想办法习惯宿舍吧_j! \emptyset_j 太吵可以用耳塞, \emptyset_j 房间太小没关系, \emptyset_i 可以送掉一些他的书和衣服。 \emptyset_j 没地方做饭最不成问题, \emptyset_i 在中国上街吃饭可方便了!

[(a) He_i then decided still think method get used to dorm_j! (b) \emptyset_j Too loud can use earbuds, (c) \emptyset_j room too small doesn't matter, (d) \emptyset_i can give away some he DE books and clothes. (e) \emptyset_j No places to cook most not become problem, (f) \emptyset_i at China street eat very convenient!]

He then decided he'd better work to get used to life in the dorm! He can use earbuds when it gets too loud. He can also give away some of his books and clothes since the room is too small. It is the least problem that there is no kitchen. Because it is so very convenient get food on the street in China.

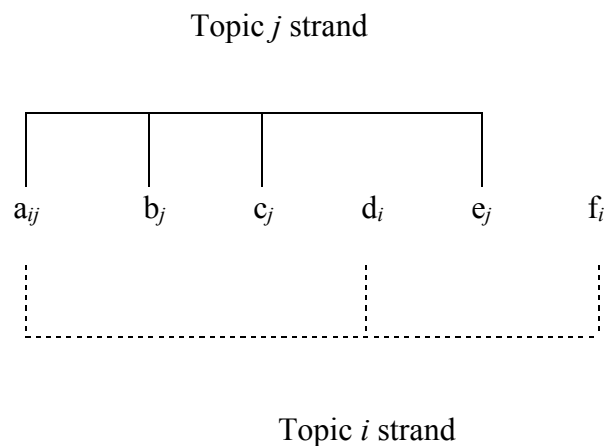


Figure 5. Illustration of the structure of Telescope Chain (54).

2.3 Linguistic complexity and task influences

2.3.1 Cognitive task complexity

Besides the linguistics complexity, another dimension of cognitive task complexity is along the line of researches on testing two rival models of task complexity as they affect language performance: the *Limited Attentional Capacity Model* (Skehan, 1998) and the *Multiple Resources Attentional Model* (Robinson, 2001, 2005). These two models conflict in whether L2 language complexity goes up along an increase of cognitive task complexity at the price of lower L2 accuracy, or whether L2 language complexity and accuracy actually go up simultaneously. Both models, however, hypothesize that L2 language complexity goes up along an increase of cognitive task complexity. By manipulating these cognitive factors that are required in completing a task, such as the amount of describing elements, reasoning demands, planning time of the tasks, etc., cognitive task complexity can be adjusted and sequenced. As part of the triad of task complexity, condition and difficulty factors, Robinson (2001) categorized the cognitive factors into two dimensions: resource-directing and resource-depleting (also addressed as resource-dispersing). Resource-directing dimension variables make reference to conceptual cognitive demands whereas resource-dispersing dimension variables refers to performative/procedural cognitive demands. Along the line of resource-directing, factors are such as element amount, time and place, and reasoning demands. Tasks with less elements to be described and distinguished, requiring simple description of events happening now and here, with no consuming of attentional, memory and reasoning resources are less cognitively complex comparing to those tasks which require description of more elements, events happened in the past, and demand more reasoning effort. Along the other line of resource-dispersing, factors include task planning, single or dual task, and learners' prior knowledge, etc. With more planning

time given, prior knowledge available, and only a single activity required, such tasks are less cognitively complex comparing to those tasks with less planning time, no prior knowledge, and dual task conduction required.

When investigated how cognitive task complexity affects the CAF of language output, more research chose to operate the resource-directing dimension factors and not so much operated on the resource-dispersing dimension, much less both resource-directing and resource-dispersing dimension factors in one study. The manipulation of the resource-directing dimension variables more directly corresponds to particular form-meaning mapping, which plays an important role in sequencing tasks predicting the language code produced “such as conjunctive coordinators to establish causality, past tense morphology and temporal expressions, and complex nominalizations to distinguish numerous similar elements” (Robinson, 2003, p. 648). In contrast, resource-dispersing dimension variables affect attention to various dimensions of languages, which matches more to task as a timeline in a pedagogical cycle. For one task planned to be conducted in class, resource-dispersing dimension variables shed insights on the actual enactment of a task, such as what background information should be provided and how to present it to students, how much planning time and what type of planning should be designed, and what form of task conduction should be chosen, etc.

Among the resource-dispersing dimension variables, planning has been mostly discussed. Both planning time and planning type has been empirically studied yet debated conclusions were drawn.

Forster and Skehan (1996) found planning to be a positive effect on the complexity of L2 speech performance. Mehnert (1998) focused on the length of

planning time by dividing English speaking German learners into four groups respectively with 0, 1, 5, and 10 minutes of planning time. For the three groups with planning time, it was reported that instruction was as well provided but it was not elaborated in the article that what kind of instruction was provided. In term of language complexity among the speech output produced with increased planning time, which was measured by words per c-unit, number of subordinate clauses, and number of s-nodes, no statistically significant difference was found. The speech output complexity was much higher for the 10-minute planning condition comparing with the nonplanned condition, yet the 1- and 5-minute planners showed same or even slightly lower complexity than nonplanners. Substantial planning time seemed to play a role in increasing the speech output complexity.

However, in classroom teaching context, it is not often applicable to provide substantial planning time as restricted by the pressure to follow along the tight teaching schedule. In a test condition, no hours or days interval is allowed. What's more, the substantial planning time also was questioned to deminish the authenticity of task since there is unlikely 10 minutes planning time in real conversation. In a testing context for L2 English speakers who take *Australian Assessment of Communicative English Skills* (ACCESS) in order to apply for an Australian visa, Wigglesworth (1995, 1997) investigated how limited planning time effects on complexity. With 1-minute planning time provided before each speaking task, she found that “planning time may allow higher-proficiency-level candidates to produce more complex language in the more difficult tasks” (Wigglesworth, 1997, p. 95).

As one type of task planning, task repetition has been utilized in language teaching classroom, especially for speaking, in order to push out better performance. It can be that the complete task or part of the task gets repeated, for twice or more times. The interval between two enactments of task performance can vary from a few hours to weeks, or even months (Bygate, 1996; Gass, Mackey, Fernandez, & Alvarez-Torres, 1999; Lynch & McLean, 2000; Bygate, 2001). One of the rationales behind the use of task repetition in classroom teaching is the Levelt's (1989) model of speech production. According to Levelt (1989), there are three sequential but overlapped stages speakers go through when produce any speech: conceptualizing the idea, formulating the language representation, and articulating the language form. Following the information processing theory, human being's attentional capacity is limited and selective (Anderson, 1995; Schmidt, 2001). When we repeat the task, since the conceptualization stage has mainly accomplished at the first task enactment, more attentional focus can be shifted to the language formulating and articulating. Positive effect of task repetition on language complexity has been empirically confirmed (Gass et al., 1999; Bygate, 2001). In their repetition, "part of the work of conceptualization, formulation and articulation carried out on the first occasion is kept in the learners' memory store and can be reused on the second occasion" (Bygate, 2001, p. 29). Along such reasoning, immediate task repetition can maximally store the learners' memory of their initial task enactment and therefore maximally take advantage of the form of task repetition. What's more, during the interval of weeks or months, there can be other factors come into play contributing to learners' language level besides task repetition, whereas immediate task repetition can potentially avoid such distracting factors.

2.3.2 Learner affective variables

In line with Spilsbury, Stankov and Roberts (1990), Robinson (2001) distinguished cognitive task complexity from task difficulty. Task difficulty can be interpreted as how learners think themselves can do, affective variables, and what learners are capable to do, ability factors. Learners' ability variables are learners' aptitude, proficiency, and intelligence, etc. Learners' affective variables include learners' anxiety, motivation, and confidence, etc.

Research has shown that L2 learners' perceived competence and communication anxiety affect their willingness to communicate (WTC). Zhou (2012)'s empirical study showed that "a language learners' perceived competence directly predicts their willingness to communicate and their communication anxiety contributes to the willingness to communicate indirectly through perceived competence" (p. 166). Learners' perceptions of their own L2 competence and demands of the task affect their anxiety in communication. Learners who perceive themselves of higher L2 language competence will be less anxious to use L2. Such learners will be more willing to speak therefore possibly get more opportunity to practice which leads to their actual L2 competence increase. While students who perceive themselves of lower L2 competence as well as hold higher debilitating anxiety are therefore less willing to communicate, which leads to less or slower L2 competence grow.

Task repetition as a form of task planning is hypothesized to lower L2 learners' communication anxiety by upgrading their self-perceived performance in the second enactment. Though L2 learners' language competence is not expected increased within one test or one class, their self-projection of the repeated performance can be greatly

raised. Instead of reinventing the wheel, when repeat the same task L2 learners have prepared themselves with what they have accomplished across the three stages of conceptualization, formulation and articulation from the first enactment. The first completion of task therefore functions as the rehearsal. With such, learners are expected to be more confident and ease at the second completion. Taking away the satisfaction from self-perceived progress, learners can be more motivated to communicate outside the classroom. While the immediate task repetition is conducted after the online review of their initial task performance, it is supposed to be perceived as a better performance.

2.3.3 Immediate task repetition for task planning

This dissertation introduces immediate task repetition as a form of task planning to operationalize the task planning variable along the resource-dispersing dimension. Right after the initial enactment of task completion, the participants are granted a second chance in order to improve their performance. Such immediate repetition of task is presumed to meet a collection of following expectation. First, with the initial task completion functioning as a rehearsal, what participants developed in their initial enactment across all the three stages of conceptualization, formulation, and articulation can be maximally kept in their short-term memory store and reused in their immediate repetition. When repeating a task, the task performers are able to load less cognitive attention in idea development, structure organization, and lexical and grammatical choosing, but focus more attention in expanding, restructuring, or polishing at lexical, phrasal, clausal, and sentential level to achieve higher complexity. Right after the initial trial, task performers are presumed to hold the freshest memory from their own

performance, as well as a first-hand self-evaluation which helps guide further self-correction and autonomous improvement. Throughout the first trial and after it ends, the L2 learners, especially advanced learners, run an online review of their own performance, which directs them how to improve their own language right after the first trial. Second, with much higher time efficiency, immediate task repetition can be more widely used in classroom teaching and applicable in testing context. The interval between two-time task enactments does not have to be days nor weeks. The repetition can start right after the initial trial, providing students ample practice without taking up too much class time. This is especially helpful since language classes are usually assigned with a lot to cover in order to keep up with the course schedule. Third, immediate task completion partly amends to the authenticity of long interval task repetition. Immediate task repetition functions as self-assessment and self-correction for a longer piece of language output. Last, immediate task repetition is also hypothesized to lower L2 learners' communication anxiety and improve their self-perceived performance. With the first task completion functions as a rehearsal, L2 learners are expected to perceive themselves better prepared therefore lower their communication anxiety. The immediate second chance within a grasp is also supposed to better motivate task performers when repeat the same task.

CHAPTER 3

OPERATIONALIZING CHINESE SYNTACTIC COMPLEXITY

3.1 T-unit as the unit of analysis

While current complexity measures are developed in a global fashion, little attention is paid to tailoring to the typological differences like topic prominence of the Chinese language, and therefore such measures are not as valid for Chinese complexity measurement as they are for Indo-European languages (Jin, 2006, 2007; Yuan, 2009). Jin (2006) adopted widely-used measurements for Indo-European languages like mean length of the T-unit and mean length of the clause to measure the complexity of Chinese language. It was found that the mean length of the T-unit or clause went up as the L2 Chinese speakers' proficiency level went up, but dropped down when it was used to measure the native speakers' output. Yuan (2009) confirmed Jin (2006)'s conclusion that the T-unit may not be a reliable indicator for Chinese, finding that native speakers, when compared with L2 speakers, produced shorter T-units and lower clause/T-unit ratio. Unfortunately, providing only limited examples, no current available studies have proposed with clarity an operationalized definition of T-unit or topic chain for Chinese syntactic complexity analysis. For the first time, Jiang (2013) proposed a working definition of the T-unit in Chinese:

A single main clause that contains one independent predicate plus whatever other subordinate clauses or non-clauses are attached to, or embedded within, that one main clause. (p. 5)

However, given this working definition, it was still not clear how the T-unit was operationalized. It was especially unclear how the subordination clauses were defined and operationalized in this definition, with no correspondent coding examples provided. One example Jiang (2013) provided was a compound sentence, (55), which was simply analyzed as three T-units and not clarified if there was subordination included.

(55) *Wǒ jiějie jiào mǎlì, jīnnián èrshí suì, zài běijīng shàng dàxué.* (Analyzed as 3 T-units)

我姐姐叫玛丽，今年二十岁，在北京上大学。

[I elder sister call Mary, this year twenty year, at Beijing go to college.]

My elder sister is called Mary. She is twenty this year. She goes to college in Beijing. (p. 9)

Provided with this working definition, the length of T-unit was attempted to be an as reliable measure for Chinese syntactic complexity. Such a definition of T-unit, while seemed hands-on, was however reductionist. Many empirical CAF studies have taken “a rather narrow, reductionist, perhaps even simplistic view on and approach to what constitutes L2 complexity” (Bulté & Housen, 2012, p. 34). In this case, first, there was neither explanation nor illustration of how subordination in Chinese is defined. Since Chinese language is paratactic-prominent and lacks overt connectives, a clear definition of subordination is required to operationalize such a working definition. Especially given the mismatched subcategorization of coordination and subordination between English and Chinese, it requires more articulation how subordination is defined in Chinese in order to apply Jiang’s working definition of T-unit. Furthermore, Jiang’s working

definition of T-unit defined any self-standing clause as a T-unit whether connected by conjunction words such as ‘and’ or by a comma, so that any clauses with a predicate that are connected via covert connective would be analyzed as individual T-units. In this way, the ratio of clause amount per T-unit for any written output can potentially be equal to a constant 1. Third, while there are three levels of complexity: global complexity, clausal complexity and subclausal complexity, the mean length of a T-unit in such working definition was only able to detect clausal complexity. As illustrated in (55), the 3 T-units in Jiang’s working definition are actually three clauses that form a topic chain as the global unit of complexity. Since syntactic complexity itself is a multidimensional construct, indices that tap into complexity at different levels are required. Such aforementioned reductionist practice of Chinese complexity assessing was caused by mechanically applying the “universal” indices based on T-unit or clause in Chinese. To provide the T-unit with an adjusted definition for Chinese still looks at the Chinese language through an Indo-European language lens, failing to consider Chinese syntactic features themselves as a standpoint. In the end, such working definition of T-unit did not catch hold of the essential typological structural difference between Chinese and Indo-European languages.

Most importantly, applying T-unit for Chinese complexity assessing in written Chinese with explicit punctuations averted the question of blurred sentence boundaries in Chinese. With sentence boundaries pre-marked by punctuation marks in written Chinese output, sentence may still manage to function as the basic unit of analysis. However, for spoken Chinese with no explicit punctuation marks provided and of which sentence boundaries can be varied even for different native speakers, applying sentence as the basic

unit of analysis is unavertably problematic. As we discussed in Section 2.2.3, the sentence in Chinese is a notion with no viable structural definition and no generally accepted criteria. Even native Chinese speakers showed considerable disagreement in punctuating the same written piece. With no viable sentence boundaries set, the starting and end points of each T-unit for even the same spoken Chinese data can be indentified inconsistently among multiple ratings or different raters. Instead, as discussed in 2.2.1.2, applying coreferential zero to detect topic chain is able to breach the blurred sentence boundaries and therefore provide an operationable unit for Chinese syntactic complexity.

3.2 Other units of analysis

In addition to T-unit, a wide range of segmentation units for quantitative analysis of language output exists. An inventory of linguistic complexity measures including T-unit, turn, AS-unit, c-unit, clause, utterance, subordinate clause, dependent clauses, and (noun, verb) phrase, was summarized in Bulté and Housen (2012), which surveyed a representative sample of forty empirical L2 studies on task-based language learning and was published between 1995 and 2008. Foster, Tonkyn, and Wigglesworth (2000) reflected on the lack of a comprehensive and accessible definition of the segmentation unit on spoken language analysis and broadly categorized units of spoken language quantitative analysis into three groups: semantic, intonational, and syntactic. Semantic units include *proposition*, *C-unit* (semantic focus), and *idea unit* (semantic focus), etc. C-unit (semantic focus) can be defined as “utterance, for example, words, phrases and sentences, grammatical and ungrammatical, which provided referential or pragmatic meaning” (Pica *et al.*, 1989, p. 72). Lacking a definite formal indicator, unit

segmentations that exclusively rely on semantic criteria carry the risk of analysis reliability. Intonational units include *tone unit/phonemic clause*, *idea unit* (intonation focus), and *utterance*, etc. Pause and semantic criteria are often incorporated in addition to intonation when defining intonational units. However, when applied to analyzing the speech of non-native-speakers, the intonational features do not necessarily indicate unit boundaries due to a lack of proficiency. Syntactic units, more widely used, include *sentence*, *idea unit* (structurally defined), and *T-unit*, etc., among which T-unit is clearly the most widely applied.

Though provided with adjusted definitions tailored to varied research data, T-unit is still found inadequate when analysing complete spoken discourses especially for interactional spoken data due to its common incomplete sentences, hesitation, and repetition. AS-unit was more suitable in capturing the characteristic of spoken discourse. “An AS-unit is a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (Foster, et al., 2000, p. 365). According to such a definition, however, the AS-unit is still based on the coordination and subordination structure of language. Such a clause complexing mechanism of coordination and subordination, as illustrated in Figure 4 (see Section 2.2.2.3), is not appropriate for languages of other typological features, such as the topic-prominent Chinese language. In addition, for AS-unit a finite verb is crucial to defining a *minimal independent*, which again is not in line with the topic-prominence of the Chinese language since a Chinese sentence might not need a finite verb as exemplified in (1) (see Section 2.2.1.1).

One crucial question to be answered before applying any segmentation unit in Chinese language analysis surrounds the arbitrariness of Chinese sentence boundaries. While other complexity analysis units designed in a global fashion are segmented by sentence boundaries, the sentence in Chinese, as discussed in Section 2.2.3, is a notion with no viable structural definition and no generally accepted criteria. Of the very few Chinese syntactic complexity studies currently available, only written Chinese was analysed and bypassed the arbitrariness of Chinese sentence boundaries (Jin, 2006; Jiang, 2013), or only clause-level segmentation was provided (Yuan, 2009). Punctuation marks are provided in written texts, withstanding a lack of reliability in terms of sentence segmentation. However, adding intonation marks for spoken Chinese output can be very arbitrary. The same Chinese spoken output recording can be transcribed into texts with different intonation marks by different raters. The reliability of such an analysis of Chinese syntactic complexity is uncertain if the identification of the beginning and end points of the unit of analysis relies only on intonation marks. Segmenting the unit of analysis based solely on punctuation marks thus fails to determine behavior identification for the Chinese syntactic complexity construct conceptualization. In other words, with its unit of analysis based on the notion of sentence with no viable structural definition and no consistently adopted criteria, the validity and reliability of Chinese syntactic complexity assessing is questionable. Without a clarification on the unit of analysis segmentation, Chinese syntactic complexity unit based on the notion of sentence also hinders the operationalization of Chinese syntactic complexity in terms of coding at the stage of observation scoring. Therefore, to provide a comprehensive and accessible

definition to TC-unit segmentation in this dissertation, coreferential zero is proposed (see Section 2.2.1.2) to consistently identify the starting and end point of a topic chain.

3.3 TC-unit as the unit of analysis

With the prototypical Chinese syntactic unit is that of the topic-comment structure but not sentence, the nature of the discrete components that the language consists of is as well different from English. As we conceptualized syntactic complexity to be the number and the nature of the discrete components as well as their constituent relationship, the unit of analysis for Chinese syntactic complexity therefore should not copy sentence. Adapting the definition of topic chain (Chu, 1998; Li, 2005), Jin proposed a *Terminal Topic-Comment Unit (TTCU)* as the unit of analysis for Chinese syntactic complexity assessment. “TTCU refers to two or more sequential clauses which shares one topic. The topic only shows once in its full form at the beginning position of the first clause and takes the form of empty pronouns or empty nouns in the remaining clauses” (translated from Jin, 2006, p. 123). In addition to five T-unit based measures, (a) mean length of sentences (MLS), (b) mean length of T-unit (MLT), (c) mean length of clauses (MLC), (d) clauses per T-unit (C/T), (e) T-unit per sentences (T/S), Jin (2006) for the first time applied three additional TTCU-based measures: (f) mean length of TTCU, (g) clause per TTCU, and (h) empty category per topic chain.

Jin’s research pointed to a breakthrough in Chinese complexity defining and assessing. However, there are still many questions regarding the conceptualization and operationalization of her TTCU remain unanswered. The first question is related to identifying the topic. Since topic is a thematic role or information unit, it lacks applicable inflectional cues in order to identify the topic. Second question is how to identify the

beginning as well as the end point of a topic chain. Li's (2005) definition noted, "topic chain includes at least two sentences which are connected by an indispensable NP topic and one or more anaphorically or cataphorically omitted NP topic" (p. 67). When a topic is "anaphorically or cataphorically omitted" in the topic-comment structures, the operationalization of TTCU should state whether such consecutive omissions are mandatory. The boundary of a topic chain is also confused with a sentence boundary. For TTCU proposed by Jin (2006), there was not a clear segmentation criterion, however, the boundaries of the exemplified TTCUs provided were overlapped with the sentence boundaries indicated by punctuation marks. However, as discussed in 2.2.3, not to mention the fact that Chinese has arbitrary sentence boundaries unlike most Indo-European languages, restraint in written Chinese and adoption of the sentence boundary marked by punctuation marks, still presents a confirmed mismatch between the sentence and topic chain boundary. Spoken output, which lacks punctuation, will require a data coder that inserts punctuation in data transcribing. With the arbitrary feature of punctuation marks in Chinese language, following punctuation marks in dividing terminable TC-units can jeopardize research, harming replication and knowledge accumulation. The third question surrounds the embeddedness and compositionality of a topic chain and its relationship with other non-topic-chain output. Research has seen some proposals and amendments on defining the topic chain, yet those definitions were generally from a theoretical instead of an operationalizable perspective. Therefore scant discussion brought together the topic chains and other sentential level unit not forming a topic chain. Without clarifying the relationship between non-topic-chain output and topic chains, TTCU was applicable to the topic chains identified, and not the total language

output collected in Jin (2006). Example (56), (57), and (58) below are examples listed in Jin (2006, p. 132), with pinyin and English translation added by this dissertation. (56) was identified as 0 topic chain, (57) as 1 topic chain of 2 clauses, and (58) as 1 topic chain of 4 clauses. In addition, there are many coding level questions need to be clarified. First, it was quite confusing how the empty category is identified. In (57), the topic, *tāmen de fángzi* (他们的房子, their house), is repeated in unpronounced form for 2 times and there was identified 2 empty category; while in (58) the topic, *tāmen* (他们, they), is repeated in unpronounced form for 3 times but there was identified 1 empty category. Second, in (57) the amount of clauses within one topic chain equals to the total times that the topic got repeated in unpronounced form; while in (58) the amount of clauses within one topic chain equals to the total times that the topic got repeated in unpronounced form plus one. Last, the question that how the topic omitting was accounted also remained unaddressed. While there were 3 “∅” marked in (58), it was counted as 2 topic omitting. In order to make the topic chain operationalizable as a unit of analysis for Chinese syntactic complexity, a definition of topic chain is required that clarifies its topic identification, inner compositionality and embeddedness, and boundary with the topic-prominent, paratactic-prominent, and discourse-oriented typological features of Chinese language syntagma considered.

(56) *Tāmen de shēnghuó tiáojiàn bǐjiào xiàndàihuà, tāmen de fángzi yǒu diàn, zài fángzi li yǒu wèishēngjiān, *zhè yǒu kòng diào.*

他们的生活条件比较现代化，他们的房子有电，在房子里有卫生间，*这有空调。(0 topic chain, and 0 empty category.)

[They DE life condition relatively modernized, they DE house have electricity, at house inside have bathroom, here have air conditioner.]

Their life condition is more modernized. Their house has electricity. Their house has bathroom. There is air conditioner.

(57) *Tāmen de xīn de fángzi hěn xiàndàihuà. Tāmen de fángzi yòu yǒu diàn, *yòu yǒu wèishēngjiān, *shènzhì hái yǒu kòng diào.*

他们的新的房子很现代化。他们的房子又有电，*又有卫生间，*甚至还有空调。(1 topic chain of 2 clauses, 2 empty category, and 3 connectives.)

[They DE new DE house very modernized. They DE house also have electricity, also have bathroom, even again have air conditioner.]

Their new house is very modernized. Their house has electricity as well as bathroom and even air conditioner.

(58) *Tāmen dōu shì nóngmín, *kào zhǒng lìzǐ wéi shēng, *shēnghuó hěn kǔ, *jiālǐ lián diàn yě méiyǒu, suǒyǐ tāmen hěn nǔlì dì gōngzuò.*

他们都是农民，*靠种栗子为生，*生活很苦，*家里连电也没有，所以他们很努力地工作。(1 overt connective, 1 empty category, 2 topic omitting, and 1 topic chain of 4 clauses.)

[They all are peasant, rely on plant chestnut as living, life very bitter, house inside even for electricity also none, so they very hard DE work.]

They are all peasants who make a living by planting chestnuts. They have a very hard life. There is not even electricity at their house. So they really work hard. (p. 131)

Note: *marks the unaccurate expressions in Jin (2006).

This dissertation hereby proposes a taxonomy of TC-units, as illustrated in the Figure 6, to capture the sentential and clausal level structures in Chinese. In this taxonomy, a *terminable TC-unit*, as a sentential level unit, is the unit of analysis for Chinese syntactic complexity. As a clausal level unit, a single TC-unit in Chinese syntactic complexity analyses refers to each individual topic-comment structure. A terminable TC-unit refers to a minimal terminable single TC-unit or a topic chain. Depending on the amount of single TC-unit it consists of, a terminable TC-unit is categorized into a *simple terminable TC-unit* or a *complex terminable TC-unit*. A simple terminable TC-unit consists of one *independent single TC-unit*. The topic of a simple terminable TC-unit is not repeated in the form of *coreferential zeros* in the preceding or subsequent topic-comment structures. A complex terminable TC-unit consists of two or more successive *dependent single TC-units*. The topic of these dependent single TC-units only shows once in its full form and is repeated in the form of coreferential zero in the rest of dependent single TC-units. Coreferential zero is proposed in this dissertation referring to an element that does not have any phonological form and is unpronounced but corefers to the full-form topic mentioned in previous or subsequent single TC-unit within one terminable TC-unit. Coreferential zero is the requisite part integrating single TC-units into a topic chain. Upon the introduction of a new topic, or repetition of a topic in its full form, a pronoun, or demonstrative, instead of coreferential zero, a new TC-unit is then activated. The beginning and end points of a terminable TC-unit does not necessarily take the form of a conventional sentence with punctuation marks. Provided such taxonomy of TC-units, clause combining refers to the composition of single TC-units within each terminable TC-unit. Clause embedding refers to the elaboration within

each single TC-unit via adding modifiers preceding the head, complex post-verbal arguments, and complex serial verb constructions, etc.

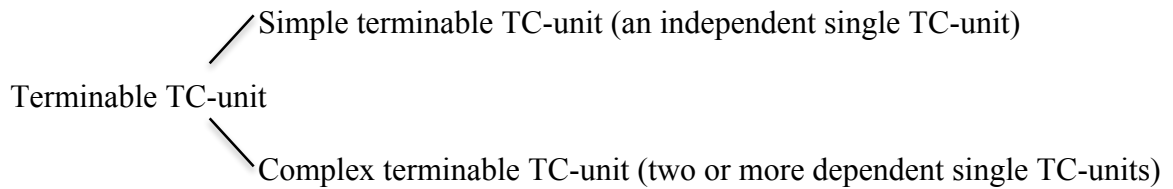


Figure 6. Taxonomy of TC-units.

Taken the aforementioned (55) as an example, Jiang (2013) analyzed it as a three T-units sentence. Applying TC-unit, (55) is analyzed as 1 terminable TC-unit consisting of 3 dependent single TC-units. “*Wǒ jiějie*” is the topic and repeated in the form of coreferential zero twice. (55a) below added another sentence to follow the original sentence of (55). When “*wǒ*” is introduced as different topic, it starts a second terminable TC-unit which consists of one independent single TC-unit.

(55a) *Wǒ jiějie jiào mǎlì, jīnnián èrshí suì, zài běijīng shàng dàxué. Wǒ hěn xǐhuan tā.*

我姐姐_i叫玛丽， \emptyset_i 今年二十岁， \emptyset_i 在北京上大学。我_j很喜欢她。

[I elder sister_i call Mary, \emptyset_i this year twenty year, \emptyset_i at Beijing go to college. I_j very like her.]

My elder sister is called Mary. She is twenty this year. She goes to college in Beijing. I like her very much.

Applying terminable TC-unit as the unit of analysis for Chinese syntactic complexity, the question of relationship between the main clause and subordinated clause

in complexity analysis shifts to the investigation on how the dependent single TC-unit is composed in a terminable TC-unit replaced. Adopting the framework of complexity investigation (Norris & Ortega, 2009), this dissertation suggests tackle Chinese syntactic complexity at at least three levels: overall complexity, clausal complexity, and phrasal complexity, with the potential addition of specific form complexity, depending on the research question. Reviewing the conceptualization of complexity discussed in Section 2.1 as well as the Chinese typological linguistic features in Section 2.2, this dissertation suggests at least 7 indices as listed in Table 4 for validation in empirical studies via their correlation to L2 Chinese language proficiency. These measures here proposed can be categorized into three types: length, frequency and ratio. Measure mean length of terminable TC-unit (MLTTCU) and mean length of single TC-unit (both independent and dependent) (MLSTCU) are length measures. Complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU) and ratio of different types of terminable TC-unit, single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU), and dependents per head are ratio measures. Frequency of a specific form, as befits the name, is a frequency measure.

For overall Chinese syntactic complexity, or global complexity, three indices are proposed. First, the mean length of a terminable TC-unit (MLTTCU) is measured in terms of the average number of characters of the terminable TC-units. All the simple terminable TC-units and complex terminable TC-units are counted. Second, the ratio of the total amount of complex terminable TC-unit divided by the total amount of all the terminable TC-units (both simple and complex) (CTTCU/ATTCU) is also considered as a global complexity indicator. A complex terminable TC-unit consists of two or more

Table 4

Inventory of Chinese syntactic complexity measures

| Complexity | Measurements |
|--------------------|--|
| Global | <ul style="list-style-type: none"> • Mean length of terminable TC-unit (MLTTCU) • Complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU) • Ratio of different types of terminable TC-unit |
| Clausal | <ul style="list-style-type: none"> • Mean length of single TC-unit (both independent and dependent) (MLSTCU) • Single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU) |
| Subclausal/phrasal | <ul style="list-style-type: none"> • Dependents per head |
| Specific form | <ul style="list-style-type: none"> • Frequency of a specific form |

depended single TC-units connected via correferential zero(s), in which way it manages to choreograph more information in a more intricate and coherent way. A simple terminable TC-unit consists of one and only one independent single TC-unit. The higher ratio of complex terminable TC-units to all terminable TC-unit is, the higher overall complexity is estimated. Again, the number of characters is used to measure the length of the unit for analysis. Third, the ratio of different types of terminable TC-unit is proposed to be one of the overall Chinese syntactic complexity measures in this dissertation. Since both the number and the nature of the relationships between constituent components is a dimension of syntactic complexity, the different types of terminable TC-unit shall also be

checked. There are scant studies that have looked at the difficulty or acquisition sequence of these patterns for L2 Chinese speakers. However, when checking the types of terminable TC-unit, factors like genre of the discourse and speakers' individual differences should also be taken into account.

At the clausal level, two measures are proposed: Mean length of single TC-unit (both independent and dependent) (MLSTCU) and Single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU). Single TC-units (both independent and dependent) are the subordinate units that compose minimally terminable TC-units. Mean length of single TC-units (both independent and dependent) in terms of the number of characters shows clausal level complexity. Therefore the length measure MLSTCU is used to measure the clausal complexity. Applying the ratio measure STCU/TTCU, the total amount of single TC-units (independent or dependent) is divided by the total amount of terminable TC-units (both simple and complex). For each complex terminable TC-unit in which dependent single TC-units are combined via coreferential zero, the more dependent single TC-units there are, the higher the clausal complexity is. A simple terminable TC-unit, which consists of only one independent single TC-unit, contributes less clausal complexity than a complex TC-unit does. In this sense, the number of coreferential zero per terminable TC-unit can also be used to indicate clausal complexity. The number of coreferential zero in each terminable TC-unit equals to the number of dependent single TC-units minus one.

As discussed in 2.2.2.2, a relative clause and a complement clause in English are classified as clause embedding but not clause combining, thus they contribute to the subclausal or phrasal level complexity in English. The Chinese equivalents of English

relative clauses are placed preceding the head as modifiers, as exemplified in (29a) and (30a) in Section 2.2.2.2. Therefore, subclausal or phrasal complexity is measured by nominal or verbal modifiers per head. A phrase of which the head is modified with multiple dependents is considered as more complex than those with one single or no modifier.

As listed in Table 4, the additional level of specific form complexity is included in the proposed indices for specific research purpose. If a study focuses on the complexity caused by the use of some specific form, the indice of the frequency of a specific form can be employed.

For the length measurement MLTTCU and MLSTCU, this dissertation codes them in terms of the number of characters. According to morphological level typology categorization, Chinese belongs to the isolating type of which a word is composed with morpheme(s) with no affixes. A word can be a single syllable or multiple syllables. One syllable often equates to one morpheme, which mostly in written Chinese takes the form of one character. As discussed in 2.2.2, if a complexity measure is applied in terms of the number of words or morphemes, it simultaneously includes phrasal complexity. Therefore, the length measure MLTTCU proposed to detect the global complexity is a hybrid measure which simultaneously taps into clausal complexity. The other generic length measure MLSTCU also taps into clausal complexity.

There are several reasons for choosing the amount of characters over words. First, there is not yet an unanimously approved definition of Chinese word demarcating. A syntactic word is usually a bigram but it is also common to have unigram, trigram, and quadgram. Word boundary and word-class related studies are debated and always in flux.

In written Chinese, text is presented in equally spaced characters with no additional spatial demarcation between words. The boundary of a Chinese syntactic word can be hard to identify. It can sometimes even be challenging to differentiate words from phrases in Chinese. The majority of Chinese syntactic words are bigram, but a Chinese phrase can also be bigram. For example, a phrase báimǎ (白马, white horse) is a bigram which can be easily considered as a word, but according to the syntactic analysis consider it as a phrase since it consists of two free morphemes. Such lack of clarity on syntactic word boundary adds to the difficulty to operationalizing the mean length of a Chinese unit of analysis in terms of the word. Second, there still goes on a fundamental debate if word should be considered as the essential unit of analysis in Chinese syntax. A unit “*Zi*”, which is mostly overlapped with the character, is argued to be the unit of analysis for Chinese language (Chao, 1968; Xu, 1991, 1997, 2005; Wang, 1994a, 1994b; 2000; Pan, 2002). Last, Jiang (2013, p. 17) measured Chinese T-unit length both in terms of words/T-unit and characters/T-unit, across L2 proficiencies, and recommended characters/T-unit for future research on L2 Chinese when results do not have to be compared with those of T-unit analyses in other languages because characters/T-unit was found more reliable for coding.

To empirically validate these measures, first need to quantitatively correlate the results of applying these indices measuring L2 Chinese written and spoken output with a full range of L2 proficiency levels from elementary to superior, or plus Chinese output of L1 Chinese speakers. In addition, qualitative analysis of L2 Chinese learners’ longitudinal development in terms of syntactic complexity and maybe CAF can be

employed to provide a thorough and comprehensive picture of L2 Chinese syntactic complexity development.

3.4 Elicited Imitation (EI) test

The EI test has been used in second language proficiency assessment and has proved to be an effective language test with high correlations with the *Oral Proficiency Interview* (Erlam, 2006; Ortega et al., 2002). A Mandarin EI test was developed by Zhou and Wu (2009)ⁱ in which the participants listen to 30 Chinese sentences of varied length, vocabulary, and grammar structures in sequence and are asked to repeat each sentence as much as possible after a short pause. All of the 30 sentences were listed in the *Appendix C Mandarin Repetition Task* in Zhou (2012) and were cited in Appendix D of this dissertation.

To reach an optimal design of EI test, three major concerns are to be addressed: 1) serial position effect, 2) the memory effect, and 3) the ceiling and floor effect. Zhou (2012) minimized each of these three effects in the Mandarin EI test.

Serial position effect refers to the tendency that the first and last items in a series best will be recalled the most comparing to the middle items. Zhou (2012) summarized that EI test had been used to test as a measure tool by comparison to other tools, to test the theory of Universal Grammar (UG), the effect of pedagogical interventions, and the implicit knowledge of a language learner. While serial position effect should be taken into consideration in test a specific structure, it does not affect EI test to be a reasonable measure of global competence of language learning (Bley-Vroman & Chaudron, 1994).

This dissertation employs EI test to assess the global Chinese proficiency, therefore does not have to consider the serial position effect.

Memory effect takes place when a sentence is to be repeated. The EI test taker can have a great memory and repeat the prompt sentence given without even being able to comprehend it, or a native speaker can fail repeating the prompt sentence due to limited memory capacity. To minimize the memory effect, seven was found to be the magic number. It was found that a person's immediate memory span is "usually somewhat in the neighborhood of seven" units of information (Miller, 1956, p.90). Bley-Vroman and Chaudron (1994) clarified that this unit of information can contain more than seven syllables but chunks, yet the length of the chunks may vary. To minimize the rote memory effect, the length of the stimuli sentence in the Mandarin EI test (Zhou & Wu, 2009) was designed between 7~19 syllables (See the complete Mandarin EI task, including instructions and stimuli, in Appendix D). In addition, a short pause was inserted between the end of the stimuli sentence and participants' repetition. This short pause lasted for 2.5 second, which consisted of two parts: 1) 2 second pause, and 2) a 0.5 second ring tone immediately followed the 2 second pause. In addition to delay the participants' repetition, the 0.5 second ring tone also functioned to signal the participants to start the repetition. This Mandarin EI test was presented through taped recording. All the instructions and stimuli sentences, in English or Mandarin, were pre-recorded with a female's voice. The response time allotted of each sentence for participants were as well pre-saved, which consisted of two parts: (a) the time that a seven-syllable stimuli sentence takes native speakers to repeat, and (b) additional 2 seconds. In addition, for sentences that include more than seven syllables, 0.5 second was added for each

additional syllable. For example, the response time saved for a 19 syllable stimuli sentence was added up by three parts: (a) the time that a seven syllable stimuli sentence takes native speakers to repeat; (b) additional 2 seconds; and (c) $0.5 \times 12 (= 19 - 7)$ additional syllable = 6 seconds.

At the beginning of this Mandarin EI test, a trial English session was given prior to the Chinese part to familiarize the English speaking participants with the EI test procedure. Instructions were given that the participants were going to hear several English sentences, after each sentence there would be a short pause, followed by a tone sound. The participants were instructed “don’t start repeating the sentence until you hear the tone sound”. They were also told to “try to repeat exactly what you hear” and “repeat as much as you can”. Six English stimuli sentences of varied length were given to repeat as a trial session. After participants finished repeating these six English stimuli sentences, the participants were told, “that was the last English sentence” and “now you are going to hear a number of sentences in Mandarin”. They were reminded again to not start repeating until they hear the tone sound. The Mandarin part includes 30 sentences from the minimum of 7 syllables to the maximum of 19 syllables. After finished repeating all the 30 sentences, the participants were then told, “This is the end of the repetition task. Thank you”.

Ortega, Iwashita, Rabie, and Norris (in preparation) developed a scoring rubric for EI test to assign points of 0 to 4 to each sentence repetition based on how well both meaning and form were preserved in the repetition. The complete form is cited in Appendix E. A summarized version was described in Zhou (2012) as below:

(a) when there is a silence, unintelligible garbles, or only one word is repeated, zero points are awarded; (b) when half or barely half of lexical words and meaning are retained in the repetition, one point is assigned; (c) two points are assigned when at least more than half of the idea units are maintained; (d) when original and complete meaning is preserved, but some synonymous substitutions occurs without changing the meaning, three points are assigned; (e) four points are only given to perfect repetition, where both form and meaning are faithfully preserved and repeated. (p. 110)

A satisfactory internal reliability for this Mandarin EI test was reported with Cronbach's alpha $\alpha = .968$ (Zhou & Wu, 2009). The inter-rater reliability of the final coding was high ($r = .985, p < .01$). Zhou (2012) administered this Mandarin EI test in assessing the global competence of language learners of Mandarin Chinese. The EI performance and language contact of heritage/foreign learners of Chinese also showed a moderate correlation at $r = .566 (p < .05)$. To efficiently assess the global Chinese proficiency of participants, this dissertation included this Mandarin EI test as adopted from Zhou and Wu (2009).

Besides the serial position effects and the memory effect discussed above, another concern of EI test design is the ceiling and floor effect, that is to say, if the EI test can well differentiate speakers of different global proficiency. A good spread of the item difficulty is required in order to test a wide range of proficiency. Zhou (2012) conducted item discrimination (ID) analysis of all the 30 stimuli sentences to examine the validity of this Mandarin EI test. The ID analysis of this Mandarin EI test showed that seventeen

items showed ID indexes between .32 - .58 ($M = .39$), seven items showed ID indexes between .20 - .28 ($M = .24$), and six items (Item 1, 10, 18, 20, 23, and 26) showed ID indexes between .11 - .18, ($M = .15$). Brown (2005, p. 75) concluded that items with ID indexes of .30 and above are usually considered acceptable items and those at .09 or below are poor items and need to be improved by revision. According to this classification, in this Mandarin EI test twenty-four items were acceptable or better while six items needed to improve. Zhou (2012) therefore observed that “the items in this Mandarin EI task are reasonably able to distinguish high and low level of Mandarin learners, and there are no items which are too difficult or too easy to generate the floor and ceiling effects (p. 114).”

As aforementioned in Section 2.1.3, taking students’ class standing and institutional status as a grouping variable is not always reliable due to intervening factors such as heritage speaking background of the target language, the gap between the true levels of the same titled courses in different institutions, or even the variation of students’ proficiency distribution in the same class. For example, both are the second year level Chinese courses, course in Institution A might be designed much advanced than the second year level Chinese course in Institution B. Even within the same class, a student with a grade of A+ could far outperform a student with a grade of F. A student with a grade of F in the second year level course might even be less proficient than a student with a grade of A in the first year level course. With the language exposure at home since early childhood, heritage learners might perform differently from other non-heritage students. Given such possible variability among participants, this dissertation included the Mandarin EI test to access the L2 Chinese speakers’ global proficiency levels. This

dissertation therefore employed this Mandarin EI test to assess participants' global Chinese proficiency in order to classify proficiency groups.

CHAPTER 4

RESEARCH DESIGN AND METHODS

4.1 Research questions

Four main research questions are investigated in this dissertation:

RQ 1: To what extent do the measures of Chinese spoken and written syntactic complexity of L1 and L2 Chinese speakers correlate with their global proficiency levels?

RQ 2: To what extent do combinations of the four measures of Chinese syntactic complexity distinguish between low, high, and native proficiency speakers of Chinese?

RQ 3: How does cognitive task complexity affect the Chinese syntactic complexity of Chinese speakers across low, high, and native proficiency levels?

RQ 4: How does immediate task repetition affect task performance and self-perception in terms of Chinese syntactic complexity for Chinese speakers across low, high, and native proficiency levels?

4.2 Data collection

4.2.1 Participants

A total of 109 English speaking Chinese L2 speakers and 32 Chinese L1 speakers completed a *Chinese Timed Writing and Speaking Test* (TW&ST) which was designed in this dissertation to collect both spoken and written Chinese output. The participants were recruited in both the U.S. and China. The English speaking Chinese L2 speakers consisted of three groups:

(a) English speaking undergraduate and graduate students who were taking various levels of Chinese language courses or Chinese major courses at universities in the U.S. They had all taken a minimum of one school year of Chinese language courses at college. Of these English speaking undergraduate and graduate students, the advanced Chinese learners had taken more than four years of Chinese language courses as well as Chinese major courses in such areas as Chinese linguistics, literature, culture, and Asian studies.

(b) English speaking high school or university faculty and staff in the U.S. whose work or research was related to Chinese. Chinese was a working language for these participants. They had been teaching Chinese language or Chinese-related subjects such as Chinese literature and social studies, or had been working in a Chinese program in the U.S. using Chinese as a working language.

(c) English speaking L2 Chinese speakers, who were taking advanced Chinese courses or working at universities in Beijing, China. These participants had been studying abroad in Beijing taking advanced level Mandarin courses, or working in Chinese programs in China using Chinese as their working language.

The Chinese L1 speaker participants consisted of two groups: (d) Chinese undergraduate or graduate students who majored in different fields of study at top-ranked national universities in Beijing; and (e) L1 Chinese speakers who already held a Bachelor's or Master's degree from top-ranked national universities in Beijing and had been working at different professions in Beijing. All of the L1 Chinese speaker participants were studying or working in Beijing at the time of the study. It is important to note that Mandarin is the national lingua franca of Chinese people of different dialects

and different language minority groups. It is likely that a northern-dialect speaker and a southern-dialect speaker can hardly communicate except through the national standard dialect of Mandarin. Beijing is the capital of China and the phonology of standard Mandarin is based on the Beijing dialect. Further, the Beijing dialect belongs to the northern dialect of China, on which the vocabulary and syntax of Mandarin is based. Therefore, this dissertation recruited people who were studying and working in Beijing to be the L1 Chinese participants for data collection. Furthermore, all of the L1 Chinese speaker participants held *College English Test Band 6* (CET-6) Certificates, *Test for English Major Grade 4* (TEM-4) Certificates, a *Test of English as a Foreign Language* (TOEFL) score of 80 and higher, or an *International English Language Testing System* (IELTS) score of 5.5 and higher. This advanced English proficiency level was an additional requirement in order to make sure that the participants are able to comprehend the 3 minute 40 second video clip in English when completing the story retelling task in the TW&ST.

Out of the data collected from the 141 participants completing the TW&ST online, 13 of them claimed their dominant language not to be English, or not only English but also Cantonese or Mandarin. 2 participants' background information survey was not successfully sent to the server due to Internet problems. In addition, the Elicited Imitation (EI) score for 6 participants, the spoken data for 5 participants, and the written data for another 4 participants were incomplete due to the technical failure at audio recording or mp3 file uploading online. One participant misunderstood the instructions of the EI test resulting in invalid data. One participant gave up at the half way point due to lack of confidence in achieving a good score. And one participant misused the microphone of the

headset, which resulted in null recordings. Therefore, a total of 5 speaking data sets and 4 written data sets were excluded, as listed in Table 5 below. Excluding the invalid or incomplete data sets, a total of 115 sets of complete and valid spoken data and 116 sets of complete and valid written data were included for analyses in this dissertation.

Table 5

Calculation of complete and valid spoken and written data sets

| | Spoken output | Written output |
|---|----------------------|-----------------------|
| Total participants | | 141 |
| Dominant language mismatch | | 13 |
| EI test recording missing | | 6 |
| Survey missing | | 2 |
| Task output missing | 5 | 4 |
| Total complete and valid data sets | 115 | 116 |

4.2.2 Instruments

Through collaboration between the researcher and a technology specialist, an online Chinese Timed Writing and Speaking Test (TW&ST)ⁱⁱ (See Appendix B TW&ST (English instruction) screenshots and Appendix C TW&ST (Chinese instruction) screenshots) was designed for this dissertation to collect both spoken and written Chinese output. There were two versions of the TW&ST: TW&ST (English instructions) for L2 Chinese learner participants and TW&ST (Chinese instructions) for L1 Chinese speaker participants.

The TW&ST (English instructions) for L2 Chinese speaker participants consisted of nine parts, in the following sequence: (a) a background information survey; (b) a preparation session; (c) a comic strip description task (CS); (d) a video story retelling task (V1); (e) an immediately repeated video story retelling task (V2); (f) a retrospective survey; (g) a free writing task (FW); (h) a guided re-writing task (GR); and (i) a Mandarin elicited imitation (EI) test (Zhou & Wu, 2009).

The background survey in TW&ST (English instructions) was included in order to filter out L2 Chinese speaker participants whose dominant language were not Chinese or not Chinese only. The survey asked participants' age, gender, dominant language, other language learning background, educational background, Chinese learning history, Chinese exposure at home, and time spent in Chinese speaking areas or countries.

A short preparation session was provided to familiarize the participants with the online interface and the procedure of TW&ST. It was a compressed version of the CS description task, with comparatively very short preparation time and recording time allotted. The comic strip consisted of six sequential pictures with a daily life topic. Before proceeding to the screen of the cartoon strip, an instruction screen informed the participants that they were going to have 10 seconds preparation time before the recording automatically began, and their recording should be up to 10 seconds while looking at the pictures.

There were three speaking tasks in TW&ST (English instructions): one was a CS task, and the other two were video story retelling tasks: V1 and V2. The comic strip in the CS consisted of six sequential pictures illustrating a family dinnertime story (Refer to Appendix B for the comic strip). The participants were instructed to tell a story based on

the comic strip after a short preparation time. The comic strip was displayed to the participants throughout the whole time of their preparation and storytelling. For the video story retelling tasks, participants were instructed to retell a story in Chinese after watching a short video clip narrated in English. The video clip was about folklore of the Chinese New Year. Participants were also instructed not to literally translate the exact words or patterns they heard in the video clip but instead to tell a complete story in Chinese based on the story in the video clip. V1 was followed by an immediate repeated task, V2, in which the participants repeated the V1 task by retelling the same video story for a second time without watching the video clip again.

These three speaking tasks, CS, V1, and V2, were all structured in order to control the same topic, context, and time on the task across individuals. The CS task required participants to narrate a story based on six cartoon frames arranged in sequence. By contrast, the V1 and V2 tasks were relatively less restricted for participants in that they were allowed more flexibility and creativity was allowed in organizing the storyline as well as choosing what details to include in their retelling of the story. All the speaking task outputs were in monologic not dialogic condition in order to avoid reduction of forms, which often appears in highly interactional language. As the aforementioned example (3) in Section 2.2.1.2 showed, the reduction in Chinese conversation is not limited to nouns, as other classes of word and sentence components can also be reduced and become zero anaphors within the context provided. Though the intended reading for (3a) is what (3b) means, both the verb and part of its complement were reduced and became zero anaphors since (3a) appeared with context in a conversation. In dialogic Chinese, the reduction of subjects, predicates, and even other parts can be very common

syntactic operations (Shi, 1993). Such reduction can to a great extent affect the length of the unit of analysis for syntactic complexity. Therefore, interactional language was excluded in favor of monologic narrative language in the task design in this dissertation.

However, these three speaking tasks differed from each other in important ways, with variables of both resource-directing and resource-dispersing dimensions of the Multiple Resources Attentional Model (Robinson, 2001, 2003, 2005) applied in the task design. Along the line of resource-directing dimensions, the video story retelling tasks were considered of much higher task complexity than the comic strip description task. The complexity gap between the video story retelling tasks and the comic strip description task was embodied in three aspects: (a) the video story retelling tasks included many more elements in the story, while the comic strip description task included fewer elements in the story; (b) the video story retelling tasks required twice as long as the comic strip description task to complete; and (c) the video story retelling tasks required arguably more cognitive reasoning effort when the participants retold the story without seeing the video anymore, while the comic strip description task required less cognitive reasoning effort since the participants could describe the comic strip while looking at the pictures (a version of the Here-and-Now, There-and-Then task, see Robinson, 2001). For the video retelling tasks, V1 was immediately repeated for a second time, and this repeated task V2 was considered less cognitively complex than V1. As discussed in Section 2.3.3, the assumption is that when repeating the same task for the second time, V1 was considered a form of task planning for the V2 performance. This design, including speaking tasks of varied cognitive complexity, also aimed to minimize possible ceiling and floor effects by allowing performance of L2 Chinese learners across

the full range of proficiency levels, including elementary learners. The less complex CS task provided an entry level for participants of lower proficiency, and the more complex V1 and V2 tasks were presumed to allow enough opportunity for participants to produce a lot of language and to demonstrate their fuller range of syntactic complexity.

The writing part of the TW&ST consisted of two tasks: a free writing (FW) task and a guided rewriting (GR) task. The FW task required participants to write a well-organized paragraph in seven minutes on the topic of *my father /mother /brother /sister /friend (choose any one of them) and I*. This topic of daily life was chosen in order to have this writing task stay in the likely ability range of all the L2 speakers of various levels from low-intermediate across advanced. The GR task provided seven semantically coherent but formally incohesive sentences for the participants to connect into one coherent paragraph. Both simplified and traditional scripts were provided for participants (who may have been familiar with different scripts). These seven sentences were controlled lexically at an elementary level in order to warrant their comprehensibility to all the L2 Chinese learner participants. The participants could manipulate the sentences, change the order of words, and omit words, but were required not to leave out any of the given information. Though named free writing task and guided rewriting, these two writing tasks both belonged to the category of structured writing tasks with topic and time controlled. In addition to controlling topic and time, the GR task had the sentence skeleton provided and asked the participants to turn the seven sentences into one coherent paragraph. With the sentence skeleton provided, that allowed possibilities of different clause combinations and reformatting, the guided rewriting task was expected to allow the advanced Chinese learners and L1 Chinese speakers to produce relatively highly

complex data in a limited time. In this way the task could draw out more variability of syntactic complexity from Chinese speakers of various proficiency levels.

Since TW&ST was web-based, participants could type Pinyin on the computer to select the characters provided by the Pinyin input method, which was different from writing Chinese characters in a paper and pencil setting. Typing on the computer is taken as easier than handwriting, since typing actually enables more character recognition rather than character writing. The participants were also instructed that they could type in Pinyin if they did not know the character. In this way, this dissertation sought to exclude a character scripting variable from Chinese writing proficiency and tried to minimize the influence of character scripting in the writing tasks.

To assess global Chinese proficiency of the L2 Chinese speaker participants, a Mandarin Elicited Imitation (EI) test was included in TW&ST. As introduced in Section 3.3, in this Mandarin EI test, participants were asked to repeat 30 sentences of various length and difficulty in sequence as accurately as possible. The whole session of the EI test was pre-recorded, including instructions, stimuli sentences, and response time saved for participants.

TW&ST (Chinese instructions) was basically the same as TW&ST (English instructions), though differed moderately in terms of the following three aspects. First, as befits the name, in TW&ST (English instructions) for L2 Chinese learner participants whose dominant language was English, all the guidelines, instructions, and questions were given in English so that reading instructions would not cause any difficulty. All the guidelines, instructions, and questions in TW&ST (Chinese instructions), correspondingly, were translated into Chinese, so the L1 Chinese speaker participants

would not have any problem comprehending the instructions. Second, the EI test for establishing global Chinese proficiency was not included in the TW&ST (Chinese instructions). All the L1 Chinese speaker participants were categorized into Group Native in the study. Lastly, as for the background information survey, different questions were asked in TW&ST (Chinese instructions) from those in TW&ST (English instructions) serving the different purposes. The background information survey in TW&ST (English instructions) was designed mainly to filter out L2 Chinese speaker participants whose dominant language were not Chinese or not Chinese only. By contrast, the background information survey in TW&ST (Chinese instructions) was mainly to filter out L1 Chinese speaker participants whose English proficiency was not sufficient to comprehend the English narration in the video retelling tasks. Therefore, the background information survey in TW&ST (Chinese instructions) asked participants' age, gender, educational background, Chinese dialect background, English proficiency, foreign language learning experience, and experience of staying in non-Chinese speaking areas or countries.

4.2.3 Procedures

In order to diminish the intervening variables at the step of behavior elicitation during the measurement process (as displayed in Figure 2 in Section 2.1.3), TW&ST was designed aiming to realize task-as-process as approximately as task-as-workplan. First, the researcher designed the blueprint of TW&ST in the forms of individual screen designs with specific requirements annotated. Such screen designs included information such as the overall arrangement of the screen, the function of each screen (like graphic displaying, audio/video playing, audio recording, timing indication, text presenting, and

control panel setting, etc.), and the display time for each screen. Next, the researcher and the technology specialist conferenced multiple times to communicate and adjust researcher's expected design given the limit of existing computer technology. Then, following the screen designs, the technology specialist designed TW&ST using the programming language of HTML and the Web: JavaScript. The two surveys in TW&ST were written by the researcher via an online platform: www.jotform.com. This platform was selected because it is user-friendly and also allows the survey to be embedded into other website. Google Chrome was designed to be the default browser for TW&ST for best display. After the technology specialist connected TW&ST on to a local server and launched it online, pilot studies were then conducted by the researcher on L2 Chinese learners of varied proficiencie levels and L1 Chinese speakers in a laboratory setting. The researcher then took notes on the troubles occured and participants' feedback from the pilot studies. Based on the detailded revision requests from the researcher on each screen, the technology specialist revised accordingly the TW&ST design and data saving. The revised TW&ST was then put through the next round of pilot studies. This step of pilot-study-and-then-revision was taken multiple rounds until no more necessary revision was seen from the researcher's end. Participants were enabled to complete TW&ST without any assistance from the TW&ST administrator throughout the whole process.

The TW&ST (English Instructions) took the participants around 45 minutes to complete. The TW&ST (Chinese Instructions), with no EI test part, took about 35 minutes to complete. All parts of the TW&ST were completed online. Throughout the whole test, no note-taking was allowed. Adult L2 Chinese learners and L1 Chinese speakers completed the TW&ST in a computer laboratory setting. Except for the

background information survey and instruction screens, the display time of each screen of the TW&ST was pre-set. Therefore, time for participants to prepare for each task, as well as the time for participants to complete each task was strictly regulated. In addition, as shown in an example screenshot in Figure 7 below, for every timed screen, the remaining seconds for showing this screen were displayed at the top right of the screen. The test was designed such that pause, forward, and backward functions were all disabled in order to keep the timing of the test standardized across individuals. Upon the time limit of each screen, the computer automatically proceeded to the next slide. Upon the time limit of each speaking session of the TW&ST, the recording of the participants' speaking was automatically converted into mp3 files and immediately uploaded to the database online. The tasks in the TW&ST were all piloted with L2 Chinese learners of varying proficiency and adult L1 Chinese speakers in order to determine sufficient time for participants to complete each task. Finally, the two surveys were designed on an online platform, jotform, and embedded into the TW&ST.

When beginning the TW&ST, participants were shown the guidelines first. The guidelines introduced participants to the main sections of the TW&ST and a time estimate was provided for them to complete the test. They were reminded about the time limit for each task preparation and completion. In addition, they were also requested to remain patient while the computer took some time to convert and upload their recordings after they finished each speaking task.

The first part of TW&ST was the background information survey. The survey took 2-3 minutes to complete but it was not timed. Participants could use their time to complete the survey and click to go to the next page when they finished it. Upon

participants' submitting the background information survey, the short preparation session, which was a compressed CS task, was then given to familiarize the participants with the test procedure. For this compressed CS task, the preparation time given was 10 seconds. When it reached 10 seconds, the computer automatically started recording for 10 seconds. A volume bar was shown on the screen when participants spoke. Therefore, the participants were aware of how to do any online adjusting during their speaking. When the time for recording reached the limit, the computer then automatically stopped recording, promptly saved the recording, converted it into an mp3 file, uploaded the mp3 file to the database, and proceeded to the next slide.


After the preparation session, the three structured speaking tasks were given to elicit spoken data from the participants: first the CS task and then the V1 and V2. The preparation time set for the CS task was 30 seconds. Throughout these 30 seconds, the participants were presented the comic strip and they prepared for telling a story based on the comic strip. When 30 seconds were up, the computer automatically proceeded to the next slide and started recording participants. The comic strip was displayed throughout the participants' speaking. The speaking time given was upto 90 seconds. When 90 seconds were up, the computer then proceeded to the next speaking task, which was, the V1 task. In the V1 task, before playing the video, participants were instructed that they were going to watch a 3-minute-and-40-second video about Chinese New Year, which was narrated in English, and after they finished watching the video, they would be given 30 seconds to prepare before they retold the story in Chinese for up to 3 minutes. After reading the instruction slide, participants clicked to start playing the video. During the playing of the video clip, participants could not forward, pause, or rewind the video clip

Screen 14

TWIST (14/22)
(help with this question)

help

173
Seconds Remaining



Video Story Retelling
You are recording now...

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

Notice that it takes quite some time to upload your recording after you finish.

home

Figure 7. Screenshot of one slide in the TW&ST.

as preset by the TW&ST. Being aware of the expected task outcome, participants could have actually started planning while watching the video clip. After the video clip reached the end, 30 seconds were given for the participants to prepare before recording started. When it reached 30 seconds, the computer then automatically proceeded to the next slide and started recording. The time given for participants to retell the story in Chinese was up to 3 minutes. When it reached 3 minutes, the computer automatically stopped recording and uploaded the audio file.

An immediate task repetition, V2, followed after the first retelling of V1. Upon finishing uploading of their V1 recording, the computer proceeded to the next slide on which there were instructions informing the participants that they had just gotten a second chance. They were told that they would be given another 30-second preparation time, and then they could repeat their retelling to improve on their performance. After another 30 seconds, the computer automatically started recording. Speaking time given for the V2 task was also up to 3 minutes.

A retrospective survey was given in the TW&ST as soon as participants completed both enactments of video story retelling, V1 and V2. In the survey, questions were asked on participants' self-perceived performance, strategy use, and affective state during the story retelling tasks. The retrospective survey was not timed.

Upon participants' submitting the retrospective survey, the two writing tasks followed. The FW task allowed up to 7 minutes for participants. When it reached 7 minutes, the computer automatically saved all that the participants had typed and then proceeded to the GR task. Time allocated for the GR task was up to 5 minutes.

When it reached the 5-minute limit for the GR task, the computer automatically proceeded to the EI test that was the last part of the TW&ST. The participants were instructed that they would hear 30 Chinese sentences, one by one, of various length and complexity. Their task was to repeat each sentence as exactly as possible in the time provided after hearing each sentence. Prior to the Chinese part, the English speaking L2 Chinese speaker participants first completed a practice round repeating English sentences. To complete the EI test required 10 minutes and 40 seconds.

The design of TW&ST in this dissertation mainly aimed for the following goals. First, TW&ST was designed to put together all the test components such as test instructions, individual tasks, surveys, and the EI test in a way that would enable the group administration of such test. The test administrator does not have to conduct individual tasks and surveys at a one-on-one research setting with individual participants in order to control everyone's task preparation and performance time. More efficiently, TW&ST enabled many participants to simultaneously take a series of sequenced tasks and surveys in a laboratory setting with only one administrator presented. In addition, with its completion time designed around 45 minutes, TW&ST enabled participants to complete it within the class time since one language class session is usually 45-50 minutes.

Second, compared with conducting individual tasks and surveys at a one-on-one setting, TW&ST aimed to diminish the differences among the repeated enactments of individual tasks and surveys for different participants. The administration setting of TW&ST was conformed in a laboratory setting. Since each screen of TW&ST was timed except for the instruction screens and survey screens, all the pre-task planning time, task

performing time, and task order were automatically conformed online. In addition, the computer-based design of TW&ST minimalized the interaction between the TW&ST administrator and participants which might have brought in other intervening variables such as the solidarity between test-administrator and test-taker. TW&ST was designed in a way that participants were able to complete it following the instructions online without any interaction with the TW&ST administrator/researcher. In this dissertation, the TW&ST administrator/researcher briefly introduced the purpose of the study and asked the participants to sign the consent form prior to their participation. Throughout the TW&ST administration, the administrator/researcher was present the whole time, but had to interact with the participants only a few times when technology trouble occurred.

Third, TW&ST was designed such that pause, forward, and backward functions between screens were all disabled in order to keep the timing of the test standardized across individuals. Except for the screens of instructions and survey, each screen was individually timed and the remaining time was indicated on the screen for TW&ST takers. Upon the time limit of each screen, the computer automatically proceeded to the next slide. In addition, the control panels for the video embedded at the video retelling task and audio embedded at the EI test were also hidden so that TW&ST takers were not able to pause, forward, and backward the video for more viewing. This way the viewing times and duration were standardized across individuals.

4.3 Scoring and analysis

4.3.1 Transcribing

All the written output collected from the FW and GR task via TW&ST were saved without the need for any transcribing. For all the spoken output collected from the CS, V1, and V2 tasks, the researcher transcribed twice to maximize the accuracy of transcription. There was a one-month time gap between the two times of transcribing in order to reduce transcribing bias. This is to say, the second time transcription which functioned as a revision of the first time transcription was not started until one month after the completion of the first. Transcribing guidelines are listed in Table 6 below with examples. When transcribing participants' spoken data, if there was self-repair, the corrected language form was saved without the part before self-correction. False starts, fillers, and back channel cues and fillers were not included in the transcription,

Table 6

Spoken data transcribing guidelines

| | Spoken form | Transcribed form |
|--|--|-------------------------|
| 1 False starts | <i>Bàba kàn...kàn dào shū</i> | 爸爸看到书 |
| 2 Self-corrections | <i>(Bùnéng zài qù, o,) bùnéng zàilái</i> | 不能再来 |
| 3 Back channel cues and fillers | <i>Nàgè; zhège; a</i> | Not coded |
| 4 Numbers coded in syllables | <i>365 tiān</i> | 三百六十五天 |
| 5 English substitution in the middle of a sentence | <i>Māma shuō háizi qù go get bàba</i> | Sentence excluded |
| 6 Incomplete final sentence of audio recording or written data | <i>Due to time limit</i> | Sentence not coded |

given the focus of this dissertation on syntactic phenomena. Sentences in English, or sentences with word(s) in English substitution were not counted as valid data, hence, they were not included in the transcription. Since there was a time limit for each task, the last sentence of the audio recording or typed writing might have been cut off and incomplete. These incomplete sentences were also excluded from the data transcription. However, if incomplete sentence existed in the middle of the recording, where the incompleteness was due to participants' language proficiency but not the time limit, such incomplete sentences were still counted as valid data and included in the transcription. The numbers in spoken and written data were coded in characters. For the transcribed spoken data, the number of characters is mostly equal to the number of syllables in speaking.

4.3.2 Measures

Four of the Chinese syntactic complexity measures proposed in Section 3.3 were tested in this dissertation. They were investigated primarily by correlating the syntactic complexity measures of spoken and written output of L1 and L2 Chinese speakers with their various proficiency levels. These four measures are: ① Mean length of terminable TC-unit (MLTTCU), ② Complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU), ③ Mean length of single TC-unit (both independent and dependent) (MLSTCU), and ④ Single TC-units per terminable TC-unit (STCU/TTCU).

These four measures tap into different dimensions of syntactic complexity, MLTTCU and CTTCU/ATTCU addressing global complexity while MLSTCU and STCU/TTCU addressing the clausal level. As discussed in Section 3.3, the length

measures MLTTCU and MLSTCU are counted in the amount of characters, which is mostly equal to the number of syllables in Chinese.

Taking the previously listed (55a) (see Section 3.3) as an example, there are two terminable TC-units marked below as TTCU-1 and TTCU-2. TTCU-1, as a complex terminable TC-unit, consists of 3 single dependent TC-units as marked STCU-1, STCU-2, and STCU-3. The length of TTCU-1 is 17 characters. TTCU-2, as a simple terminable TC-unit, consists of only 1 independent TC-unit marked as STCU-1. The length of TTCU-2 is 5 characters. Applying the four measures, then, at global complexity level, (55a) is scored 11 characters at measure ① MLTTCU ($= (17+ 5)/ 2$) and 5 ($= 1/ 2$) at measure ② CTTCU/ATTCU; at clausal complexity level, (55a) is scored 5.5 characters ($= (6+ 5+ 6+ 5)/ 4$) at measure ③ MLSTCU and 2 at measure ④ STCU/TTCU ($= (3+ 1)/ 2$).

(55a) *Wǒ jiějie jiào mǎlì, jīnnián èrshí suì, zài běijīng shàng dàxué. Wǒ hěn xǐhuan tā.*

我姐姐_i叫玛丽， \emptyset_i 今年二十岁， \emptyset_i 在北京上大学。我_j很喜欢她。

I elder sister_i call Mary (STCU-1), \emptyset_i this year twenty year (STCU-1), \emptyset_i at Beijing go to college/(STCU-3, TTCU-1). I_j very like her (STCU-1, TTCU-2).

My elder sister is called Mary. She is twenty this year. She goes to college in Beijing. I like her very much.

4.3.3 Coding

All the spoken transcription and written data were coded by the researcher according to the four measures this dissertation proposed in Table 4: ① Mean length of terminable TC-unit (MLTTCU), ② Complex terminable TC-unit/all the TC-units

(CTTCU/ATTCU), ③ Mean length of single TC-unit (MLSTCU), and ④ Single TC-units per terminable TC-unit (STCU/TTCU).

All the transcribed spoken output as well as the written output collected in this dissertation were listed by task in five separate Microsoft Excel files. The coding conducted by the researcher followed three sequential steps: 1) identifying the phenomena, 2) counting the phenomena, and 3) calculating scores on each of the four measures. In each Excel file, the output of different participants was saved in individual cells under the same column. First, the researcher identified and marked the boundary of each terminable TC-unit and each single TC-unit. Provided with the marked boundaries, in the Excel file each single TC-unit was saved in separate cells under one column. All the punctuation marks or any additional spaces were removed to assure they would not confuse the length count in characters. Second, a formula “=LEN(A1)” was then applied in the Excel file to generate the length of each single TC-unit (dependent and independent) in the amount of characters. Some of these single TC-units were independent single TC-units which themselves each formed a simple terminable TC-unit. Others were dependent single TC-units, of which two or more composed a complex terminable TC-unit via correferential zero(s). The length of each terminable TC-unit (simple and complex) was also calculated in the amount of characters. The total amount of complex terminable TC-units and the total amount of the single TC-units (dependent and independent) in each participant’s output were counted by task as well. Lastly, scores on each of the four measures were calculated. For the length measures of ① MLTTCU and ③ MLSTCU, the average length of all the terminable TC-units (simple and complex) or all the single TC-units (dependent and independent) was calculated in the form of a

mean score. For the ratio measure ② $CTTCU/ATTCU$, of each participant's output by task, the total amount of complex terminable TC-units was divided by the the total amount of terminable TC-units (simple and complex). For scores on the other ratio measure ④ $STCU/TTCU$, of each participant's output by task, the total amount of single TC-units (dependent and independent) was divided by the total amount of the terminable TC-units (simple and complex). A coding sample is provided next to illustrate these three steps.

Output (59) is the first time transcription of the spoken output produced by a participant for the CS task. Table 7 shows the first time coding for (59). As the first coding step, the boundaries of each single TC-unit were segmented and marked by the researcher for further character counting and score calculation. As discussed in Section 3.3, the beginning and end points of a terminable TC-unit are not necessarily concurrent with the conventional sentence boundaries marked by punctuation marks. Therefore, when identifying the boundary of a terminable TC-unit or a single TC-unit, punctuation marks were not the major concern due to the arbitrariness of a Chinese sentence boundary. In the first column of Table 7, each single TC-unit was segmented and saved in an individual cell. When the topic of a single TC-unit is noted only once in its full form and is repeated in the form of coreferential zero in the preceding or following single TC-units, these consecutive single TC-units then become dependent single TC-units and together compose a complex terminable TC-unit. Upon the introduction of a new topic, or repetition of a topic in its full form, a pronoun, or demonstrative, instead of coreferential zero, a new terminable TC-unit is then activated. A total of 12 single TC-units were segmented in Table 7. In the column designated *STCU*, in each cell the numeral "1" was

assigned to mark each single TC-unit listed. The length of each single TC-unit was then counted in characters and marked in the column of *STCU length (in characters)*. Each terminable TC-unit was then assigned the numeral “1” in the column of *TTCU*. For example, the topic of the first single TC-unit, *māma* (妈妈, mom), was repeated in the form of coreferential zero in the second single TC-unit and therefore, these two consecutive single TC-unit thus became dependent single TC-units which then composed one complex terminable TC-unit as marked in the column of *CTTCU*. The length of each terminable TC-unit was then counted in the column of *TTCU length (in characters)*. In Table 7 all the punctuation marks or additional spaces for each single TC-unit listed in the first column were removed. Note the topics of each single TC-unit were marked in Table 7 only to exemplify the segmentation; they were however not included in the actual coding for length count in the Excel files themselves. In addition to the first terminable TC-unit which consisted of two single dependent TC-units and was marked by the numeral “1” in the column designated *CTTCU*, another complex terminable TC-unit in this particular output transcription was composed by the last two single TC-units as marked in Table 7.

(59) *Zài zhège mǎnhuà lǐmiàn, māma háoxiàng gāng zhǔnbèi hǎo wǎncān, bà tā dài dào zhuōzi shàng. Bàba yǐjīng zuò xià zhǔn bèi chīfànle. Nà kěshì hái zǐ huán méi lái, suǒyǐ ne, māma jiào bàba qù zhǎo zài gébifángjiān de háizi. Bàba fāxiàn háizi zài nà'er pāzhe kànshū, tā bǎ háizi jiào qù chīfàn qùle. Kěshì ne, shū yě yīnqǐle bàba de zhùyì, hòulái wǒmen kàn dào le māma hé háizi zuò zài zhuōzi pángbiān dēngzhe bùzài de bàba, ránhòu māma jiào háizi qù zhǎo bàba. Hái zǐ huí dào yuán chù kànshū dì dìfāng, fāxiàn bàba zhèng pā zài dīshàng kànzhe háizi gāng zài kàn de shū.*

在这个漫画里面，妈妈好像刚准备好晚餐，把它带到桌子上。爸爸已经坐下准备吃饭了。那可是孩子还没来，所以呢，妈妈叫爸爸去找在隔壁房间的孩子。爸爸发现孩子在那儿趴着看书，他把孩子叫去吃饭去了。可是呢，书也引起了爸爸的注意，后来我们看到了妈妈和孩子坐在桌子旁边等着不在的爸爸，然后妈妈叫孩子去找爸爸。孩子回到原处看书的地方，发现爸爸正趴在地上看着孩子刚在看的书。

[At this cartoon strip, the mother seem to prepare-done dinner, BA it bring to table top. The father already sit down ready to eat LE. Then but the child still not come, so, the mother ask farther go look for at next-door room DE child. The father find the the child at there lie face down ZHE read book, he BA the child called to eat meal LE. However, the book also catch the father's attention, later we see LE the mother and the child sit at table side wait Zhe not there DE father, and then the mother ask the child go look for the father. The Child return to original place read-book DE place, finding the father in the course of lying face down read ZHE the child just at read DE book.]

In this cartoon strip, the mother seems to have just served the dinner and brought them to the table. The father has already sat down and is ready to eat. Yet the child is not here yet, so the mother asks the father to look for the child next door. The father finds the child there lying face down and reading a book. He calls the child to go to the dinner. However, the book also catches the father's attention. Later we see the mother and child sitting at the table and waiting for the father to come. After that the mother asks the child to look for the father. The child returns to where he was and finds

out that the father is lying face down and reading the book that the child was reading just now.

With the first two steps of identifying and counting completed, the third step was to then calculate the score on each of the four measures. The mean scores of the length measures ① MLTTCU and ③ MLSTCU were calculated by using a formula in the Excel file, in this case “=average(A2:A13)”. The two ratio measures were also calculated by using a formula. In this case “=2/10” for ② CTTCU/ATTCU, and “=12/10” for ④ STCU/TTCU. Examples for the results of these calculations are listed in Table 8.

Six months after the first coding, a second coding was conducted by the researcher. For each task, the scores each participant received in the two separate codings were correlated in order to investigate intra-rater reliability. For example, for the CS task, all the participants’ scores on the measure ① MLTTCU were correlated to see how consistent the first and second codings were, and then the same was done with the measure ② CTTCU/ATTCU, ③ MLSTCU, and ④ STCU/TTCU. In Table 9 below, the intra-rater reliability indices between the two codings for each task of each participant on each measure are listed, with correlation coefficients ranging between .85 and .99 ($p = .000$).

When coding the spoken and written output collected in terms of TC-unit, some specific language phenomena were found requiring particular attention. Overall, such special cases fell within two main categories: 1) being coded as one single TC-unit; and 2) being coded as separate single TC-units. The category of being coded as one single TC-unit subsumes four conditions: a) Chinese chengyu, b) adverbial modifier, c) verb +

Table 7

A coding sample for output (59)

| | STCU | STCU length (in characters) | TTCU | TTCU length (in characters) | CTTCU |
|---|------|--------------------------------|------|--------------------------------|-------|
| 在这个漫画里面妈妈 ₁ 好像刚准备好晚餐 at this cartoon strip the mother ₁ seem to prepare-done dinner, BA it bring to table top | 1 | 17 | 1 | 24 | 1 |
| ∅ ₁ 把它带到桌子上 ∅ ₁ BA it bring to table top. | 1 | 7 | | | |
| 爸爸 ₂ 已经坐下准备吃饭了 the father ₂ already sit down ready to eat LE | 1 | 12 | 1 | 12 | 0 |
| 那可是孩子 ₃ 还没来 then but the child ₃ still not come | 1 | 9 | 1 | 9 | 0 |
| 所以呢妈妈 ₄ 叫爸爸去找在隔壁房间的孩子 so the mother ₄ ask farther go look for at next-door room DE child | 1 | 18 | 1 | 18 | 0 |
| 爸爸 ₅ 发现孩子在那儿趴着看书 the father ₅ find the the child at there lie face down ZHE read book | 1 | 14 | 1 | 14 | 0 |
| 他 ₆ 把孩子叫去吃饭去了 he ₆ BA the child called to eat meal LE | 1 | 11 | 1 | 11 | 0 |
| 可是呢书 ₇ 也引起了爸爸的注意 | 1 | 13 | 1 | 13 | 0 |

| | | | | | |
|---|------|-------|------|-------|-----|
| However the book ₇ also catch the father's attention | | | | | |
| 后来我们 ₈ 看到了妈妈和孩子坐在桌子旁边等着不在的爸爸 later we ₈ see LE the mother and the child sit at table side wait Zhe not there DE father | 1 | 26 | 1 | 26 | 0 |
| 然后妈妈 ₉ 叫孩子去找爸爸 and then the mother ₉ ask the child go look for the father | 1 | 12 | 1 | 12 | 0 |
| 孩子 ₁₀ 回到原处看书的地方 the Child ₁₀ return to original place read-book DE place | 1 | 12 | | | |
| Ø ₁₀ 发现爸爸正趴在地上看着孩子刚在看的书 Ø ₁₀ finding the father in the course of lying face down read ZHE the child just at read DE book | 1 | 18 | 1 | 30 | 1 |
| Total | (12) | (169) | (10) | (169) | (2) |

Table 8.

Calculated scores of output (59) on four measures

| Task | ID | ① MLTTCU | ② CTTCU/ATTCU | ③ MLSTCU | ④ STCU/TTCU |
|------|----|-----------------|---------------|-----------------|---------------|
| CS | 1 | 16.90 (=169/10) | 0.20 (=2/10) | 14.08 (=169/12) | 1.20 (=12/10) |

Table 9

Correlation coefficient between the two codings for each task on each measure

| | ① MLTCU | ② CTTCU/ATTCU | ③ MLSTCU | ④ STCU/TTCU |
|-----------|---------------------|---------------------|---------------------|---------------------|
| CS | .98* ($p = .000$) | .85* ($p = .000$) | .92* ($p = .000$) | .90* ($p = .000$) |
| V1 | .99* ($p = .000$) | .89* ($p = .000$) | .95* ($p = .000$) | .92* ($p = .000$) |
| V2 | .99* ($p = .000$) | .92* ($p = .000$) | .96* ($p = .000$) | .95* ($p = .000$) |
| FW | .97* ($p = .000$) | .94* ($p = .000$) | .92* ($p = .000$) | .95* ($p = .000$) |
| GR | .98* ($p = .000$) | .99* ($p = .000$) | .95* ($p = .000$) | .94* ($p = .000$) |

Note: *. Correlation is significant at the .01 level (2-tailed).

TC-unit(s) within one intonation contour, and d) serial verb constructions of one intonation contour. The category of being coded as separate single TC-units subsumes three conditions: e) direct quotation, f) verb + TC-unit(s) outside one intonation contour, and g) serial verb constructions of different intonation contours. All of these special cases are listed in Table 10 below and analysis is provided for further clarification.

a) Chinese *chengyu*. In Chinese, words, phrases, and sentences share a similar composing structure (Zhu, 1985). The distinction between words and phrases in Chinese is therefore not as clear as it is in English or other Indo-European languages. The traditional Chinese idiomatic expressions, *chengyu*, are categorized as phrases in this dissertation. In (60) for example, *chengyu* “知书达理” was analyzed as being the comment part of a single dependent TC-unit “她_i知书达理”. The comment part “知书达理” paralleled the other comment parts in this terminable complex TC-unit connected via

correferential zeros: “ \emptyset_i 性格温和, \emptyset_i 孝顺老人, \emptyset_i 敦睦亲邻, \emptyset_i 勤于家务, \emptyset_i 在亲友中广受尊重”.

b) Adverbial modifier. As discussed in Section 2.2.2.2, Chinese adverbial modifiers in this dissertation are analyzed as being subclausal level complexifying. In (61) for example, though “年看到火” by itself was a single TC-unit, once “的时候” added to it, “年看到火的时候” as a whole served as the time modifier of “它就害怕”. The complete unit (61), “年看到火的时候它就害怕” was then analyzed as being one terminable TC-unit.

c) Verb + TC-unit(s) within one intonation contour. When a verb takes one or more TC-units within one intonation contour, such TC-units were then analyzed as the argument of the verb. Therefore, verb + TC-unit(s) under one intonation contour were coded as one single TC-unit. In the previously listed example (34) for example, “他都很騰{疼}我也愛我” was the argument of the verb “知道” within one intonation contour, and (34) was then coded as one single TC-unit. No matter how complex the object was, it was coded as an embedded part of a single TC-unit as long as it stayed within the same intonation contour as the verb. As in (62), “爸爸正像自己一样趴在地上看着原先的那一本书” was a very complex argument of the verb “看见”, so (62) was coded as one single TC-unit. In contrast, when the argument of a verb went outside the intonation contour such as in the cases (35) and (66), they were then coded as two separated single TC-units. In (35), since “我和我的哥哥出去玩” was outside of the intonation contour of the verb “計的{记得}”, it was not considered embedded within an upper-level single TC-unit, but it was coded as a second single TC-unit in (35) along with the first single TC-

Table 10

Some special marks for coding

| | | Stranscribed spoken output or collected written data | Coding with single TC-unit marked |
|-----------------------------------|---|---|--|
| One single TC-unit | a. Chinese <i>chengyu</i> | (60) 她 _i 知书达理, \emptyset_i 性格温和, \emptyset_i 孝顺老人, \emptyset_i 敦睦亲邻, \emptyset_i 勤于家务, \emptyset_i 在亲友中广受尊重。 [She _i is well-educated and show a good sense of judgement, \emptyset_i disposition gentle, \emptyset_i shows filial piety to elderlies, \emptyset_i respect and gets long with neighbors, \emptyset_i is diligent at housework, \emptyset_i at relatives and friends widely receives respect.] | (60a) 她 _i 知书达理/STCU-1, \emptyset_i 性格温和/STCU-2, \emptyset_i 孝顺老人/STCU-3, \emptyset_i 敦睦亲邻/STCU-4, \emptyset_i 勤于家务/STCU-5, \emptyset_i 在亲友中广受尊重/STCU-6。 |
| | b. Adverbial modifier | (61) 年看到火的时候它 _i 就害怕。 [Nian see fire DE time it _i then fears.] | (61a) 年看到火的时候它 _i 就害怕。/STCU-1 |
| | c. Transitive verb + TC-unit(s) within one intonation contour | (34) 我 _i 知道他都很腾{疼}我也爱我。 [I know _i he all very dear me also love me.] (62) 小明 _i 推开门看见爸爸正像自己一样趴在地上看着原先的那一本书。 [Xiaoming _i push-open the door see father is like himself same lie face down read original that one CL book.] | (34a) 我 _i 知道他都很腾{疼}我也爱我。/STCU-1 (62a) 小明 _i 推开门看见爸爸正像自己一样趴在地上看着原先的那一本书。/STCU-1 |
| | d. Serial verb constructions of one intonation contour | (36) 妈妈 _i 叫你去吃饭。 [Mother _i ask you go eat meal.] (63) 有一个怪物 _i 住在山上。 [Has a CL monster _i live at mountain top.] (64) 你 _i 越急他越开心。 [You _i more rushed he more cheerful.] | (36a) 妈妈 _i 叫你去吃饭。/STCU-1 (63a) 有一个怪物 _i 住在山上。/STCU-1 (64a) 你 _i 越急他越开心。/STCU-1 |
| | e. Direct quotation | (65) 爸爸 _i 说: “小孩子 _j 在哪儿。我 _g 去找他。” | (65a) 爸爸 _i 说/STCU-1: “小孩子 _j 在哪儿/STCU-2。 |

| | | | |
|-------------------------------------|---|---|---|
| More than one single TC-unit | | [Father _i say, "Little Child _j at where. I _g go look for him."] | 我 _g 去找他/STCU-3。” |
| | f. Transitive verb + clause outside one intonation contour | (35) 我 _i 計的{记得}有一此{次}, 我和我的哥哥 _j 出去玩。 [I _i remember has once, I and my elder brother _j go out play.] (66) 那他们 _i 就发觉年一看到鞭炮爆发它就会退步, 而且它 _j 也很不喜欢红色的颜色。 [Then they _i just realize <i>nian</i> once see fire crackers explode it then will retreat, in addition it _j also very not like red color.] | (35a)我 _i 計的{记得}有一此{次}/STCU-1, 我和我的哥哥 _j 出去玩/STCU-2。 (66a) 那他们 _i 就发觉年一看到鞭炮爆发它就会退步/STCU-1, 而且它 _j 也很不喜欢红色的颜色/STCU-2。 |
| | g. Serial verb constructions of different intonation contours | (37) 妈妈叫你 _i , ∅ _i 去吃饭。 [Mother call you, go eat meal] (67) 有一个怪物 _i , ∅ _i 住在山上。 [Has a CL monster, live at mountain top.] (68) 你 _i 越急, 他 _j 越开心。 [You _i more rushed, he _j more cheerful.] | (37a) 妈妈叫你 _i /STCU-1, ∅ _i 去吃饭/STCU-2。 (67a) 有一个怪物 _i /STCU-1, ∅ _i 住在山上/STCU-2。 (68a) 你 _i 越急/STCU-1, 他 _j 越开心/STCU-2。 |

unit “我計的{记得}有一此{次}”. Also in (66), “年一看到鞭炮爆发它就会退步” was inside the intonation contour of the verb “发觉”, so “那他们就发觉年一看到鞭炮爆发它就会退步” was coded as the first single TC-unit. The rest of (66), “而且它也很不喜欢红色的颜色” was outside the intonation contour of the verb “发觉” and therefore, it was coded as the second single TC-unit of (66).

Serial verb constructions were also coded accordingly depending on whether the verb constructions stayed within or were outside the same intonation contour. For d), serial verb constructions of one intonation contour i.e. (36), (63), and (64) were all coded as one single TC-unit each since in each of these sentences both verb structures were within the same intonation contour. For g), serial verb constructions of different intonation contours i.e. (37), (67), and (68) were each treated as two separated single TC-units since the two verb structures were not covered in one intonation contour.

Lastly, for direct quotations the quoted part was treated as a separate single TC-unit. As shown in (65), “爸爸说” was analyzed as the first single TC-unit, and what “爸爸” said within the quotation marks was analyzed as being two other single TC-units. “小孩子在哪儿” was coded as the second single TC-unit, and “我去找他” was coded as the third single TC-unit.

Note that the discussion here has remained at the coding level of single TC-units. The coding of terminable TC-units is also dependent upon the topics of the preceding or subsequent single TC-units and more specifically, whether the same topic is repeated in the form of coreferential zero. Therefore, the coding of terminable TC-units in the examples shown in Table 10 could not be completed without their corresponding complete output being provided.

4.3.4 Participant classification

For the spoken Chinese output data, this dissertation collected and utilized a total of 115 complete raw data sets for analysis (66 females; 49 males). Of these, a total of 84 raw data sets belonged to L2 Chinese speaker participants and 31 raw data sets belonged to L1 Chinese speaker participants. For the written Chinese output data, a total of 116 (65 females, 51 males) complete raw data sets were collected and of these, a total of 84 raw data sets belonged to L2 Chinese speaker participants and 32 raw spoken data sets belonged to L1 Chinese speaker participants. A complete raw speaking data set consisted of a background survey, the output of three speaking tasks (CS, V1 and V2), the retrospective survey, and the Mandarin EI test recording. A complete raw written data set consisted of a background survey, the output of two written tasks (FW and GR), a retrospective survey, and a Mandarin EI test recording.

A total of 86 L2 Chinese speaker participants' raw data sets were collected complete in the form of either spoken or written data. Every such complete raw data set of L2 participants included an EI test file where each L2 participant's response to the EI test were recorded individually. Each EI test (as described in Section 3.3) consisted of 30 responses to the 30 stimuli sentences. Two raters rated the 86 Mandarin EI test files of these 2580 (= 86 x 30) responses referring to the same rubric (See Appendix E). Before started rating all the test files, the two test raters piloted 10 participants separately, compared their rating for each item in the EI test, and made sure they were using the rubric in a consistent way. After separately completing all the scoring, it was found that 1,938 single responses out of the 2,580 single responses for all the participants were assigned the identical scores by two raters. Therefore, the inter-rater reliability was

satisfactory with a substantial agreement rate of 75.1%. The mean score of the two ratings by the two raters was then assigned to all the participants as their final EI score. All statistical data processing in this dissertation was conducted within SPSS version 21. As shown in Figure 8, the EI scores of all the L2 Chinese speaker participants fell clearly in a bimodal distribution, with a lower proficiency group and a higher proficiency group, of which the cut-point fell right in the middle of the test, at a score of 60 out of 120. All the L2 Chinese speaker participants were thus divided into two groups based on their Chinese proficiency: (a) Group Low: EI Score < 60 ($n = 38$, $M = 35.16$, $SD = 12.17$); (b) Group High: EI Score > 60 ($n = 48$, $M = 87.96$, $SD = 15.13$). Comparing the mean of the EI score of Group Low and Group High, a t -test showed a statistically significant difference with $t(84) = -17.48$, $p = .000$ ($p < .05$). There was a substantial and meaningful difference between Group Low and Group High in terms of Chinese language

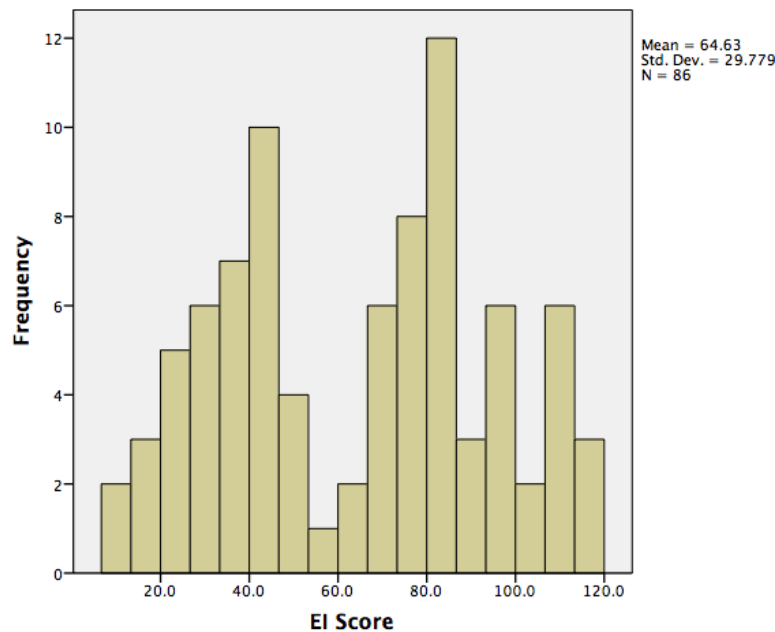


Figure 8. Mandarin EI score distribution of L2 Chinese speaker participants.

proficiency. All the L1 Chinese speaker participants were classified as Group Native ($n=32$).

Table 11

Total cases for speaking and writing tasks by proficiency group

| | Spoken | Written |
|--------------|---------------|----------------|
| Group Low | 37 | 36 |
| Group High | 47 | 48 |
| Group Native | 31 | 32 |
| Total | 115 | 116 |

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Data screening

Data screening was undertaken before performing discriminant function analysis. As stated in 4.2.1, a total of 115 complete sets of raw spoken data and 116 complete sets of raw written data were collected for this dissertation. In accordance with the different tasks, the spoken data were grouped into three sets: (a) CS task data, (b) V1 task data, and (c) V2 task data; and the written data were grouped into two sets: (d) FW task data, and (e) GR task data. According to the Chinese language proficiency of the participants, data for each task were grouped into Group Low, Group High, and Group Native.

Outliers

To identify univariate outliers, z scores of all the predictors were calculated and robustness of significance tests was investigated through SPSS DESCRIPTIVE for each task separately. In addition to the inspection of z scores, histograms of the scores on different measures of each group by task were checked for cases unattached to the rest of the distribution. Because this dissertation utilized a relatively large sample size for each group in this dissertation, cases with standardized scores in excess of 3.50 ($p < .001$, two tailed test) that also displayed disconnection with the rest of the distribution were considered univariate outliers, although there were very few such cases in the data set. Outliers were identified as follows: for the CS task data sets 1 case was identified on measure ④ STCU/TTCU in Group Low; for the V2 task 1 case was identified on

measure ④ STCU/TTCU in Group Low; for the FW task 1 case was identified on measure ④ STCU/TTCU in Group Low, 1 case on measure ① MLTTCU, and 1 case was identified on measure ③ MLSTCU in Group High; for the GR task 1 case was identified on measure ④ STCU/TTCU in Group Low, 1 case was identified on measure ② in Group High, and 1 case was identified on measure ① MLTTCU and 1 case was identified on measure ④ STCU/TTCU in Group Native. These 9 individual scores out of 60 groups of scores in the data sets for 5 tasks were adjusted by making them one point higher than their next closest score, thereby bringing them into contact with the rest of the distribution in each data set. This approach of score adjustment was taken since it was the least invasive of the options for modifying score distributions.

The remaining data for the five tasks were then checked respectively by groups for multivariate outliers through SPSS REGRESSION. Cases with too large Mahalanobis D^2 for their own group, evaluated as $\chi^2(4) > 18.467$ ($\alpha = .001$) were identified as multivariate outliers, again resulting in the identification of only a few suspect cases. The score of 2 Group Low participants and 2 Group High participants on the CS task, 1 Group Low participant on the V1 task, 1 Group Low participant and 1 Group High participant on the V2 task, 1 Group Low participant and 1 Group High participant on the FW task, and 2 Group Low participants and 1 Group High participant on the GR task were identified as multivariate outliers and eliminated from the respective data sets.

Based on the data sets with univariate outliers adjusted and multivariate outliers deleted, discriminant analyses were conducted to see how the four measures work at correctly predicting the participants' proficiency group membership. Since discriminant analysis is quite robust and not that affected by nonnormal distributions, especially when

the sample sizes in this dissertation were large enough, this dissertation therefore employed discriminant analyses on the unadjusted data.

The results of discriminant analyses on the data with or without adjustment were found to be quite similar as listed in Table 12 below. On the unadjusted data sets, the four measures were able to obtain 67.8%, 75.7%, 76.5%, 67.2%, and 61.2% correct group membership classification for the CS, V1, V2, FW, and GR tasks respectively. On the adjusted data sets, the four measures were able to obtain 66.7%, 74.6%, 77.0%, 68.4%, and 60.2% correct group membership classification for the CS, V1, V2, FW, and GR task respectively. Compared with the analyses on the unadjusted data, analyses on the adjusted data increased by 0.5% and 1.2% in correct group membership classification for the task V2 and FW, but decreased by 1.1%, 1.1%, and 1.0% for the CS, V1, and GR tasks respectively. Considering that the results were so minimal that they did not seem to make any difference, this dissertation decided to keep the analyses on the unadjusted data especially since the sample sizes were large enough.

Table 12

The results of discriminant analyses on the data with or without adjustment

| | On the unadjusted data | On the adjusted data |
|-----------|-------------------------------|-----------------------------|
| CS | 67.8% (<i>N</i> =115) | 66.7% (<i>N</i> =111) |
| V1 | 75.7% (<i>N</i> =115) | 74.6% (<i>N</i> =114) |
| V2 | 76.5% (<i>N</i> =115) | 77.0% (<i>N</i> =113) |
| FW | 67.2% (<i>N</i> =116) | 68.4% (<i>N</i> =114) |
| GR | 61.2% (<i>N</i> =116) | 60.2% (<i>N</i> =113) |

Normality The distributions of data in the 60 cells of this design (5 tasks*3 proficiency groups*4 measurements) were checked based by their skewness and kurtosis, as well as graphically by histogram. The distributions were relatively normal for most of the measures, with greater skewness and kurtosis on measure ④ STCU/TTCU for Group Low on the FW task and Group Native on the GR task. However, as Tabachnick and Fidell (2012) pointed out:

For grouped data, it is the sampling distribution of the means of variables that are to be normally distributed. The Central Limit Theorem reassures us that, with sufficiently large sample sizes, sampling distributions of means are normally distributed regardless of the distribution of variables. For example, if there are at least 20 degrees of freedom for error in a univariate ANOVA, the F test is said to be robust to violations of normality of variables (provided that there are no outliers). (p. 78)

As shown in Table 11, the sample size of each group is between 31 and 48. In each group for every task in this dissertation, the degrees of freedom all surpassed 20. Hence, the sampling distributions of means in this dissertation were considered not violating the normality assumption.

Homogeneity of variance

After the univariate and multivariate outliers were eliminated, homogeneity of variance was assessed with F_{\max} in conjunction with sample-size ratios. If sample sizes are relatively equal (within a ratio of 4 to 1 or less for largest to smallest cell size), an F_{\max} as great as 10 is acceptable (Tabachnick & Fidell, 2012, p. 86). The sample sizes for

all the cells in this dissertation were quite similar, with the largest cell size of 47 and the smallest cell size of 31. For every task, the F_{\max} values were found to be much lower than 10. Thus, there was no problem with the homogeneity of variance assumption for the data in this dissertation.

Homogeneity of variance-covariance matrices

The multivariate generalization of homogeneity of variance for individual DVs is found in the estimation of homogeneity of variance-covariance matrices. As shown in Table 13 below, Box's M analysis was statistically significant for each task, which suggested departures from homogeneity of variance-covariance. However, based on the Monte Carlo test of robustness for T^2 (Hakstian, Roed, & Lind, 1919), Tabachnick and Fidell (2012, p. 254) argued that if sample sizes are equal, robustness of significance tests is expected and it is legitimate to disregard the outcome of Box's M test, a notoriously sensitive test of homogeneity of variance-covariance matrices. With each cell having 31-47 cases, the sample sizes in this dissertation were quite similar. Therefore, robustness to violations of homogeneity of variance-covariance matrices was expected.

Table 13

Box's M output across three groups for different tasks

| | CS | V1 | V2 | FW | GR |
|----------------|--------------|--------------|--------------|--------------|--------------|
| Box's M | 119.35 | 120.39 | 83.14 | 143.95 | 193.99 |
| | ($p=.000$) | ($p=.000$) | ($p=.000$) | ($p=.000$) | ($p=.000$) |

Linearity

For each task, the linearity of relationships among all pairs of predictors was examined using SPSS PLOT for each pair. No markedly non-linear relationship was found.

Multicollinearity

As displayed in Table 14 below, the Pearson product-moment correlation matrix of all the predictors with each other was examined by task to check multicollinearity. For all the five tasks, the bivariate correlations between pairs of all the predictors were lower than .90. For the speaking tasks, the correlation between ② CTTCU/ATTCU and ④ STCU/TTCU were relatively high ranging from .856- .896. A likely cause might be that measures ② CTTCU/ATTCU and ④ STCU/TTCU were both ratio measures with relatively narrow score distributions compared with the length measures. In addition, the range of scores on measure ② CTTCU/ATTCU were fixed within 0 to 1 which potentially led to higher correlations between the scores on these two measures. However, if the only goal of the discriminant function analysis is prediction, multicollinearity can be ignored (Tabachnick & Fidell, 2012, p. 91). Since this dissertation was designed to predict proficiency group membership from the four Chinese syntactic complexity predictors, multicollinearity did not appear to be a problem.

After cases with univariate and multivariate extreme outliers were eliminated in each task, no worrisome violations of the assumptions of discriminant function analysis remained in this dissertation, therefore subsequent inferential analyses were undertaken.

Table 14

Pearson correlation matrix of all the predictors with each other by task

| | | ②CTTCU/ ATTCU | ③MLSTCU | ④STCU/ TTCU |
|-----------|--------------|------------------|---------|----------------|
| CS | ①MLTTCU | .499 | .780 | .504 |
| | ②CTTCU/ATTCU | | -.054 | .896 |
| | ③MLSTCU | | | -.097 |
| V1 | ①MLTTCU | .618 | .641 | .633 |
| | ②CTTCU/ATTCU | | -.065 | .873 |
| | ③MLSTCU | | | -.146 |
| V2 | ①MLTTCU | .597 | .548 | .637 |
| | ②CTTCU/ATTCU | | -.168 | .856 |
| | ③MLSTCU | | | -.255 |
| FW | ①MLTTCU | .470 | .654 | .643 |
| | ②CTTCU/ATTCU | | -.152 | .775 |
| | ③MLSTCU | | | -.127 |
| GR | ①MLTTCU | .423 | .453 | .726 |
| | ②CTTCU/ATTCU | | -.061 | .572 |
| | ③MLSTCU | | | -.237 |

5.2 Results

5.2.1 Descriptive statistics

The mean values for each measure on five tasks by proficiency level were checked and are listed in Table 15 below. As also graphically displayed in Figure 9 below, each measure with varied power was able to distinguish varied syntactic complexity levels among proficiency groups Low, High, and Native.

Note that especially on measure ① MLTTCU the distributions of scores for each proficiency group in all five tasks were quite evenly spread which indicated the substantial power of measure ① MLTTCU for distinguishing the syntactic complexity levels of language output across proficiency groups. ① MLTTCU seemed to be a very strong indicator of the Chinese syntactic complexity itself. For all five tasks the mean score on measure ① MLTTCU ranged from 8.88-13.27 characters for the Group Low participants, 12.91-15.45 characters for the Group High participants, and 17.83-19.94 characters for the Group Native participants. Such mean score distributions on measure ① MLTTCU indicated that overall longer terminable TC-units were produced by participants of higher proficiency, while shorter terminable TC-units were produced by those having a lower proficiency.

The mean scores on measure ③ MLSTCU were mostly evenly spread out. For the three speaking tasks the mean score of the Group Low participants on measure ③ MLSTCU ranged from 8.21-8.96 characters, 10.38-10.89 characters for Group High, and 12.27-13.14 characters for Group Native. For the three speaking tasks this generally indicated that on average longer single TC-units were produced by participants of higher proficiency, while shorter single TC-units were produced by those having a lower

proficiency. As for each of the two writing tasks however, the distribution showed some crossover between different groups. In the FW task, the mean score of 11.69 characters for the Group High participants surpassed the mean score of 10.74 characters for Group Native participants on measure ③ MLSTCU. In the GR task, the mean score of 12.23 characters for Group Low was a little higher than the mean score of 12.16 characters for Group High participants on measure ③ MLSTCU. This suggested that for the writing tasks, there might be a nonlinear development in terms of the length of single TC-units as proficiency increases. Another possible interpretation for such score crossover is

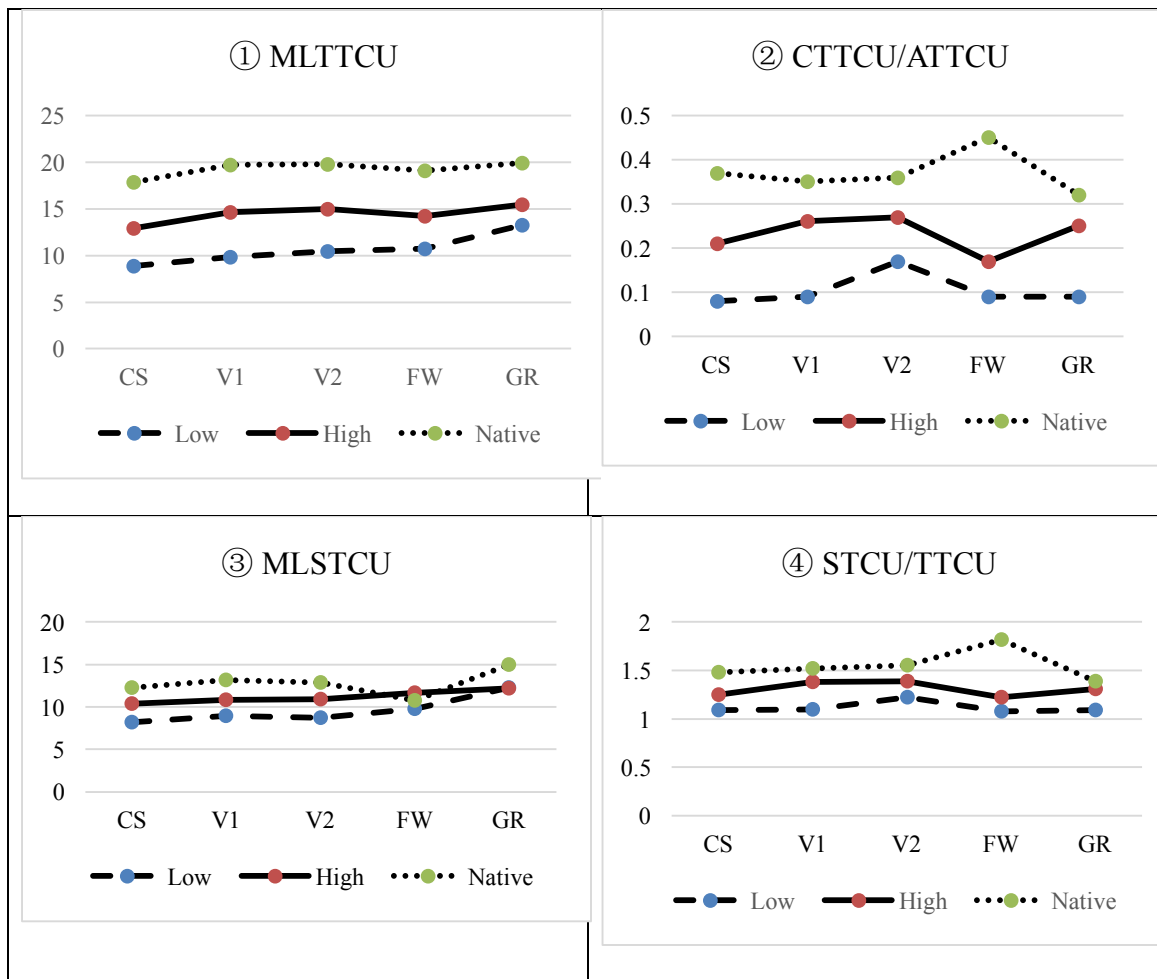


Figure 9. The mean values for each measure on five tasks by proficiency level.

Table 15

The mean values for each measure on five tasks by proficiency level

| | | ① | ② CTTCU | ③ | ④ STCU |
|-----------|-------------|--------|---------|--------|--------|
| | | MLTTCU | /ATTTCU | MLSTCU | /TTCU |
| | Low (37) | 8.88 | .08 | 8.21 | 1.09 |
| CS | High (47) | 12.91 | .21 | 10.38 | 1.25 |
| | Native (31) | 17.83 | .37 | 12.27 | 1.48 |
| | Low (37) | 9.82 | .09 | 8.96 | 1.10 |
| V1 | High (47) | 14.65 | .26 | 10.82 | 1.38 |
| | Native (31) | 19.74 | .35 | 13.14 | 1.52 |
| | Low (37) | 10.45 | .17 | 8.68 | 1.22 |
| V2 | High (47) | 14.97 | .27 | 10.89 | 1.39 |
| | Native (31) | 19.81 | .36 | 12.83 | 1.55 |
| | Low (36) | 10.73 | .09 | 9.77 | 1.08 |
| FW | High (48) | 14.23 | .17 | 11.69 | 1.22 |
| | Native (32) | 19.12 | .45 | 10.74 | 1.82 |
| | Low (36) | 13.27 | .09 | 12.23 | 1.09 |
| GW | High (48) | 15.45 | .25 | 12.16 | 1.31 |
| | Native (32) | 19.94 | .32 | 15.00 | 1.39 |

that for writing tasks, a longer single TC-unit does not equal more sophisticated or more native-like Chinese syntactic complexity.

On measure ② CTTCU/ATTCU for all five tasks, participants of higher proficiency groups scored higher than those of lower proficiency groups. For all five tasks the mean score of the Group Low participants on measure ② CTTCU/ATTCU ranged from .08-.17; .17-.27 for Group High; and .32-.45 for Group Native participants. Generally, this pattern suggested that on average a higher proportion of complex terminable TC-units was produced by higher proficiency participants than by lower proficiency participants. The FW task showed a somewhat different score distribution on this measure as shown in Figure 9 where the gap between the mean scores of Group High and Group Native was stretched when compared with the other tasks. In the FW task, the score of Group High on measure ② CTTCU/ATTCU was relatively lower than it was in the other four tasks, while the score of Group Native was higher than it was in the other four tasks.

Similarly on the other ratio measure ④ STCU/TTCU, for each task participants of higher proficiency scored higher than the participants of lower proficiency. For all five tasks, the mean score of the Group Low participants on measure ④ STCU/TTCU ranged from 1.08-1.22; 1.22-1.38 for Group High; and 1.39-1.82 for Group Native. This finding suggested that one terminable TC-unit on average consisted of more dependent single TC-units in the output of higher proficiency participants, while the terminable TC-units produced by lower proficiency participants consisted of less dependent single TC-units. Similar to the distribution on measure ② CTTCU/ATTCU for the FW task, the gap between the average scores of Group High and Group Native was stretched compared with the other four tasks on measure ④ STCU/TTCU. This as well was caused by the relatively lower average score of Group High compared with the relatively higher score

of Group Native. This indicated a relatively large gap between Group High and Group Native participants in terms of producing complex terminable TC-units consisting of more dependent single TC-units. In order to produce more sophisticated or native-like Chinese of high syntactic complexity, the advanced L2 Chinese speakers need to produce more complex terminable TC-units consisting of more dependent single TC-units.

5.2.2 Correlational Analysis

For each of the different tasks, this dissertation then tested how the scores on four complexity measures correlated with L2 Chinese speaker participants' global Chinese proficiency level in terms of their EI score. Group Native was not included in this correlational analysis since the Chinese native speakers did not take the EI tests. Table 16 below shows all the Pearson correlation coefficients between the EI scores of L2 Chinese

Table 16

Correlation between EI scores of L2 Chinese speakers and their scores on the four complexity measures by task

| EI Score | ① MLTTCU | ② CTTCU /ATTTCU | ③ MLSTCU | ④ STCU /TTCU |
|------------------|---------------|--------------------|--------------|-----------------|
| CS (N=84) | .59* (p=.000) | .44*(p=.000) | .47*(p=.000) | .48*(p=.000) |
| V1 (N=84) | .79* (p=.000) | .72* (p=.000) | .57*(p=.000) | .69* (p=.000) |
| V2 (N=84) | .77* (p=.000) | .43* (p=.000) | .63*(p=.000) | .47* (p=.000) |
| FW (N=84) | .41* (p=.000) | .24* (p=.027) | .31*(p=.004) | .33*(p=.002) |
| GR (N=84) | .38* (p=.000) | .33* (p=.002) | .01(p=.995) | .47*(p=.000) |

Note: *. Correlation is significant at the .05 level (2-tailed).

speaker participants and four complexity measure scores by task. Note that most of the correlations were moderate to strong in magnitude, and most of the correlations were statistically significant at the $p < .05$ level (2-tailed). Such positive linear relationships between global proficiency and the four complexity measure scores were consistent with the primary assumption that the higher proficiency Chinese speakers are capable of producing Chinese output of higher syntactic complexity.

First, among the four measures, scores on the global measure ① MLTTCU stood out demonstrating consistently stronger Pearson correlation coefficients with the EI proficiency scores of L2 Chinese speaker participants for each task. For the three speaking tasks in particular, the correlation between scores on measure ① MLTTCU and EI scores ranged from .59-.79. The coefficients of determination r^2 ranged from .35- .62 which indicated that about 35% - 62% of the variance on the EI test scores of L2 Chinese speaker participants can be predicted only by the scores on measure ① MLTTCU. For the two written tasks, relatively lower correlations were found in comparison to the speaking tasks. Still, the global measure ① MLTTCU showed the highest correlation with proficiency scores ranging from .38- .41. This indicated that about 14% - 17% of the variance on the EI test scores of L2 Chinese speaker participants could be predicted only by the scores on measure ① MLTTCU for the writing task.

Second, a contrast exists between the syntactic complexity of different levels. For Chinese syntactic complexity at the global and clausal levels, the length and ratio measures demonstrated correlations of varied power. At the global complexity level, the length measure ① MLTTCU showed relatively higher correlations with proficiency level than the ratio measure ② CTTCU/ATTCU. At the clausal complexity level by contrast,

the ratio measure ④ STCU/TTCU showed relatively higher correlations with proficiency level when compared with the length measure ③ MLSTCU. There was one exception that occurred on the V2 task and it will be discussed separately in Section 5.2.4.2. In addition, if compared cross-sectionally by task, length measure ① MLTTCU at the global complexity level showed higher correlation than the clausal level length measure ③ MLSTCU. However, ratio measure ④ STCU/TTCU at the clausal level showed higher correlation than the global level ratio measure ② CTTCU/ATTCU. It might have been that Chinese syntactic complexity at different levels showed uneven increases in terms of the embeddedness and compositionality of grammatical structures that were detected by measures of different types. At the global complexity level, producing generally longer simple terminable TC-units contributed more to the increase of Chinese syntactic complexity when compared with composing increased amounts of complex terminable TC-units from all the terminable TC-units. In order to produce longer simple terminable TC-units, more elaboration at the subclausal level can be applied such as adding more attributive and adverbial modifiers, forming more serial verb constructions or adding arguments within one intonation contour, and by using more sophisticated phrases or structures. By contrast, at the clausal level, it was found that when compared with lengthening each single TC-unit, combining more dependent single TC-units into a terminable TC-unit more effectively contributed to the increase of Chinese syntactic complexity. It could be that Chinese syntactic complexity at different levels showed varied patterns of continual increase with proficiency and indeed, it could even be that clausal complexity levels fall off at some point or actually drop on length measure as proficiency increases.

Third, comparing tasks in different modalities, the speaking tasks of CS, V1, and V2 showed generally higher correlations between the EI score and the four complexity measure scores than the writing tasks of FW and GR. For the three speaking tasks, the correlation coefficients ranged from .43- .79. The coefficients of determination r^2 ranged from .18- .62, which indicated that 18% to 62% of the variance on the EI test scores of L2 Chinese speaker participants can be predicted by any one of the four complexity measures. By contrast, the writing tasks FW and GR indicated lower correlations between the EI score and the four complexity measure scores and ranged from .01- .47. Among these, the GR task correlation between ③ MLSTCU and EI scores showed a relationship which was not statistically significant, $r = .01$ ($p = .995$). For the rest of the correlations between the measures and EI scores for both writing tasks, the correlation coefficients ranged from .24- .47. The coefficient of determination, r^2 , ranged from .06- .22 which indicated that 6% to 22% of the variance on the EI score could be predicted by any one of the complexity measures. To interpret such differences between speaking and writing tasks, one possible explanation is that the speaking tasks were better designed than the writing tasks for the TW&ST with regards to eliciting the Chinese syntactic complexity of L2 Chinese speakers. By nature, writing takes a longer time than speaking. To achieve higher complexity, one usually uses more time writing compared with speaking at the stage of formulation with regards to selecting, polishing, and structuring at the lexical, phrasal, and sentential levels. Though longer time was given for the two writing tasks (10 minutes for the FW task and 7 minutes for the GR task) compared with the speaking tasks (1.5 minutes for the CS task, 3 minutes for the V1 task, and another 3 minutes for the V2

task), the time allocated to the two writing tasks was still considered to be restricted when compared with the time allocated for a typical classroom writing assignment.

Fourth, among the three speaking tasks however, there were very different magnitudes of correlation for on CS task than for the V1 and V2 tasks. The CS task demonstrated lower correlations between the four measures and the EI proficiency scores than for the V1 and V2 tasks. On measure ① MLTTCU especially, the CS task showed a much lower correlation of .59 between proficiency scores compared with the higher correlations of .79 and .77 found on the V1 and V2 task. As discussed in Section 4.2.2, the three speaking tasks were designed having varied cognitive complexity according to the Multiple Resources Attentional Model. Therefore, there seemed to be a real difference in what complexity learners are able to achieve syntactically while completing tasks of varied cognitive complexity.

As one dimension of the triad of CAF measures, the syntactic complexity measure can of course at best only partly predict variance in language proficiency. By including the measures of accuracy and fluency, such predictions would likely be much higher and more able to strongly predict the variance in language proficiency. In addition, considering the multifaceted nature of syntactic complexity itself, consisting of global, clausal, and subclausal/phrasal complexity, it may be that still only part of the complexity picture has been detected by the current measures. ① MLTTCU and ② CTTCU/ATTCU detect the global complexity, and ③ MLSTCU and ④ STCU/TTCU detect clausal complexity. Therefore, when assessing Chinese complexity, a combination of multiple complexity measures can be applied to generate a more comprehensive score. Third, considering the varied correlation on each measure by task, it might have been that

different task designs played different roles in eliciting complex language output.

Cognitive task complexity along the lines of resource-directing and resource-dispersing elicited language output of varied syntactic complexity.

5.2.3 Discriminant function analysis

Direct discriminant function analysis is used to predict group membership based on a set of predictors; in the current study, discriminant analysis was undertaken within each of the five tasks. Discriminant analysis (or discriminant function analysis) is essentially MANOVA turned around. In MANOVA, the groups' performance (dependent variables) differs based on the group membership (independent variables). In discriminant analysis, the other way around, the measures of different groups' performances are used as predictors (independent variables) for group membership (dependent variable). According to Norris (2015), "Discriminant Analysis or Discriminant Function Analysis provides a statistical approach to investigate the extent to which a set of measured variables can distinguish— 'discriminant'— between members of different groups or distinct levels of another, nominal or possibly ordinal, variables" (p. 309). Brown, Robson, and Rosenkjar (2001) commented, "Unlike multiple regression analysis, which is limited to a single linear dimension, discriminant function analysis can investigate predictions along more than one dimension" (p. 377). In this dissertation, three groups, Group Low, Group High, and Group Native, were predetermined by the participants' three Chinese proficiency levels based on their Mandarin EI scores or native speaker status. The four Chinese syntactic complexity measures were then applied by task as predictors of membership in the three proficiency groups, since higher proficiency

Chinese speakers should be capable of producing Chinese output with higher syntactic complexity. By investigating how well the membership can be correctly predicted by the syntactic complexity measures through discriminant analysis, a differential group approach to validation was undertaken. Since there were five different tasks in the TW&ST completed by participants, a separate Discriminant Analysis was conducted for each task, using SPSS. For all the discriminant analyses, prior probabilities were computed from group sizes.

5.2.3.1 Speaking tasks

A direct discriminant function analysis was first conducted for the CS task output. The analysis identified two discriminant functions, the first accounting for the large majority (96.2%) of observable between-groups variance across the three proficiency groups, and the second accounting for 3.8%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .419, $\chi^2(8, N=115) = 96.201, p = .000$. This indicated that the combined predictor variables were able to account for around 58% of the actual variance in proficiency level among the three groups. On its own, the second function did not provide additional statistically significant predictions, Wilks' lambda = .952, $\chi^2(3, N=115) = 5.491, p = .139$. Figure 10 displays the individual cases and group centroids (average values for each group) in two dimensions: (a) from left to right, Function 1 clearly distinguishes between all three groups; (b) from top to bottom, Function 2 additionally provides little distinction between Group High and the other two groups.

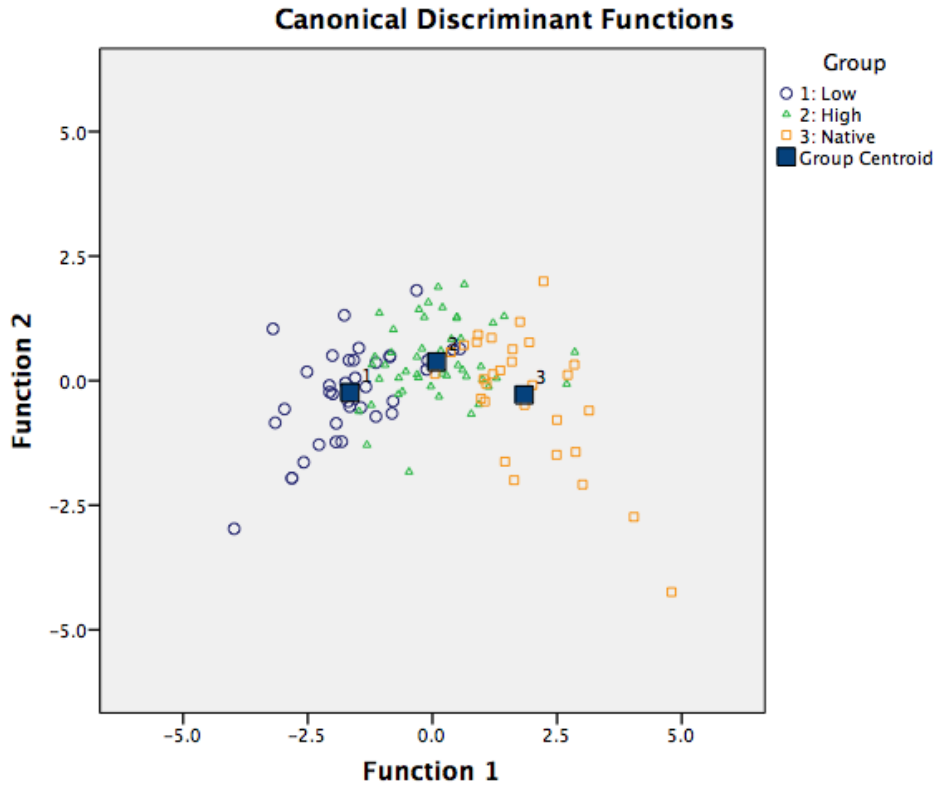


Figure 10. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the CS task.

The classification results for the CS task indicated that, overall, the combined Functions 1 and 2 were able to correctly classify 78 cases (or 67.8%) as shown in Table 17. However, the accuracy of the classifications varied for the three levels. Group Low participants were classified with 73.0% accuracy, while Group High participants were classified with 68.1% accuracy, and Group Native participants were classified correctly with 61.3% accuracy. For the CS task, the Group Native scores showed less correctly classified cases on the four complexity measures compared with Group Low and Group High with 38.7% of the Group Native cases being misclassified into Group High.

Table 17

Classification results for the CS task

| Actual Group | <i>N</i> | Predicted Group Membership | | |
|---------------------|----------|-----------------------------------|--------------|---------------|
| | | Low | High | Native |
| | | (67.8% correctly predicted) | | |
| Low | 37 | 27 | 10 | 0 |
| | | 73.0% | 27.0% | .0% |
| High | 47 | 8 | 32 | 7 |
| | | 17.0% | 68.1% | 14.9% |
| Native | 31 | 0 | 12 | 19 |
| | | .0% | 38.7% | 61.3% |

Note. The bold are the correct predictions.

For the discriminant analysis with the V1 task, both the combined functions (1 and 2) and the second function alone showed statistically significant effects in distinguishing proficiency level between the three groups. Function 1 accounted for the large majority (90.4%) of observable between-groups variance across the three proficiency groups, and Function 2 accounted for 9.6%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .270, $\chi^2(8, N=115) = 144.720, p = .000$, indicating that the combined predictor variables were able to account for around 73% of the actual variance in proficiency level between the three groups. On its own, the second function provided additional statistically significant predictions, Wilks' lambda = .822, $\chi^2(3, N=115) = 21.646, p = .000$. Figure 11 shows

the individual cases and group centroids (average values for each group) displayed in two dimensions: (a) from left to right, Function 1 distinguishes between Group Native and the other two groups; (b) from top to bottom, Function 2 additionally distinguishes Groups High from the other two groups.

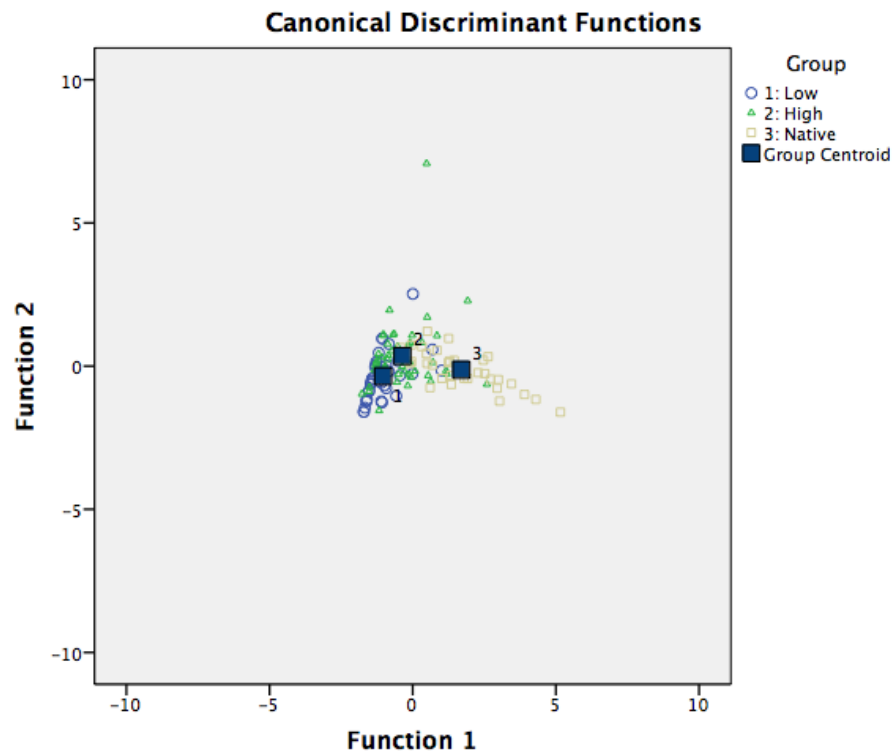


Figure 11. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the V1 task.

For the V1 task, the classification results indicated that overall, 87 (or 75.7%) were correctly classified as shown in Table 18. However, the accuracy of the classifications varied for the three levels. Group Low participants were classified with 81.1% accuracy while Group High participants were classified with 72.3% accuracy, and Group Native participants were classified correctly with 74.2% accuracy.

Table 18

Classification results for the V1 task

| Actual Group | <i>N</i> | Predicted Group Membership | | |
|---------------------|----------|-----------------------------------|--------------------|--------------------|
| | | Low | High | Native |
| | | (75.7% correctly predicted) | | |
| Low | 37 | 30 81.1% | 7 18.9% | 0 .0% |
| High | 47 | 6 12.8% | 34 72.3% | 7 14.9% |
| Native | 31 | 0 .0% | 8 25.8% | 23 74.2% |

Note. The bold are the correct predictions.

For the V2 task, both the combined functions (1 and 2) and the second function alone showed statistically significant effects in distinguishing proficiency levels between the three groups. Function 1 accounted for the large majority (94.8%) of observable between-groups variance across the three groups, and Function 2 accounted for 5.2%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .319, $\chi^2(8, N=115) = 126.222, p = .000$, indicating that the combined predictor variables were able to account for around 68% of the actual variance in proficiency level between the three groups. On its own, the second function also provided additional statistically significant predictions, Wilks' lambda = .907, $\chi^2(3, N=115) = 10.729, p = .013$. Figure 12 shows the individual cases and group centroids (average

values for each group) displayed in two dimensions: (a) from left to right, Function 1 distinguishes between all three groups with more distinction occurring between Group Native and the other two groups; (b) from top to bottom, Function 2 additionally distinguishes Group High from the other two groups.

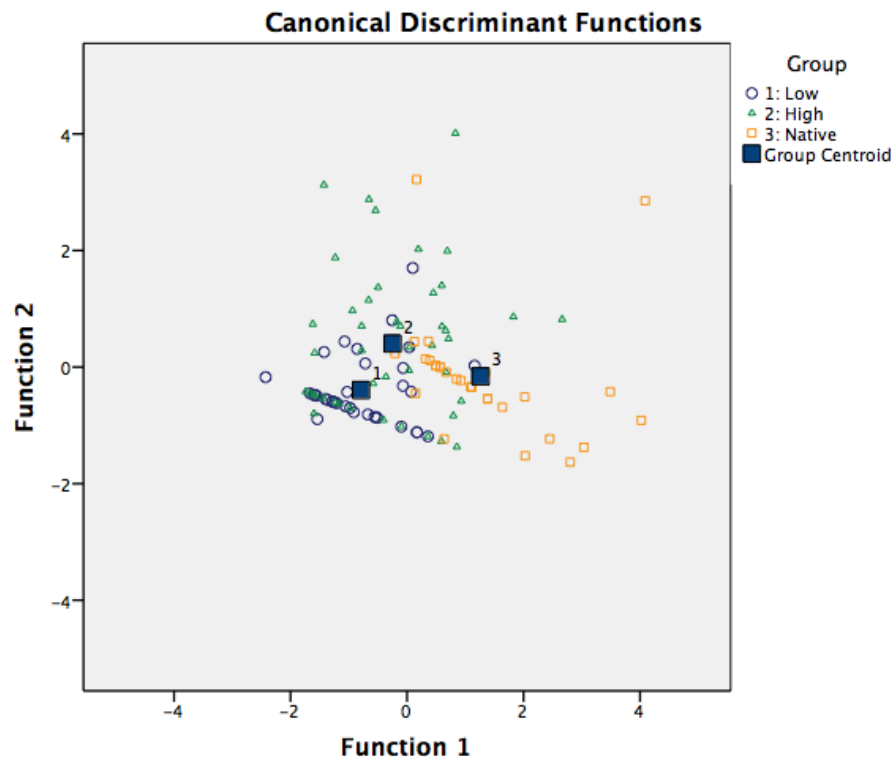


Figure 12. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the V2 task.

The classification results indicated that, overall, 88 (or 76.5%) were correctly classified as shown in Table 19. However, the accuracy of the classifications varied for the three levels. Group Low participants were classified with 81.1% accuracy, while Group High participants were classified with 78.7% accuracy, and Group Native participants were classified correctly with 67.7% accuracy.

Table 19

Classification results for the V2 task

| Actual Group | <i>N</i> | Predicted Group Membership | | |
|---------------------|----------|-----------------------------------|---------------------------|---------------------------|
| | | (76.5% correctly predicted) | | |
| | | Low | High | Native |
| Low | 37 | 30 81.1% | 7 18.9% | 0 .0% |
| High | 47 | 7 14.9% | 37 78.7% | 3 6.4% |
| Native | 31 | 0 .0% | 10 32.3% | 21 67.7% |

Note. The bold are the correct predictions.

Applying the four Chinese syntactic complexity measures, the discriminant analyses on V1 (75.7%) and V2 (76.5%) generated much higher accuracy of proficiency membership prediction than on CS (67.8%). Since CS was designed with less cognitive task complexity along the line of resource-directing than V1 and V2, these three tasks seemed to have played different roles at eliciting Chinese syntactic complexity factors. With the cognitive task complexity manipulated along the line of resource-directing, more complex task generated output for more accurate proficiency group membership prediction, whereas less complex task generated output for less accurate proficiency group membership prediction. This pattern required further consideration of the

correlations between the four predictor variables and each discriminant function in order to see where such differences were located.

As mentioned above, for the three speaking tasks, CS, V1, and V2, Function 1 accounted for 96.2% ($p = .000$), 90.4% ($p = .000$), and 94.8% ($p = .000$) of the between-groups variance across the three proficiency groups in discriminating on the three tasks respectively. In addition, for all of these speaking tasks, as shown in Table 20 below, the predictor variable ① MLTTCU showed the absolutely highest correlation, $r = .96$, $.99$, and $.96$, with the first function. This indicated that in all three discriminant analyses for the three tasks, each Function 1 was best represented by the global complexity measure ① MLTTCU. Function 2 accounted for 3.8% ($p = .139$), 9.6% ($p = .000$), 5.2% ($p = .013$) for the three speaking tasks, CS, V1, and V2. Function 2, by contrast, was best represented by varied measures in different tasks. In the CS task, Function 2 was best represented by the clausal complexity measure ③ MLSTCU, which correlated moderately ($r = .34$) with the function. In addition, note that Function 2 correlated negatively with both measure ② CTTCU/ATTCU ($r = -.12$) and ④ STCU/TTCU ($r = -.12$). In the V1 task, Function 2 was best represented by measure ② CTTCU/ATTCU and measure ④ STCU/TTCU, which each correlated moderately ($r = .40$, $.37$) with the function. In the V2 task, the correlations between Function 2 and all the measures were weak, with measure ③ MLSTCU showing the highest correlation ($r = .14$) followed by ② CTTCU/ATTCU ($r = .12$) and ④ STCU/TTCU ($r = .05$). In both task V1 and V2 which generated higher accuracy of classification in the discriminant analyses, and in each task measure ① MLTTCU correlated negatively with Function 2 in each task.

Such different correlations between the measures and Function 2 of the three tasks might point to the likelihood that the different task designs elicited quite different language performances. The higher correlation of scores on the ratio measures with Function 2 might have contributed to the increase of classification accuracy in the V1 (75.7% classification accuracy) and V2 task (76.5% classification accuracy) compared with the CS task (67.8% classification accuracy). Across the three tasks of varied cognitive task complexity, the ratio measures of ② CTTCU/ATTCU and ④ STCU/TTCU both correlated negatively with Function 2 in the CS task, but showed a relatively higher correlation out of the four measures in the V1 and V2 tasks. This might be suggesting that the CS task when compared with the V1 and V2 tasks, did not elicit a good amount of complex terminable TC-units consisting of more dependent single TC-units. By contrast, when completing the V1 and V2 tasks of higher cognitive task complexity, a higher number of complex terminable TC-units as well as more dependent single TC-units in one complex terminable TC-unit were elicited. Therefore, across the three tasks of varied cognitive task complexity, in addition to eliciting syntactic complexity on the length measure of MLTTCU, tasks of higher cognitive complexity compared to tasks of lower cognitive complexity seemed to be able to elicit more complex terminable TC-units as well as more dependent single TC-units.

In spite of the varied correlations between the four measures and Function 2 across the three tasks, the absolute highest correlations were shown between measure ① MLTTCU and Function 1 for the three speaking tasks, .96 in the CS task, .99 in the V1 task, and .96 in the V2 task. Therefore, ① MLTTCU was chosen to be applied as the

singular predictor in subsequent discriminant analyses, and it generated approximately similar classification results as applying all the four predictors together. The reliability of the discriminant function for each of the three tasks was also found to be statistically

Table 20

Two functions of each discriminant analysis for speaking tasks

| Correlations of Predictor Variables with | | | | | | |
|---|-----------|----------|-----------|----------|-----------|----------|
| Discriminant Functions | | | | | | |
| Predictor | CS | | V1 | | V2 | |
| | 1 | 2 | 1 | 2 | 1 | 2 |
| Variable | | | | | | |
| ① MLTTCU | .96 | .01 | .99 | -.01 | .96 | -.08 |
| ② CTTCU/ATTCU | .70 | -.12 | .64 | .40 | .43 | .12 |
| ③ MLSTCU | .60 | .34 | .67 | -.11 | .76 | .14 |
| ④ STCU/TTCU | .71 | -.12 | .58 | .37 | .42 | .05 |
| Canonical <i>R</i> | .75 | .22 | .82 | .42 | .81 | .30 |
| Eigen value | 1.27 | .05 | 2.05 | .22 | 1.85 | .10 |

significant when ① MLTTCU was applied as the only predictor variable. The mean and standard deviation of each group for each task are listed in Table 21 below. For the CS task, Wilks' lambda = .459, $\chi^2(2, N=115) = 87.312, p = .000$, and a total of 78 (or 67.8%) of the cases were correctly classified. For the V1 task, Wilks' lambda = .332, $\chi^2(2, N=115) = 123.369, p = .000$, and a total of 85 (or 73.9%) of the cases were correctly

classified. For the V2 task, Wilks' lambda = .370, $\chi^2(2, N=115) = 111.409, p = .000$, and a total of 89 (or 77.4%) of the cases were correctly classified. To sum up, for the CS, V1, and V2 tasks respectively, accuracy in case classification were 67.8%, 73.9%, and 77.4% when applying ① MLTTCU as the singular predictor, and 67.8%, 75.7%, and 76.5% when utilizing all four measures as predictors. Therefore, the classification accuracy were approximate between the results generated by utilizing ① MLTTCU only or all four measures. Such approximate accuracy suggested that the measure ① MLTTCU by itself may be chosen as the most effective indicator of spoken Chinese syntactic complexity.

Table 21

Classification results for the three speaking tasks with MLTTCU as the only predictor

| | | CS | | | V1 | | | V2 | | |
|--------------|----|----------------------------|--------------|--------------|----------------------------|--------------|--------------|----------------------------|--------------|--------------|
| Actual Group | N | Predicted Group Membership | | | Predicted Group Membership | | | Predicted Group Membership | | |
| | | (67.8%) | | | (73.9%) | | | (77.4%) | | |
| | | Low | High | Native | Low | High | Native | Low | High | Native |
| Low | 37 | 27 | 9 | 1 | 31 | 6 | 0 | 31 | 6 | 0 |
| | | 73.0% | 24.3% | 2.7% | 83.8% | 16.2% | .0% | 83.8% | 16.2% | .0% |
| High | 47 | 9 | 31 | 7 | 8 | 31 | 8 | 8 | 36 | 3 |
| | | 19.1% | 66.0% | 14.9% | 17.0% | 66.0% | 17.0% | 17.0% | 76.6% | 6.4% |
| Native | 31 | 1 | 10 | 20 | 0 | 8 | 23 | 0 | 9 | 22 |
| | | 3.2% | 32.3% | 64.5% | .0% | 25.8% | 74.2% | .0% | 29.0% | 71.0% |

5.2.3.2 Writing tasks

The FW task and GW task were completed in the written modality. For the FW task, Function 1 accounted for the large majority (92.7%) of the observable between-groups variance in discriminating among the three groups, and Function 2 accounted for 7.3%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .412, $\chi^2(8, N=116) = 98.848, p = .000$, indicating that the combined predictor variables were able to account for around 59% of the actual variance across the three proficiency groups. On its own, the second function provided additional statistically significant predictions, Wilks' lambda = .912, $\chi^2(3, N=116) = 10.239, p = .013$. Figure 13 shows the individual cases and group centroids (average values for each group) displayed in two dimensions: (a) from left to right, Function 1 clearly

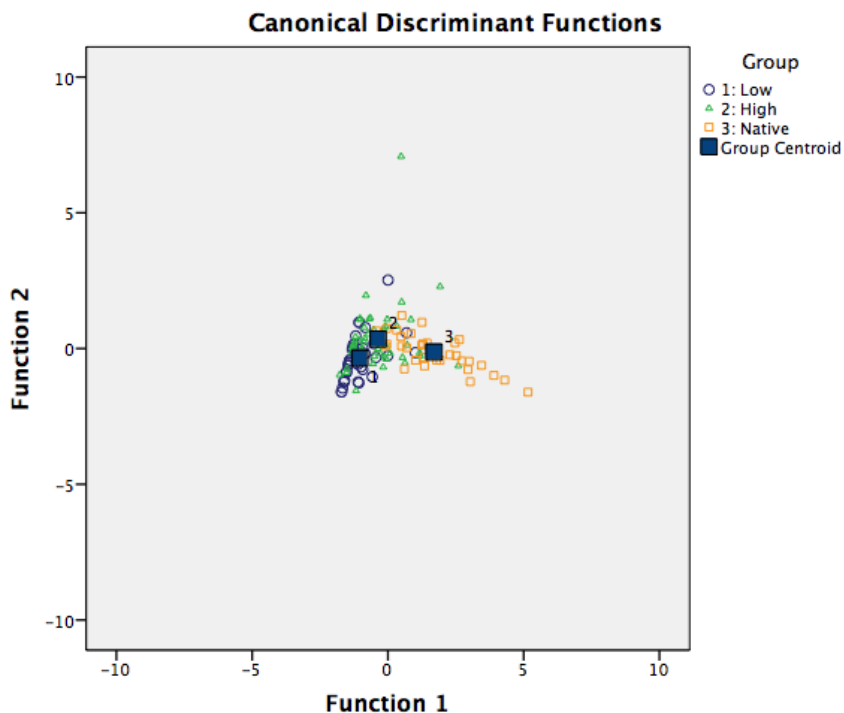


Figure 13. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the FW task.

distinguishes between Group Native and the other two groups; (b) from top to bottom, Function 2 additionally distinguishes Groups High from the other two groups.

For the FW task, the classification procedure indicated that, overall, 78 (or 67.2%) cases were correctly classified as shown in Table 22. As for each group, Group Low participants were classified with 61.1% accuracy, while Group High participants were classified with 72.9% accuracy, and Group Native participants were classified correctly with 65.6% accuracy.

Table 22

Classification results for the FW task

| Actual Group | <i>N</i> | Predicted Group Membership | | |
|---------------------|----------|-----------------------------------|---------------------------|---------------------------|
| | | Low | High | Native |
| Low | 36 | 22 61.1% | 13 36.1% | 1 2.8% |
| High | 48 | 9 18.8% | 35 72.9% | 4 8.3% |
| Native | 32 | 0 .0% | 11 34.4% | 21 65.6% |

Note. The bold are the correct predictions.

For the GR task, Function 1 accounted for the majority (84.3%) of the observable between-groups variance in discriminating among the three groups, and Function 2 accounted for 15.7%. An overall statistically significant effect was found for the

combined functions (1 and 2), Wilks' lambda = .530, $\chi^2(8, N=116) = 70.815, p = .000$, indicating that the combined predictor variables were able to account for around 47% of the actual variance in proficiency level between the three groups. On its own, the second function provided additional statistically significant predictions, Wilks' lambda = .888, $\chi^2(3, N=116) = 13.191, p = .004$. Figure 14 shows the individual cases and group centroids (average values for each group) displayed in two dimensions: (a) from left to right, Function 1 distinguishes between Group Native and the other two; (b) from top to bottom, Function 2 additionally distinguishes between the three groups.

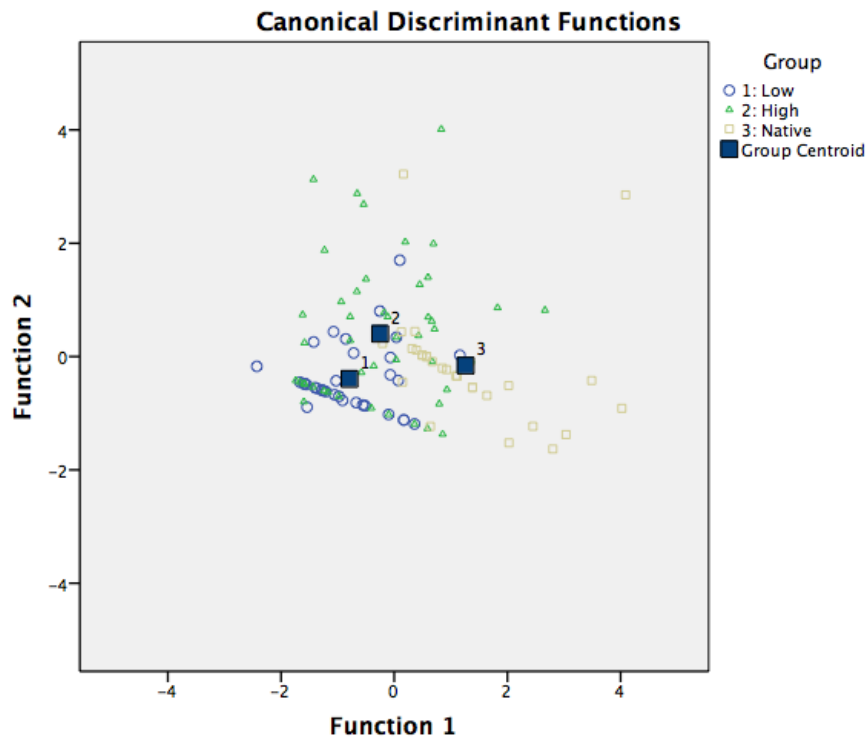


Figure 14. Predicting proficiency groups by four measures: Cases and group centroids for two discriminant functions of the GR task.

The classification procedure indicated that, overall, 71 (or 61.2%) cases were correctly classified in the GR task as shown in Table 23. However, the accuracy of the

classifications varied for the three levels. Group Low participants were classified with 69.4% accuracy, while Group High participants were classified with 56.3% accuracy, and Group Native participants were classified correctly with 59.4% accuracy.

Table 23

Classification results for the GR task

| Actual Group | | Predicted Group Membership | | |
|---------------------|----------|-----------------------------------|--------------|---------------|
| | | Low | High | Native |
| | <i>N</i> | (61.2% correctly predicted) | | |
| | | | | |
| Low | 36 | 25 | 10 | 1 |
| | | 69.4% | 27.8% | 2.8% |
| High | 48 | 14 | 27 | 7 |
| | | 29.2% | 56.4% | 14.6% |
| Native | 32 | 0 | 13 | 19 |
| | | .0% | 40.6% | 59.4% |

Note. The bold are the correct predictions.

Comparing the correlations between predictor variables and the two functions in the two writing tasks, somewhat different patterns were found from the correlations in the three speaking tasks. As shown in Table 24, for the FW task it was not the measure ① MLTTCU but the predictor variable ④ STCU/TTCU that showed the highest correlation, .95, with Function 1. In the FW task, this ratio measure worked better than other measures in differentiating the observable between-groups variance across the three groups. Another ratio measure ② CTTCU/ATTCU also demonstrated a high

correlation, .87, with Function 1 in the FW task. This might be pointing to a possibility that the composition of dependent single TC-units played a more important role in contributing to the written Chinese syntactic complexity than to the spoken Chinese syntactic complexity. In other words, the FW task might have elicited more dependent single TC-units to compose each complex terminable TC-unit as well as a higher percentage of such complex terminable TC-units out of all terminable TC-units. However, applying measure ④ STCU/TTCU as the singular predictor variable for the FW task generated only 56.9% correct group membership prediction. The length measure ① MLTTCU also showed a quite strong correlation, $r = .76$, with Function 2. Function 2 was best represented by the predictor variable ③ MLSTCU, with a correlation $r = .96$.

Table 24

Two functions of each discriminant analysis for the writing tasks

| Correlations of Predictor Variables | | | | |
|--|------------|------|------------|------|
| with Discriminant Functions | | | | |
| Predictor Variable | FW | | GR | |
| | 1 | 2 | 1 | 2 |
| ① MLTTCU | .72 | .55 | .93 | .17 |
| ② CTTCU/ATTCU | .87 | -.07 | .47 | .65 |
| ③ MLSTCU | .05 | .96 | .70 | -.52 |
| ④ STCU/TTCU | .95 | -.23 | .48 | .73 |
| Canonical <i>R</i> | .74 | .30 | .64 | .33 |
| Eigen value | 1.21 | .10 | .68 | .13 |

For the GR task, measure ① MLTTCU showed the highest correlation, .93 with Function 1. Function 2 by contrast, was best represented by predictor variable measure ④ STCU/TTCU, with a correlation of $r = .73$. Applying the length measure ① MLTTCU as the singular predictor variable for the GR task generated a lower accuracy rate of 53.4% at group membership prediction, not very different from chance.

Applying both the length measure ① MLTTCU and the ratio measure ④ STCU/TTCU as predictors through SPSS CLASSIFY generated approximately similar accuracy of group membership prediction for the two writing tasks as when all the four measures were applied. For the FW task, applying both measure ① MLTTCU and measure ④ STCU/TTCU as the predictor variables showed statistically significant reliability of both two discriminant functions. Function 1 accounted for the large majority (93.0%) of the observable between-groups variance across the three groups, and Function 2 accounting for 7.0%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .434, $\chi^2(4, N=116) = 93.933, p = .000$, indicating that the combined predictor variables were able to account for around 57% of the actual variance in proficiency level between the three groups. On its own, Function 2 provided additional statistically significant predictions, Wilks' lambda = .922, $\chi^2(1, N=116) = 9.170, p = .002$. Figure 15 shows the individual cases and group centroids (average values for each group) displayed in two dimensions: (a) from left to right, Function 1 clearly distinguishes between Group Native and the other two; (b) from top to bottom, Function 2 additionally distinguishes Group High from the other two. The classification procedure indicated that, overall, 79 (or 68.1%) cases were correctly classified, as shown in Table 25. At the three group membership levels, Group Low

participants were classified with 61.1% accuracy, while Group High participants were classified with 72.9% accuracy, and Group Native participants were classified with 68.8% accuracy.

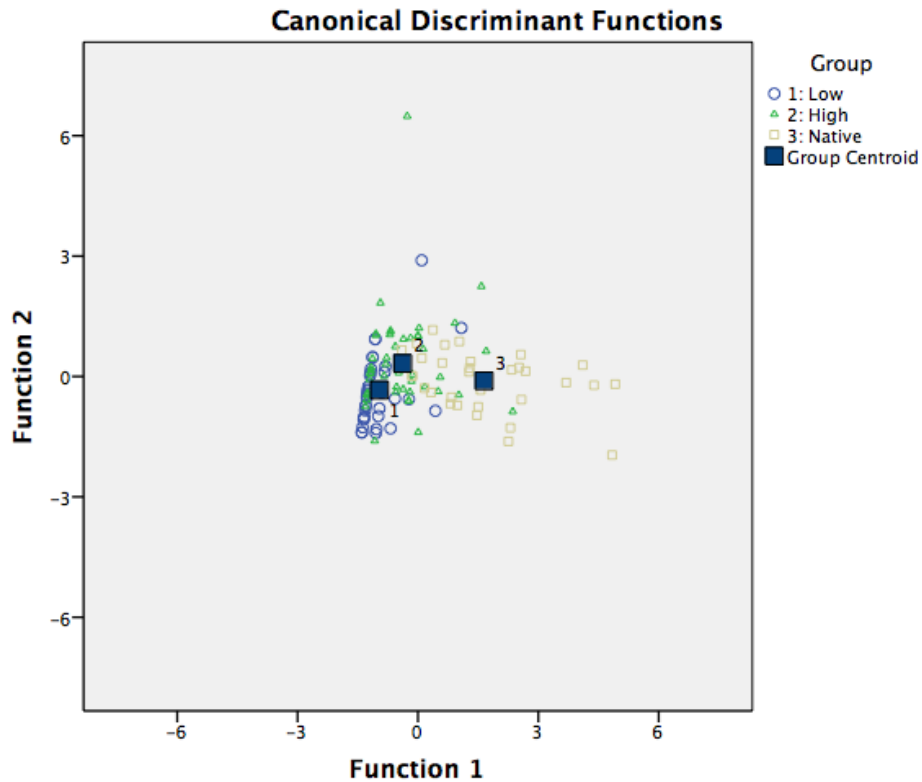


Figure 15. Predicting proficiency groups by measure MLTTCU and STCU/TTCU: Cases and group centroids for two discriminant functions of the FW task.

For the GR task, applying measure ① MLTTCU and measure ④ STCU/TTCU as the two predictor variables showed statistically significant reliability of both discriminant functions. Function 1 accounted for the majority (86.4%) of the observable between-groups variance across the three proficiency groups, and Function 2 accounting for 13.6%. An overall statistically significant effect was found for the combined functions (1 and 2), Wilks' lambda = .554, $\chi^2(4, N=116) = 66.354, p = .000$, indicating that the

combined predictor variables were able to account for around 45% of the actual variance in proficiency level between the three groups. On its own, the second function provided additional statistically significant predictions, Wilks' lambda = .908, $\chi^2(1, N=116) = 10.804, p = .001$. Figure 16 shows the individual cases and group centroids (average values for each group) displayed in two dimensions: (a) from left to right, Function 1 clearly distinguishes between Group Native and the other two; (b) from top to bottom, Function 2 additionally distinguishes between the three groups. The classification procedure indicated that, overall, 73 (or 62.9%) cases were correctly classified, as shown in Table 25. At the three group membership levels, Group Low participants were classified with 72.2% accuracy, while Group High participants were classified with

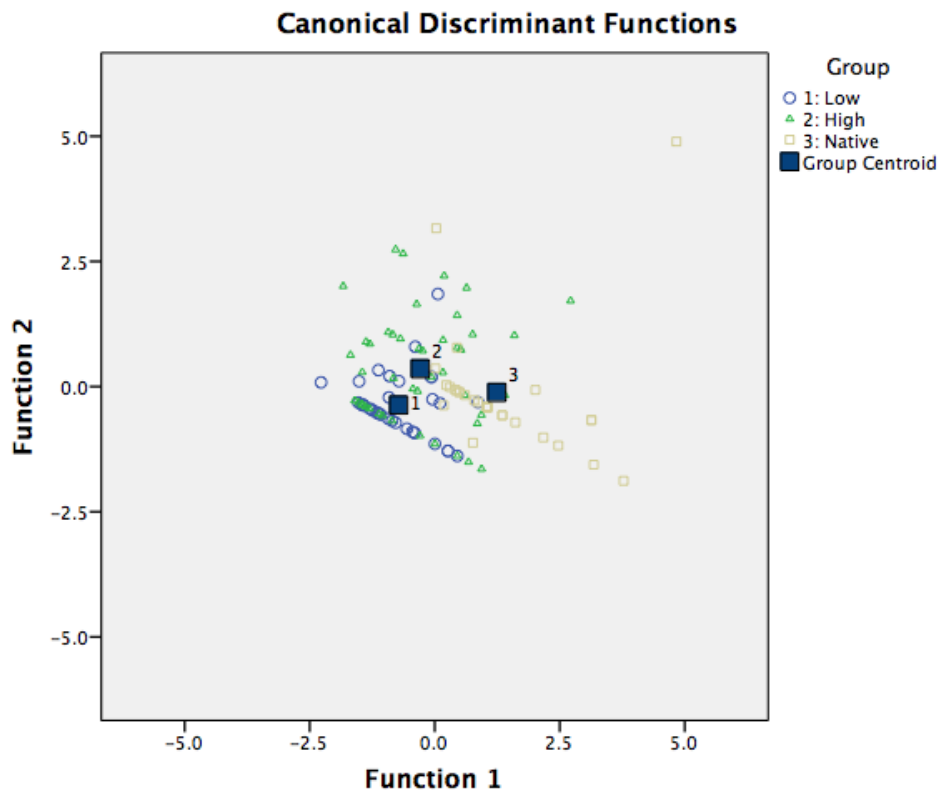


Figure 16. Predicting proficiency groups by measure MLTTCU and STCU/TTCU: Cases and group centroids for two discriminant functions of the GR task.

Table 25

Classification results for the writing tasks with MLTTCU and STCU/TTCU as the predictors

| | | FW | | | GR | | |
|--------------|----|---|--------------|--------------|---|--------------|--------------|
| Actual Group | N | Predicted Group Membership (68.1% correctly predicted) | | | Predicted Group Membership (62.9% correctly predicted) | | |
| | | Low | High | Native | Low | High | Native |
| Low | 36 | 22 | 13 | 1 | 26 | 9 | 1 |
| | | 61.1% | 36.1% | 2.8% | 72.2% | 25.0% | 2.8% |
| High | 48 | 9 | 35 | 4 | 13 | 28 | 7 |
| | | 18.8% | 72.9% | 8.3% | 27.1% | 58.3% | 14.6% |
| Native | 32 | 0 | 10 | 22 | 0 | 13 | 19 |
| | | .0% | 31.3% | 68.8% | .0% | 40.6% | 59.4% |

Note. The bold are the correct predictions.

58.3% accuracy, and Group Native participants were classified correctly with 59.4% accuracy.

As shown in Table 26 below, for the FW task, measure ④ STCU/TTCU still showed the highest correlation, $r = .93$, with Function 1. Function 2 was best represented by the predictor variable ① MLTTCU with a correlation of $r = .67$. For the GR task, Function 1 was best represented by the predictor variable ① MLTTCU with a strong correlation of $r = .95$. Function 2 was best represented by measure ④ STCU/TTCU with

another strong correlation of $r = .88$. For the GR task, Function 2 accounted for 15.7% when all measures were applied as predictor variables, and 13.6% when measures ① MLTTCU and ④ STCU/TTCU were applied as the two predictor variables for participants' global proficiency level. This is great accountability by Function 2 alone. While Function 2 was best represented by measure ④ STCU/TTCU having a strong correlation of $r = .73$ and $.88$ respectively. This indicated that the ratio of single TC-units per terminable TC-unit contributed greatly to the syntactic complexity in the GR task. If participants composed complex terminable TC-unit consisting of more dependent single TC-units, to some extent this quite effectively distinguished their global language proficiency level.

Table 26

Two functions of each discriminant analysis for the writing tasks

| Predictor Variable | Correlations of Predictor Variables with Discriminant Functions (FW) | | | Correlations of Predictor Variables with Discriminant Functions (GR) | | |
|--------------------|--|------|-----------------------|--|-----|-----------------------|
| | 1 | 2 | Univariate $F(2,113)$ | 1 | 2 | Univariate $F(2,113)$ |
| | ① MLTTCU | .75 | .67 | 37.41 | .95 | .31 |
| ④ STCU/TTCU | .90 | -.14 | 62.34 | .48 | .88 | 12.54 |
| Canonical R | .73 | .35 | | .62 | .30 | |
| Eigen value | 1.16 | .14 | | .64 | .10 | |

When comparing the accuracy of group membership predictions for the writing tasks and speaking tasks by all four measures, predictions of 67.2% (FW) and 61.2% (GR) were relatively lower than 67.8% (CS), 75.7% (V1), and 76.5% (V2). A closer look was then taken at Table 22 and Table 23 for the misclassification of each proficiency group separately. For the FW task, the highest misclassification occurred in Group Low of which 13 (36.1%) cases were misclassified into Group High, followed by 11 (34.4%) cases in Group Native being misclassified into Group High. Figure 13 displayed quite an overlap among the three proficiency groups, especially between Group Low and Group High. This pattern suggested that the scores of three groups were not varied enough to better distinguish each from the other. Some Group High participants did not sufficiently outperform the Group Low participants on the four complexity measures in the FW task. There were some Group Native participants who did not quite outperform Group High on the four complexity measures in the FW task either. For the GR task, the highest misclassification fell in Group Native of which 13 (40.6%) cases were misclassified into Group High. In addition, and 14 (29.2%) Group High cases were misclassified into Group Low, and 10 (27.8%) Group Low cases were misclassified into Group High. As the cases of the three groups for the GR task displayed in Figure 14 graphically indicate, the scores of each case were quite coarsely distributed, especially for the Group High and Group Native participants. This pattern suggests that on the four complexity measures the performance of all the participants on the GR task showed more variations at the individual level than on the group level.

Several possible interpretations might account for such accuracy differences when applying the same four measures at classifying spoken and written Chinese syntactic complexity.

On the one hand, a possible explanation could be the varying developmental trajectory between spoken and written Chinese complexity in terms of composition versus lengthening of the single TC-units. As discussed in Section 5.2.2 Correlational Analysis, in this dissertation the length measures better represented the discriminant functions for spoken Chinese complexity, whereas the ratio measures showed higher efficiency at classifying written Chinese complexity. It might be that spoken Chinese complexity showed more salient development alongside the lengthening of terminable TC-units and single TC-units. While for the written Chinese complexity, in addition to the lengthening of TC-units, more salient development was shown alongside composing more single TC-units into a complex terminable TC-unit as well as composing more such complex terminable TC-units out of all the terminable TC-units.

As previously shown in Figure 9 (See Section 5.2.1), the distribution of the mean score on each complexity measure by task also confirmed such difference between spoken and written Chinese output. For the spoken Chinese output elicited, the mean scores on all four measures across proficiency groups were quite evenly spread. For the two writing tasks, the distributions of both showed some overlap on measure ③ MLSTCU among proficiency groups. For the FW task output, Group High scored 11.69 characters which surpassed the mean score of 10.74 characters for Group Native. The two ratio measures ② CTTCU/ATTCU and ④ STCU/TTCU, by contrast, showed better distinction across the three proficiency groups, especially between Group Native and

Group High in the FW task output. The FW task showed a higher score on measure ② CTTCU/ATTCU and ④ STCU/TTCU than with other tasks. Participants on the FW task seemed to have produced a bigger amount of complex terminable TC-units as well as more dependent single TC-units per terminable TC-unit than on any of the speaking tasks. In addition, the gap between the mean scores of Group Native and Group High on the two ratio measures in the FW task were both relatively larger than the score gaps in the speaking task output, while measure ② CTTCU/ATTCU and ④ STCU/TTCU showed better distinctions between Group Native and Group High for the FW task output. In the GR task output, there also was an overlap between Group Low, 12.23 characters, and Group High, 12.16 characters on the length measure ③ MLSTCU. However, the score distribution on the two ratio measure were more dispersed.

On the other hand, it could be that the task design in this dissertation functioned differentially in eliciting Chinese spoken and written syntactic complexity.

There was good predictability on the speaking tasks, and in particular on the two video retelling tasks: V1 and V2. In all of the three speaking tasks, no cases in Group Low were misplaced into Group Native, nor vice versa. As shown in Table 17, for the CS task, the highest misclassifications were placed in Group Native and Group Low when all the four measures were applied. 12 (38.7%) Group Native cases were misplaced into Group High, and 10 (27.0%) Group Low cases were also misplaced into Group High. It suggested that the dispersion between the scores of three proficiency groups, especially between Group High and the other two, could be more stretched to better separate the three groups. Such dispersion between Group High and the other two groups was much more stretched in the V1 and V2 tasks of higher cognitive task complexity along the line

of resource-directing. As also shown in Figure 11 and Figure 12, there was less overlap of the scores in the V1 and V2 task between the Group High and the other two. In the task V1, 8 (25.8%) Group Native cases were misplaced into Group High, and 7 (18.9%) Group Low cases were misplaced into Group High. In the V2 task, 7 (18.9%) Group Low cases were misplaced into Group High, and 10 (32.3%) Group Native cases were misplaced into Group High. Along the line of resource-directing, speaking tasks of higher cognitive complexity as compared with tasks of lower cognitive complexity have elicited language output for better group membership classification on the four complexity measures.

One likely cause for the relatively lower accurate group membership prediction for the writing tasks than with the speaking task was the allocated time limit. Different from the spontaneity of speaking output, for writing output people had the chance to review what they were writing, weigh their words, and revise the structures. All of this required more contemplation. While 1.5 – 3 minutes might have been sufficient to produce a great range of language complexity in the form of spoken output, 7 minutes for the FW task and 5 minutes for the GR task in the TW&ST could be extended in order to elicit written output of higher syntactic complexity from the participants of higher proficiency levels, especially the native speaker participants. With more time allocated the writing tasks might be able to spread out the scores on syntactic complexity measures for written Chinese output of the three proficiency groups.

In addition, for the FW task a daily life topic such as “my relationship with my father / mother / brother / sister / friend (choose any one of them)” was selected in order to minimize the floor effect. However, this might have also indulged the written output of

the higher proficiency group to a threshold complexity level. Words and syntactic structures of lower complexity were able to express such a daily life topic, which might have kept the higher proficiency participants from using words and structures of higher complexity. In other words, a topic such as “Confucius and Immanuel Kant: a comparison and contrast of their philosophy” targeting at the higher proficiency level might induce a much higher level of language complexity level. However, this would not be without the potential drawback of having lower level participants not being able to produce any written output for such an advanced topic.

For the relatively coarsely distributed scores in the GR task, there could have additionally existed other possible causes. Instead of writing from scratch, a certain baseline level of syntactic complexity (especially the length of single TC-units), seemed to be achievable for all the participants by starting with the provided seven semantically coherent yet formally incohesive sentences. More justification and revision of the provided sentences, especially composing more independent single TC-units into a complex terminable TC-unit, would be required in order to achieve higher complexity. This need for composing more single TC-units into one complex terminable TC-unit in order to differentiate proficiency level on the GR task was in line with the aforementioned result – that Function 2 on the GR task contributed quite some accountability and was best represented by measure ④ STCU/TTCU. If participants composed complex terminable TC-unit consisting of more dependent single TC-units, this then to some extent quite effectively distinguished their global language proficiency level. Another possible cause for the varied performance for Group Native participants might be due to the lack of such test practice. As native speakers, Group Native

participants were much less experienced with such task types when compared with the L2 participants who usually experience such exercises in their language classes. Therefore, the test scores of the Group Native participants might have underestimated their actual Chinese syntactic complexity level. Task priming would have also helped the native speaker participants to become more familiar with the guided rewriting task.

5.2.3.3 Interim summary

As clarified in Section 2.1.3, this dissertation adopted the definition of complexity from Bulté and Housen (2012): “at the most basic level, complexity refers to a property or quality of a phenomenon or entity in terms of (a) the number and the nature of the discrete components that the entity consists of, and (b) the number and the nature of the relationships between the constituent components” (p. 22). As we have seen through discriminant function analyses, applying the proposed measures of Chinese syntactic complexity as predictors of Chinese language proficiency, the TC-unit based measures proved high efficiency (61.2%~76.5%) at proficiency group membership classification. Therefore, the TC-unit seems to be proved across tasks an appropriate unit of analysis for Chinese syntactic complexity. As the scores on the length and ratio measures shown, both length and internal compositionality of the terminable TC-units can reveal the syntactic complexity of the Chinese language. Generally, with higher Chinese language proficiency, longer terminable TC-units consisting of more dependent single TC-units can be produced. By contrast, with lower language proficiency, shorter complex terminable TC-units consisting of less dependent single TC-units or shorter simple terminable TC-units consisting of singular independent single TC-units were produced. A

complex terminable TC-unit is composed by dependent TC-units via coreferential zero. Therefore, proactive use of coreferential zero also shows higher Chinese syntactic complexity. However, the development of syntactic complexity across global proficiency levels varied between spoken and written Chinese. For spoken Chinese syntactic complexity, the length of the terminable TC-units seemed to be the most salient feature distinguishable among Chinese global proficiency levels. The written Chinese syntactic complexity additionally showed more combining of single TC-units in order to form complex terminable TC-units. In addition to lengthening the terminable TC-units as a result of proficiency level increases, a greater proportion of complex terminable TC-units were discovered within all of the terminable TC-units for written Chinese.

5.2.4 Chinese syntactic complexity and cognitive task complexity

5.2.4.1 Higher language complexity produced in more complex tasks

There were two dimensions applied in varying the cognitive complexity of the three speaking tasks. The complexity variance between the cartoon strip task and the video retelling tasks was along the resource-directing dimension. In contrast, the cognitive complexity difference of the two video retelling tasks varied along the resource-dispersing dimension.

A repeated measures analysis showed that higher language complexity was produced in the tasks of higher cognitive complexity along the resource-directing dimension in this dissertation. For the output of three speaking tasks of different cognitive complexity, CS, V1, and V2, a repeated measures analysis was conducted. As listed in Table 11 (in Section 4.3.4), the amount of complete spoken data sets consisting

of all these three speaking tasks was $N = 115$. The amount of valid data sets for each proficiency group were Group Low: $n = 37$, Group High: $n = 47$, and Group Native: $n = 31$. The difference of the cognitive task complexity among the three speaking tasks was therefore taken as the “within-subjects” variable. The four complexity measures, ① MLTTCU, ② CTTCU/ATTTCU, ③ MLSTCU, and ④ STCU/TTCU were the dependent variables. The proficiency group was the “between-subjects” variable. A clear effect for “task” ($F(8, 105) = 7.443, p = .000, \eta^2 = .362$), and a clear effect for proficiency level ($F(8, 218) = 29.704, p = .000, \eta^2 = .522$) were found. This finding showed that the cognitive complexity of the task as well as the proficiency level of the participants did have a statistically significant effect on the Chinese syntactic complexity of their output in general. As analyzed in Section 4.2.2, the CS task was less cognitively complex than the two video retelling tasks in terms of fewer elements involved in the story, shorter time of performance required, and much less cognitive reasoning effort demanded. In the video retelling tasks that were more cognitively complex along the resource-directing line, syntactic complexity of the Chinese output was higher (on all measures) than the output of the CS task that was less cognitively complex. There was also likely an interaction effect between proficiency and task ($F(16, 210) = 2.567, p = .001, \eta^2 = .164$) suggesting that cognitive task complexity played different roles across different Chinese proficiency groups in affecting Chinese syntactic complexity. On the one hand, the syntactic complexity development across proficiency levels might be nonlinear or uneven at different complexity levels. On the other hand, tasks of varied cognitive complexity design can function differently with regards to eliciting Chinese syntactic complexity

from different levels. The score distance on Chinese syntactic complexity measures between the three tasks varied for participants of different proficiency levels.

In the posthoc analyses graphically displayed in Figure 17 below, on all the four complexity measures participants from each proficiency group scored much lower in the CS task than they did in the V1 and V2 tasks, but tasks V1 and V2 were similar on all four measures. It was clear that the differences all occurred between the CS task and the two video retelling tasks (V1 and V2) on all four measures. There was no statistically significant difference found for any of the measures between task V1 and V2 ($F(4, 104) = 1.953, p = .107, \eta^2 = .070$). Nevertheless, there were still clear differences between proficiency groups ($F(8, 208) = 27.379, p = .000, \eta^2 = .513$). No interaction effects between proficiency level and task ($F(8, 208) = 1.308, p = .241, \eta^2 = .048$) were found between the V1 and V2 task.

Since V2 is less cognitively complex than V1 along the resource-dispersing dimension, it can be concluded that in this dissertation syntactic complexity of the Chinese output produce in the higher cognitively complex task V1 was not higher than those produced in the less cognitively complex task V2. Generally, along the resource-dispersing line, in terms of planning time and prior knowledge, V2, as an immediate repetition of task V1, was less cognitively complex than V1. As we discussed in Section 2.3.3, immediate task repetition was introduced as in this dissertation to operationalize the task planning variable along the resource-dispersing dimension. As for prior knowledge, in the retrospective survey (See Appendix B and C) that participants completed upon after both V1 and V2, Question 14 asked participants if they “were familiar with this story about *Nián* before watching the video”. Only 13 (14.3%)

participants reported “False” to this question, since this *Nian* story is a widely disseminated and well-known legend. However, because of the great popularity of this legend, there are a good number of varied versions with different details of the plot. It was not until participants finished watching the video that they knew the exact story to retell. Prior knowledge of the story content therefore did not seem to exert a meaningful difference on the language complexity of their performances.

However, the score difference between V1 and V2 on each measure seemed to vary by proficiency group. As shown in Figure 17, on all four measures V2 outperformed V1 with a slight advantage overall. On measure ① MLTTCU, V1 and V2 were almost identical on the four measures with Group Low showing a little improvement over the other two groups in V2. On measures ② CTTCU/ATTCU and ④ STCU/TTCU, Group Low showed similarly low scores in CS and V1 but more improvement in V2, whereas the other two groups showed almost identical scores for V1 and V2 tasks. For the V1 and V2 task designed with cognitive complexity differences along the resource-directing line, the Chinese output complexity did not show too much difference on these two ratio measures. When completing the V2 task (a less cognitively complex task with more planning time and prior knowledge), compared with the V1 task, it was found that lower proficiency Chinese speakers produced more complex terminable TC-units consisting of more single dependent TC-units. The reason for this could be that the progress of lower proficiency Chinese speakers is more salient in this case due to their lower base score in the V1 task compared with the other proficiency groups. On measure ③ MLSTCU, each group again scored almost identical in V1 and V2. However, different from the scores on the other three measures, both Group Low and Native participants performed better in V1

than V2 on measure ③ MLSTCU. For Group High, participants scored almost the same on V2 as on V1 with a very slightly advantage over V1 on measure ③ MLSTCU. Such score distributions indicated a possible trade-off effect at the clausal complexity level between the ratio measure ④ STCU/TTCU and length measure ③ MLSTCU. When

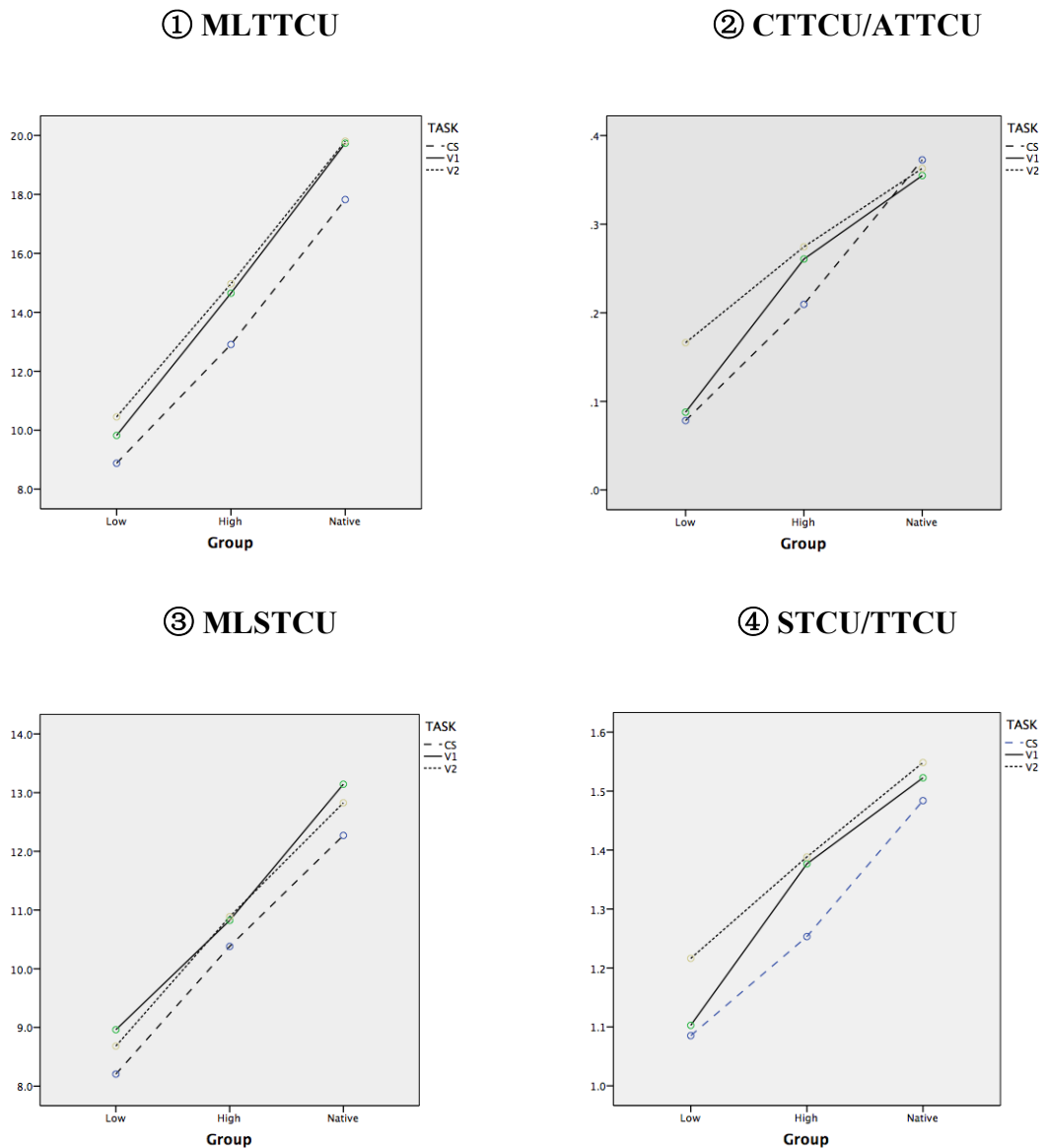


Figure 17. Estimated marginal means of four complexity measures on three tasks by proficiency groups.

longer single TC-units were produced, fewer single TC-units were combined into one terminable TC-unit. Conversely, in general when shorter TC-units were produced, more single TC-units were combined into one terminable TC-unit.

To sum up, by comparing syntactic complexity in the performance of the three tasks of varying presumed cognitive complexity, statistically significant more complex Chinese output was produced in the video retelling tasks of higher cognitive complexity along the resource-directing dimension, compared with the Chinese language output produced in the cartoon strip task of lower cognitive complexity along the resource-directing dimension. For the two video retelling tasks of which cognitive complexity varied along the resource-dispersing dimension, there was no statistically significant difference found in the output complexity of tasks V1 and V2. Overall, the Chinese output in the immediate task repetition V2 was not more complex than the V1 output although the low proficiency participants showed relatively more progress than the others in terms of complexity by immediately repeating the same video retelling task. By comparing participants' performance on V2 to their performance on V1 on the four measures, a possible trade-off effect was observed between measure ③ MLSTCU and measure ④ STCU/TTCU on the clausal complexity level.

5.2.4.2 Immediate task repetition effects on learners' self-perception

On the two retellings, V1 and V2, it was interesting to see that almost equal numbers of participants increased their complexity scores as those who decreased (including no increase) their complexity scores. As shown in Table 32 below, for all the participants, 58 participants increased 2.2 characters on average along the measure of ①

MLTTCU, while 52 participants decreased 1.7 characters on average. This was consistent within each proficiency group. The score on ① MLTTCU of V2 output was compared with V1 output in different proficiency groups. In Group Low, 20 participants increased 1.5 characters while 15 participants decreased 0.8 characters on average. In Group High, 23 participants increased 2.4 characters while 21 participants decreased 1.8 characters on average. In Group Native, 15 participants increased 2.6 characters while 16 participants decreased 2.3 characters on average.

It was also interesting to see that regardless of whether or not their Chinese syntactic complexity actually increased or decreased, participants' perceptions of their own performance and strategies applied to increase language complexity were mostly positive. Upon their completion of both task V1 and V2, all the participants filled out a retrospective survey. Table 31 below lists the seven questions in this retrospective survey. Of these seven questions, 1 and 5 were about participants' perceptions on their performance in terms of Chinese language complexity; 2 and 3 were about participants' perceptions on the strategies they applied to increase Chinese complexity; 4 was about participants' self-perceived anxiety during the test; and 6 and 7 were about their perceptions of the effect of immediate task repetition. For Questions 1-4, there were three optional answers provided for participants to choose: 1) the 1st time, 2) same for both times, and 3) the 2nd time. For Questions 5-7, the three optional answers were: 1) true, 2) not sure, and 3) false. For Questions 1-7 in Table 31, participants' positive perceptions of their V2 performance and strategies applied were coded 1; if negative, they were coded -1; if there was no difference or not sure, they were coded 0.

Questions 1-7 asked participants about their own perceived performance, strategy use, anxiety, and potential in V1, V2, and a hypothetical V3. In Table 32, participants' self-perception was coded according to different proficiency groups as well as the actual variation of ① MLTTCU comparing participants' output in V1 and V2. For Questions 1-7 in Table 32, out of a total of 56 (= 7 x 8) scores, there were only 4 (7.1%) minus scores that showed participants' negative perception regarding their repeated task performance, compared with the previous time. Regardless of the proficiency groups, and regardless of whether the score on ① MLTTCU actually increased or decreased, the 52 (92.9%) positive scores out of all the scores in Table 32 show that most participants were positive about their language complexity, strategy use, and confidence and ease in V2 compared with V1, regardless of their actual performance.

For Questions 1 and 5, scores of 0.3-0.6 show participants' self-perceived language complexity increase in task repetition, regardless of their actual performance. For Questions 2 and 3, scores of 0.3-0.7 show participants' positive self-perceived strategy application to increase complexity in task repetition. For Questions 4 and 6, scores of 0.2-0.6 (with one score value of -0.1) show participants' self-perceived increase of ease and confidence instead of weariness and impatience in task repetition. For Question 4, a total of 71 (64.5%) participants reported that they felt more ease and confidence at the 2nd time of retelling. For Question 6, 29 (26.4%) participants reported that they got tired and impatient the 2nd time as a result of speaking the 1st time. Participants' positive perception of their performance, strategy use, ease and confidence in V2 show that the immediate task repetition helped to lower test takers' anxiety, increasing their confidence instead of causing weariness in learners' test performance.

Table 31

Question list in the retrospective survey and coding of the answers

| | The 1st time | Same for both times | The 2nd time |
|---|------------------------------------|--------------------------------|------------------------------------|
| Q1. My language was more complex/sophisticated. | -1 | 0 | 1 |
| Q2. I deliberately tried harder to use more advanced words/structures. | -1 | 0 | 1 |
| Q3. I deliberately tried harder to use more diverse words/structures. | -1 | 0 | 1 |
| Q4. I felt more ease and confidence | -1 | 0 | 1 |
| | True | Not sure | False |
| Q5. As a result of the 1 st time speaking, I better structured words/expressions the 2 nd time. | 1 | 0 | -1 |
| Q6. As a result of the 1 st time speaking, I got tired and impatient the 2 nd time. | -1 | 0 | 1 |
| Q7. I could perform better if I had a 3 rd retelling opportunity. | 1 | 0 | -1 |

Table 32

Participants' retrospective perception on their performance and strategy in V1 and V2

| | | MLTTCU | | | | | | | | |
|---------------|---------------|----------|---------|-----|-----|-----|-----|-----|------|------|
| | | <i>n</i> | (V2-V1) | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 |
| Group | Not increased | 15 | -0.8 | 0.6 | 0.6 | 0.7 | 0.6 | 0.4 | 0.7 | 0.0 |
| | Increased | 20 | 1.5 | 0.6 | 0.6 | 0.6 | 0.6 | 0.4 | 0.4 | -0.1 |
| Low | Not increased | 20 | -1.8 | 0.4 | 0.6 | 0.7 | 0.4 | 0.3 | -0.1 | 0 |
| | Increased | 23 | 2.4 | 0.5 | 0.4 | 0.5 | 0.6 | 0.4 | 0.2 | -0.1 |
| High | Not increased | 16 | -2.3 | 0.6 | 0.6 | 0.4 | 0.5 | 0.5 | 0.5 | 0.2 |
| | Increased | 15 | 2.6 | 0.3 | 0.3 | 0.4 | 0.7 | 0.4 | 0.5 | 0.1 |
| Native | Not increased | 52 | -1.7 | 0.5 | 0.6 | 0.6 | 0.5 | 0.4 | 0.3 | 0 |
| | Increased | 58 | 2.2 | 0.5 | 0.4 | 0.5 | 0.6 | 0.4 | 0.3 | -0.1 |
| Total | Not increased | | | | | | | | | |
| | Increased | | | | | | | | | |

Question 7 was about participants' prediction of their V3 performance (if there was one). Self-perception was less positive here, and participants received scores of -0.1- 0.2. For Question 7, 28 (25.5%) did not think they could perform better with a 3rd retelling opportunity; half of the participants, 55 (50.0%), were not sure, and 27 (24.5%) believed that they could have performed better with a 3rd retelling opportunity.

These findings suggested that immediate task repetition lowered learners' communication anxiety and increased positive self-perception of their own performance. Nevertheless, the actual language complexity did not show any clear increase along with task repetition. Yet, learners' positive perception of their V2 performance and lower

anxiety in V2 may support the application of such immediate task repetition in L2 learning and teaching. When repeating the same task immediately, learners with lower communication anxiety were more willing to communicate. The sense of accomplishment, as a result of a higher self- perception of their performance, might also motivate learners to be more willing to communicate outside the classroom.

As discussed in Section 2.3.2, the self-perceived competence and lowered communication anxiety directly and indirectly predict L2 learners' willingness to communicate. In the long run, then, conducting immediate task repetition in classroom teaching can potentially contribute to the increase of their L2 competence.

As shown in Table 32, of all the participants, 52 (47.3%) showed no increase along the complexity measure ① MLTTCU in V2 compared to V1 performance; 58 (52.7%) participants' language complexity along measure ① MLTTCU increased in V2 compared to V1. This was the same for each proficiency group. There was a roughly equal number of participants who increased their complexity scores as participants who decreased their complexity scores on the retelling. However, for different proficiency groups, the major concerns of participants varied while completing the video story retelling task.

The last question of the retrospective survey solicited more open-ended reflection from the participants by asking the following: “Comparing your performance and strategies used for the two retellings, are there any other thoughts or comments you would like to share?” In their responses different concerns were addressed by different proficiency groups. The participants' text responses by proficiency group were visualized by pasting them into an online data visualization tool named *TagCloud*

(<http://tagcrowd.com>). This process provided visualization for the frequency distribution of keyword metadata that described the text input. The more frequently a keyword is repeated, the bigger the font by which it is presented in the visualized output. The similar words in the responses were grouped into sets. For instance, the words “*learn*”, “*learned*”, and “*learning*” were grouped into the set “*learn*”. Since the retrospective survey for the Chinese native speakers was in Chinese, the participants’ responses were translated into English in order to be consistent with the other two groups (see the generated visualization of the text by proficiency groups presented in Figure 18 below).

Group Low L2 Chinese speakers were more concerned with linguistic items when they completed the video retelling task. As shown in Figure 18, the words “*vocabulary*”, “*words*”, and “*sentence*” were among the keywords mostly mentioned in Group Low’s responses. The primary difficulty and frustration encountered by participants was their lack of necessary vocabulary, patterns, and grammar structures. As quoted from Group Low participants, “*I didn’t know some of the key words, which have frustrated me*”; and “*Substantial lack of the vocabulary I wanted to use created a slight frustration.*” Some of them were confident that they could have done a much better job if provided access to more vocabulary and grammar structures. As quoted, “*I would have liked a chance to look up certain vocabulary I didn’t know during either retelling*”; and “*I wish there was more vocabulary available. I could have done a much better job. I could have also done a better job if grammar structures were available, available in Pinyin.*” Group Low also showed a lower cognitive capacity when trying to recall the story. As a participant reflected, “*It was really difficult to retell the story even if I heard it many times. I think if I heard it in Chinese first, it would be easier. I think retelling an*

English story in Chinese is one of the hardest parts to teach in Chinese because there're many ways of doing it. Different structures/patterns can make it difficult for people to think quickly. Sometimes people might get caught up in translating it directly." Some Group Low participants expressed the wish to be able to take notes during the video or to be able to watch the video one more time before completing task V2. Another participant wrote, *"I feel that being able to take notes during the video will help test takers greatly in recalling the story as well as putting together complex and more accurate sentences. The problem I had encountered while doing this retelling was recalling the sequence of events as well as small details that would have improved my retelling greatly."*

Instead of being frustrated with not knowing the necessary vocabulary and grammar patterns as Group Low, Group High participants having a much more sufficient language repertoire were trying to use more precise language and better expressions. Among the mostly repeated keywords were *"vocabulary"*, *"language"*, *"organization"*, *"story"*, and *"details"*. Equipped with higher Chinese language proficiency to handle the linguistic challenge, Group High participants were more aware of the content of their narration. They tried to locate more attentional resources on the organization of their narration and worried more about leaving out too much detail. As quoted from Group High participants, *"Some details of the story I missed in the retelling because either I do not remember it or it was not important to the main idea of the story"*; *"I have a bad memory, so I'm sure I left out a lot of the details"*; and *"It was difficult for me because there was so much to remember and so much to organize in such a short time."* In addition, they were able to divert attention to the time limit when they were completing the task, especially for the second retelling. As the participants wrote, *"I paid more*

attention to the timing (the running numbers on the clock) the second time. I think I was able to do this because my thinking power was more available, since I had already thought about the story and how to tell it”; “The time limit gave me a lot of pressure”; and “I found myself simplifying a lot of vocabulary because I was being timed.” Also, some of them expressed that they were familiar with the use of repetition. One participant wrote, *“repetition and drilling even in story telling is how I would practice by myself in my room in college.”* Some of them performed immediate task repetition based on their initial task performance leaving them with a sense of self-satisfactory. As a participant reflected, *“I had practice after I did it the first time so I performed better during the second time. When I told the story the second time, I based it mostly on what I said the first time.”*

Group Native participants were mainly concerned with whether or not their story retelling included all the information from the video. Though the story of *Nian* is very well known, its storyline varied among different versions of this popular story. Among the mostly repeated words are “*story*”, “*content*”, and “*forgot*”. One participant stated, *“I did not deliberately use strategy as a native speaker. I just tried to tell the story more complete with more details.”* In addition to content completion, Group Native paid greater attention to the language quality used in terms of the story organization, fluency, language genre, and vividness. Participants stated, *“The first time retelling the story may have been disorganized because I was focusing on telling the whole story without missing details, and what I mainly did in 30 seconds preparation time was to recall the main points. I focused more on how to tell the story the second time because I was familiar what I was going to say. So it felt more like telling a story instead of describing an event.”*

What I did in the 30 seconds preparation time is to organize the storyline”; “I remembered many details the first time, but forgot them the second time. I paid attention to the narration genre the second time, paid attention mainly at the content the first time.



Figure 18. Text visualization of participants' response by proficiency groups.

The difference of the video story with the story I knew caused influence in my narration”;
and *“As the result of the first retelling, the second retelling was more fluent without too*
much pause. In addition, I paid more attention to the language the second retelling,
trying to use more vivid language.”

CHAPTER 6

DISCUSSION AND IMPLICATIONS

6.1 Qualitative analysis

By taking the same five tasks in TW&ST, Chinese speakers of varied proficiency levels demonstrated corresponding levels of Chinese syntactic complexity in their output by task. Generally, participants with higher level of Chinese language proficiency were able to produce more complex Chinese output compared with the participants having a lower level of language proficiency. The increase of Chinese syntactic complexity along with Chinese proficiency development suggested support for the proposed three stages of Chinese complexity development outlined and described by Jin (2006) (translated as): *threshold, growth, and leap*.

At the *threshold* stage, learners are not aware of the syntagmatic difference between Chinese and English. Transferring English syntagmatic mechanism to Chinese leads to overuse of subordinate structures, conjunctions, and explicit particulars and definite referring markers. Therefore, threshold stage can also be referred as subject-prominent stage. As their complexity development *grows*, learners start noticing the role that empty categories (clarified and referred as correferential zero in this dissertation) play when form a topic chain. They start conservatively employing empty categories to produce short topic chains. However, their use of empty category is limited to the theme/subject position but not patient/object position. Also, though learners at this stage are able to apply

empty category and demonstrative pronouns to produce some conventional simple topic-comment structures, they still produce redundant overt conjunctions, paralleled structures, complete form repetition of the same subject, and incohesive subject-predicate sentences. Language at this grow stage is a hodgepodge of subject-prominent and topic-prominent. At the final stage, *leap*, learners can be expected to lengthen Terminal Topic-Comment Units with more complex related clauses. They can proactively use empty categories, not limited to the subject position, but also employ empty categories as other components of the clause as well. In addition, more covert conjunctions are employed in topic chains. Topic-comment structures are also applied to form passive sentences. Learners' language complexity finally enters to the topic-prominent stage. (p. 134-5)

Based on the analysis in this dissertation, such developmental feature of Chinese complexity development can be more consistently illustrated in terms of TC-units. As designed in the TW&ST, the same cartoon strip was provided to all the participants in the CS task, the same video clip was played in the V1 and V2 tasks, the same topic was given for FW task, and the same series of Chinese sentences were provided in the GR task for the participants to rewrite. In both speaking and writing tasks, however, participants of higher Chinese proficiency produced longer terminable TC-units consisting of more dependent single TC-units, while participants of lower Chinese proficiency produced shorter terminable TC-units consisting of less single TC-units. These differences along

the proposed Chinese syntactic complexity developmental stages are further illustrated below with examples of the participants' output by task.

6.1.1 Overuse of complete sentences

In the output of the lower Chinese proficiency group, possibly owing to L1 transfer from English, the sentences were often complete in terms of consisting both subject and predicate. In the output of the lower Chinese proficiency participants, when the same subject is shared by the consecutive clauses or sentences, it is typically repeated in the form of either a complete repetition or a pronoun, but there is almost no evidence of correferential zero. However, a complex terminable TC-unit cannot be composed without a topic being repeated in the form of correferential zero. Therefore, with the topic redundantly repeated in the form of its full form, a pronoun, or a demonstrative, such a simple terminable TC-unit only consists of one independent single TC-unit and is not lengthened by connecting it with more dependent single TC-units to form a complex terminable TC-unit. Lacking the use of correferential zero which enables the forming of complex terminable TC-units as shown in Figure 9 and Table 15 (See Section 5.2.1), Group Low scored only .08 - .17 on measure ② CTTCU/ATTCU. Out of all the terminable TC-units, only 8 - 17% were complex terminable TC-units that consisted of two or more dependent single TC-units, while the remaining 83 - 92% were simple terminable TC-units each consisting of only one independent single TC-unit. On measure ④ STCU/TTCU, Group Low scored 1.08 - 1.22 which is the ratio of the total number of single TC-units (dependent and independent) divided by the total number of terminable TC-units (simple and complex). In other words, an average of 1.08 - 1.22 single TC-units

composed each terminable TC-unit in the output by Group Low. As a result of lacking complex terminable TC-units with each consisting more single TC-units, the length of a terminable TC-unit was very likely found to be limited to the length of one independent single TC-unit. Therefore, the score of 8.88 - 13.27 for Group Low on measure ① MLTTCU mostly overlapped with its score of 8.21-12.23 on measure ③ MLSTCU. As proficiency increased in Group High, scores on the two ratio measures ② CTTCU/ATTCU and ④ STCU/TTCU increased. Group High scored .17 - .27 on measure ② CTTCU/ATTCU, and 1.22 - 1.38 on measure ④ STCU/TTCU. Out of all the terminable TC-units, the proportion of complex terminable TC-units increased to 17 - 27%, and each terminable TC-unit on average consisted of 1.22 - 1.38 single TC-units in the output by Group High. Therefore, Group High's score of 12.91 - 15.45 on measure ① MLTTCU showed more deviation from its score of 10.38 -12.16 on measure ③ MLSTCU. Group Native scored .32 - .45 on measure ② CTTCU/ATTCU, and 1.39 - 1.82 on measure ④ STCU/TTCU. In other words, the proportion of complex terminable TC-units out of all the terminable TC-units increased to 32 - 45%, and an average of 1.39 - 1.82 single TC-units composed each terminable TC-unit. At this point, the score of 17.83 - 19.94 for Group Native on measure ① MLTTCU substantially surpassed its score of 10.74 -15.00 on measure ③ MLSTCU.

The following three quotes (69), (70), and (71) are all collected from participants' output in the V1 task to illustrate such development, (69) from Group Low, (70) from Group High, and (71) from Group Native. They all corresponded to the same line in the

video clip: “*Nian ate everything. He ate vegetables, meat, fruit, whatever he could find. And he even ate people.*”

(69) Group Low (EI Score: 43/120): (3 terminable TC-units; MLTTCU = 5.67

characters)

Dànshì tā shénme dōu chī. Tā chī qīngcài hé ròu. Tā dōu chī rén.

但是它_i什么都吃。它_j吃青菜和肉。它_k都吃人。

[But it_i whatever eats. It_j eats vegetables and meat. It_k eats people.]

(70) Group High (EI Score: 83/120): (1 terminable TC-units; MLTTCU=13

characters)

Tā chī cài, chī ròu, chī guǒzi, yěshì huì chī rén.

它_i吃菜， \emptyset_i 吃肉， \emptyset_i 吃果子， \emptyset_i 也是会吃人。

[It_i eats vegetables, \emptyset_i eats meat, \emptyset_i eats fruit, \emptyset_i also is able to eat people.]

(71) Group Native: (1 terminable TC-units; MLTTCU=24 characters)

Tā₁ huì chī shūcài, huì chī shuǐguǒ, huì chī ròu, zǒngzhī shì chī tā zhǎodào de yíqiè dōngxi.

它_i会吃蔬菜、 \emptyset_i 会吃水果， \emptyset_i 会吃肉，总之是 \emptyset_i 吃它找到的一切东西。

[It_i will eat vegetables, \emptyset_i will eat fruit, \emptyset_i will eat meat, in short is \emptyset_i eat it find DE all things.]

A comparison of (69), (70), and (71) on the four Chinese syntactic complexity measures is summarized in Table 27 below. Produced by a Group Low participant, (see end of sentence) (69) consisted of 3 terminable TC-units with each of these consisting of 1 independent single TC-unit. The topic *tā* (‘它’, it) occupied the beginning position of

each of the 3 single TC-units of (69), which kept these 3 single TC-units independent from composing a complex terminable TC-unit via correferential zero. For each terminable TC-unit in (69), its length was equal to the length of the one single independent TC-unit it consisted of. The mean length of terminable TC-units therefore was equal to the mean length of single TC-units in (69). In contrast, in (70) and (71), the topic *tā* (它, it) only showed up once in its full form at the beginning of the first single TC-unit, while the subsequent 3 single TC-units all repeated the same topic *tā* (它, it) via coreferential zero which formed these 3 subsequent single TC-units dependent. With a total of 4 dependent single TC-units each, (70) and (71) formed two complex terminable TC-units. Terminable TC-unit (70), the output of a Group High participant, consisted of 4 dependent single TC-units which contributed to a MLTTCU of 13 characters. Produced by a L1 Chinese speaker, (71) had 1 terminable TC-unit consisting of 4 dependent single TC-units of 24 total characters. The length of the terminable TC-unit in (70) and (71) was much increased by applying coreferential zero instead of full forms of the same topic taking the form of complete sentences. The output of Group High and Native participants showed a higher score on ④ STCU/TTCU, which means their terminable TC-units on average consisted of more dependent single TC-units. Comparing the output of Group High (70) and Group Native (71), (71) not only consisted of more but also longer dependent single TC-units with an attributive modifier included.

As shown in Table 27, the transition from redundantly applying complete sentences to that of composing a greater amount of complex terminable TC-units showed a positive effect for Chinese syntactic complexity on measure ① MLTTCU. Such a positive effect was also shown to take place on ratio measure ② CTTCU/ATTCU and ④

STCU/TTCU. However, on measure ③ MLSTCU the development seemed not to be as linear. As Group High participants tried to avoid using redundantly complete sentences and by noticing the different syntagma between English and Chinese, the mean length of their single TC-units might have been shortened since the topic was not redundantly repeated but instead replaced with no phonological content.

Table 27

Comparison among (69), (70), and (71) on four Chinese syntactic complexity measures

| | EI | ① | ② CTTCU | ③ | ④ STCU |
|---|--------------|---------------|----------------|---------------|--------------|
| | Score | MLTTCU | /ATTTCU | MLSTCU | /TTCU |
| Group Low output (69) | 43 | 5.7 | 0 | 5.7 | 1 |
| Group High output (70) | 83 | 13.0 | 1 | 3.3 | 4 |
| Group Native output (71) | N/A | 24.0 | 1 | 6.0 | 4 |

6.1.2 Single TC-unit lengthening and combining

To produce more complex sentences, lower proficiency English speaking L2 Chinese speakers tended to rely on lengthening single TC-units in order to increase global Chinese syntactic complexity. This reliance on subclausal level complexity syntagma was likely transferred from their use of relative, complement, and adverbial clauses in English. More lengthening was undertaken within each independent single TC-

unit by embedding more complex structures, attributive and adverbial modifiers, serial verb constructions, and verb arguments. There was less awareness in Group Low participants to increase global Chinese syntactic complexity by composing more dependent single TC-units into a terminable TC-unit via coreferential zero. For example, the following three excerpts from the CS task output, (72), (73), and (74), were all describing the same two pictures below.



(72) Group Low (EI score 33/120): (2 terminable TC-units; MLTTCU= 9 characters)

Māma ràng háizi zhǎo bàba. He, bàba juéde shū hěn yǒuyìsi.

妈妈_i让孩子找爸爸。嗨，爸爸_j觉得书很有意思。

[Mother_i let child look for father. Ah, father_j felt the book very interesting.]

The mother sent the child to look for his father. Ah, (it seems that) father found the book very interesting.

(73) Group High (EI Score: 110/120): (2 terminable TC-units; MLTTCU= 14 characters)

Ránhòu māma jiù jiào érzi qù jiào bàba lái chīfàn, búguò bàba háishì hěn tóurù de zài kànshū.

然后妈妈_i就叫儿子去叫爸爸来吃饭，不过爸爸_j还是很投入地在看书。

[Then mother_i just ask son go call father come eat meal, but father_j still very devotedly DE at read book.]

Then the mother asked the son to call his father for dinner, but the father was still very engaged in reading the book.

(74) Group Native: (4 terminable TC-units; MLTTCU= 19.3 characters)

Ránhòu māma jiù duì lǐ kǎ shuō: “Nǐ qù kàn yí xià bàba dàodǐ shì zài gàn ma ne, zěnmē hái bù guòlái chīfàn.” Lǐ kǎ gǎnjǐn pǎo chūqù, ránhòu tuī kāile fángmén, a, yuánlái bàba₄yě zhèngzài dūn zài dìshàng kàn zìjǐ méiyǒu kàn wán dì nà běnshū ne, érqǐě kàn dé tèbié guānzhù.

然后妈妈₁就对里卡说：“你去看一下爸爸₂到底是在干嘛呢， \emptyset_2 怎么还不过来吃饭？”里卡₃赶紧跑出去， \emptyset_3 然后推开了房门，啊，原来爸爸₄也正在蹲在地上看自己没有看完的那本书呢， \emptyset_4 而且看得特别关注。

[Then mother₁ to Rica said, “you go see a bit father₂ really is at doing what, \emptyset_2 how come still not come to the dinner?” Rica₃ rushed ran out, then \emptyset_3 pushed open the door, ah, it turned out that father₄ also in the course of squatting on the floor read self not yet finish book, \emptyset_4 plus read particularly concentrated.]

Then Mother said to Rica, “Go and see what your father is doing. Why isn’t he coming for dinner yet?” Rica rushed out, and then pushed the door opened. Oh, it turned out that father was squatting on the floor reading that book that he has not finished yet, and he was especially engrossed in his reading.

Produced by a Group Low participant, (72) consisted of 2 terminable TC-units each of which consisted of 1 independent single TC-unit. Produced by a Group High

participant, (73) also consisted of 2 terminable TC-units each of which consisted of 1 independent single TC-unit as well. All 4 of single TC-units in (73) and (74) have complete sentence structures in terms of employing both subject and predicate. However, the 2 terminable TC-units in (73) compared with (72) were better coordinated by the conjunction *búguò* (不过, however). In addition, with more complex structures or modifiers embedded, the terminable TC-units in (73) were longer than those in (72). In (73), the first terminable TC-unit embedded a double-layered serial verb construction which took the form of two causative structures, and the second terminable TC-unit included an adverbial modifier “*tóurù de* (投入地, engrossed in)” preceding the predicate “*zài kànshū* (在看书, at reading)”. As a result, the mean length of terminable TC-units in (73), 14 characters, was much higher than the mean length of terminable TC-units in (72), 9 characters. As for (74) produced by a Group Native participant, it consisted of 4 terminable TC-units with a mean length of terminable TC-units of 19.3 characters. While in (72) and (73) all the terminable TC-units were simple terminable TC-units consisting of 1 independent single TC-unit each. The score of (74) on measure STCU/TTCU was 1.8 indicating that each terminable TC-unit in (74) on average consisted 1.8 dependent single TC-units. However, the mean length of single TC-units for (74) was 11 characters which was longer than the 9 characters in (72), but shorter than 14 characters in (73). It seemed that while English speaking L2 Chinese speakers of lower proficiency were rather limited in lengthening single TC-units as the primary method to increase global complexity, Chinese speakers of native or native-like proficiency relied more on combining more single TC-units into a complex terminable TC-unit. Complexity making in the form of combining more dependent single TC-units via correferential zero

compared with lengthening each single TC-unit, played a more crucial role in syntactic complexity making for Chinese speakers of native or native-like proficiency.

Table 28

Comparison among (72), (73), and (74) on four Chinese syntactic complexity measures

| | EI | ① | ② CTTCU | ③ | ④ STCU |
|---------------------------------|--------------|---------------|----------------|---------------|---------------|
| | Score | MLTTCU | /ATTTCU | MLSTCU | /TTCU |
| Group Low output (72) | 33 | 9.0 | .0 | 9.0 | 1.0 |
| Group High output (73) | 118 | 14.0 | .0 | 14.0 | 1.0 |
| Group Native output (74) | N/A | 19.3 | 0.8 | 11.0 | 1.8 |

For English speaking L2 Chinese speakers then, there seemed to have been a transition from mainly relying on borrowing English syntagmatic mechanisms to that of eventually realizing and applying the Chinese syntagmatic mechanisms in developing their Chinese syntactic complexity. At the initial stage of complexity development, lacking sufficient knowledge of the contrast between topic-prominent Chinese versus subject-prominent English, the English speaking L2 Chinese speakers' clause complexing primarily relied on adding the Chinese equivalent of English relative, complement, or adverbial clause, which resulted the fact that each single TC-unit is lengthened. As L2 Chinese speakers obtained higher language proficiency, more Chinese syntactic

complexing took place by combining more single TC-units into a complex terminable TC-unit via correferential zero(s).

The descriptive statistics in Section 5.2.1 confirmed an emerging transitional preference away from lengthening to that of combining single TC-units along proficiency increase. As shown in Table 15 in Section 5.2.1, for the three spoken Chinese data sets the mean score distance between Group Low and Group High on measure ③ MLSTCU was 2.08 characters, and between Group High and Group Native it was 2.05 characters. But on measure ① MLTTCU, a correspondingly much bigger mean score gap of 4.46 characters was found between Group Low and Group High, and a mean score gap 4.95 characters was found between Group High and Group Native. While on measure ③ MLSTCU Group Native did not surpass Group High as much as Group High surpassed Group Low, the extra length of .49 characters on ① MLTTCU between Group Native and Group High was compensated by more dependent single TC-units composed into one terminable TC-unit, as well as by the increased number of complex terminable TC-units in total.

The transitional preference away from lengthening to that of combining single TC-units along with a proficiency increase was also confirmed by the correlational analysis in Section 5.2.2 between the Chinese syntactic complexity and L2 Chinese speakers' EI proficiency scores. As shown in Table 16, while at the global complexity level the length measure ① MLTTCU generally indicated a higher correlation with proficiency level when compared with the ratio measure ② CTTCU/ATTCU, at the clausal complexity level the ratio measure ④ STCU/TTCU when compared with the length measure ③ MLSTCU indicated a higher correlation with proficiency level. The increased combining

of single TC-units into complex terminable TC-units contributed to the higher correlations on the ratio measure ④ STCU/TTCU. Meanwhile, not as much of an increase noted on the length of single TC-units and this might have led to the relatively lower correlation on the length measure ③ MLSTCU. Along with the proficiency increase, a global Chinese complexity increase via lengthening the single TC-units and combining more single TC-units into one terminable TC-unit showed an uneven and perhaps nonlinear developmental trajectory. In addition to the difference shown on the length and ratio measures, such uneven development was also displayed across the different complexity levels of global, clausal, and subclausal phenomena. While the lengthening mechanism was less favored at the clausal level and led to some loss at both clausal and global complexity levels, its loss at the global level might have been compensated by more combining more of the single TC-units which helped form longer terminable TC-units. This might have explained the high correlation with proficiency level on the global measure ① MLTTCU regardless of the variation on other measures for each task.

Comparing the correlations on four complexity measures between the V1 and V2 tasks as shown in Table 16, the correlation of .79 on ① MLTTCU for V1 and .77 on V2 were similar. However, the V2 task showed higher correlation on ③ MLSTCU, and much lower correlations on both ② CTTCU/ATTCU and ④ STCU/TTCU when compared with the V1 task. Such relative higher correlation on ③ MLSTCU and lower correlations on both ratio measures ② CTTCU/ATTCU and ④ STCU/TTCU in the V2 data set indicated that overall longer single TC-units were produced along with the

proficiency increase in the V2 task. However, the amount of complex terminable TC-units and the amount of single TC-units a complex terminable TC-unit consisted of were not as consistently produced along with the proficiency increase as in the V1 task. The inconsistency between the growth on ③ MLSTCU and the drops on ② CTTCU/ATTCU and ④ STCU/TTCU in V2 might be an indication of some trade-off effect between the lengthening and combining of single TC-units at the clausal complexity level. Since the V2 task was an immediate repetition of the V1 task and the participants were told to take this second chance to improve on their performance, some participants might have been subconsciously trying to lengthen their single TC-units in order to add complexity to their language while at the same time being unaware of the preference for single TC-unit combining at higher proficiency levels. Given that longer single TC-units were produced as proficiency increased, as a trade-off effect, fewer single TC-units were combined into one terminable TC-unit which also led to fewer complex terminable TC-units being produced. On measure ③ MLSTCU, comparing the performance of L2 Chinese speaker participants in the V1 and V2 task, an almost equal number of the participants within each proficiency group increased their complexity scores as those who decreased their complexity scores. Though some L2 Chinese speakers of higher proficiency employed the combining of more single TC-units to produce more complex language output, they were not equipped with the declarative/explicit knowledge of such a Chinese syntactic complexity making preference. If given more practice combining single TC-units into a complex terminable TC-unit via coreferential zero, English speaking L2 Chinese speakers can then more efficiently I the targeted Chinese syntactic complexity development without too much of a detour.

Below, (75) and (76) represents spoken output data elicited from the V1 and V2 task by the same Group High participant who attained a score of 115/120 on the EI test. (75) and (76) were about the same for content based on the story played in the given video clip. When comparing the four complexity measures in Table 29, it can be seen that there were a total of 5 complex terminable TC-units in (75) with a mean length of 19.6 characters. Each terminable TC-unit consisted of an average of 1.8 single TC-units of 9.3 characters. By contrast, (76) consisted of only two complex terminable TC-units of which the mean length was 32.5 characters which was much longer than the mean 19.6-character length of terminable TC-units for (75). However, the 7.2-character mean length of the single TC-units in (76) was shorter than the 9.3-character mean length of single TC-units in (75). There was an average of 4.5 dependent single TC-units in each terminable TC-unit in (76) which greatly surpassed the average of 1.8 dependent single TC-units in each terminable TC-unit in (75) and contributed to the much longer mean length of terminable TC-unit of (76). Therefore, the combining more single TC-units into one terminable TC-unit can effectively increase Chinese syntactic complexity, and this seemed to occur more frequently among learners at higher proficiency levels.

(75) *Nián xià shān yǐhòu lái dào xiāngcūn lǐ kànjiàn yǒu huǒ, tā bù xǐhuān huǒ.*

Kànjiàn xiāngcūn lǐ de rén chuān de shì hóngsè de yīfú, tā bù xǐhuān hóngsè. Ránhòu tīngjiàn nàgè zhúzi zài shāo de shíhòu huì xiǎng, tā yě bù xǐhuān zhè zhǒng bàozhú de shēngyīn. Suǒyǐ nián jiù tuìle, wǎng shāndǐng shàng tuì, zài yě méi xiàláile.

年₁下山以后来到乡村里看见有火，它₂不喜欢火。Ø₃看见乡村里的人穿的是红色的衣服，它₃不喜欢红色。Ø₄然后听见那个竹子在烧的时候会响，它₄也不喜欢这种爆竹的声音。所以年就退了₅，Ø₅往山顶上退，Ø₅再也没下来了。

[Nian₁ down the mountain after arrive village-in saw existing fire, it₂ not like fire. Ø₃Saw village-in people wearing red DE clothes, it₃ not like red color. Ø₄ Then heard that bamboo at burning time will sound, it₄ also not like this kind craking bamboo DE sound. Therefore Nian₅ retreated, Ø₅ toward mountain top-on retreat, Ø₅ again also not down-come LE.]

Nian came down from the mountain to the village and saw fire there. It did not like fire. *Nian* saw the village people wearing red clothes, but it did not like the red color. Afterwards, *Nian* heard the bamboo making noise when being burnined, but it did not like the noise of bamboo craking either. So *Nian* retreated back to the mountain top and has never come down ever since.

(76) *Nián xià shān yǐ hòu, kànjiàn huǒ, tā jiù hài pà wǎng hòu tuì. Kànjiàn rén chuān de yīfú dōu shì hóngsè de, yǒudiǎn xiàng huǒ, tā gèng hài pà. Ránhòu zuìhòu tīngjiàn zhúzi bào de shēngyīn, jiù sātuǐ jiù pǎole, wǎng shàng shāndǐng shàng tuì, zài yě méi xiàshān le.*

年下山以后，Ø₁看见火，它₁就害怕往后退。Ø₂看见人穿的衣服都是红色的，有点像火，它₂更害怕。Ø₂然后最后听见竹子爆的声音，Ø₂就撒腿就跑了，Ø₂往山顶上退，Ø₂再也没下山了。

[Nian down the mountain afterwards, Ø₁saw fire, it₁ then scared toward back retreat. Ø₂ Saw people wear DE clothes all are red, a little looked like fire, it₂ even more scared. Thereafter at last Ø₂ heard bamboo cracking noice, then Ø₂ let go legs and ran away LE, Ø₂ toward up mountain top retreat, Ø₂ again also not down mountain LE.]

After it got down from the mountain, Nian saw fire. Nian got scared and retreated back. Seeing the clothes people wearing were all red which looked like fire, Nian got even more scared. And then at last as Nian heard the cracking noise of bamboo, it ran away up to the mountain top. Nian had never came down from the mountain ever since.

Table 29

Comparison among (75) and (76) on four Chinese syntactic complexity measures

| | EI | ① | ②CTTCU | ③ | ④STCU |
|-----------------------|--------------|---------------|---------------|---------------|--------------|
| | Score | MLTTCU | /ATTCU | MLSTCU | /TTCU |
| V1 output (75) | 115 | 19.6 | 1.0 | 9.3 | 1.8 |
| V2 output (76) | | 32.5 | 1.0 | 7.2 | 4.5 |

The transitional preference from more lengthening to more combining of the single TC-unit along with a proficiency increase seemed particularly salient in written Chinese. As shown in Figure 9 (See Section 5.2.1), the FW task displayed evenly distributed mean scores on measure ① MLTTCU in line with other task data sets. However, for the FW task, the mean scores on the other three measures were unevenly distributed in contrast to speaking task data sets. Group Native scored lower on measure ③ MLSTCU but much higher on both measure ② CTTCU/ATTCU and measure ④ STCU/TTCU than Group High did. In other words, in the FW task for Group Native, the increase on measure ① MLTTCU was mostly the result of contributions being made by the increases on measure ② CTTCU/ATTCU and ④ STCU/TTCU despite the decreases

on measure ③ MLSTCU as compared with Group High. The writing modality may have provided Chinese speakers more chances to produce complex terminable TC-units that consisted of more single TC-units than did the speaking modality. Still, such single TC-units could be relatively shorter compared with the spoken output.

6.1.3 Hypotactic and paratactic syntagma

For parataxis-prominent languages or hypotaxis-prominent languages, rather than exclusively applying parataxis or hypotaxis, the more frequent case is for both of these two mechanisms to be employed with varying partial adoption. As a parataxis-prominent language, Chinese speakers prefer applying covert conjunctions but do not exclude overt conjunctions. There are two types of overt conjunctions used in Chinese language. One is to connect single TC-units without specified semantic references such as *hé* (和, and), *gēn* (跟, and), *yě* (也, also) at the clausal level which, however, are incorrect when coordinating Chinese single TC-units. The other conjunctions such as *kěshì* (可是, but), *yīnwèi* (因为, because), *rúguǒ* (如果, if), *suǒyǐ* (所以, so), and *chúle* (除了, except), have more concrete semantic references. It is perfectly grammatical to apply such overt conjunctions in Chinese. The way such overt conjunctions specify the semantic relation between single TC-units contributes to clarity and preciseness in Chinese. However as discussed in Section 2.2.2.4, when sufficient contextual clues were provided, an implicit connection in the form of juxtaposition like in (51) was preferred by native speakers as compared with the explicit marking of coordination or subordination found in (52). As Lian (1993) suggested, English sentences were more precise and Chinese sentences were more concise. In this case, on condition that sufficient contextual clues are provided,

conciseness via covert conjunctions between single TC-units is preferred than preciseness via overt conjunctions between single TC-units.

English is a hypotaxis-prominent language; English speakers therefore tend to transfer their English hypotactic syntagma to Chinese speaking and writing. As English speaking L2 Chinese speakers learned to combine more dependent single TC-units in addition to lengthening each single TC-unit to achieve higher global Chinese syntactic complexity, the syntagmatic mechanism they applied to connect such dependent single TC-units still remained a transfer from the hypotaxis-prominence of English. Therefore, instead of applying covert conjunctions, redundant overt conjunctions were employed by L2 learners to connect their Chinese dependent single TC-units, such as *hé* (和, and), *gēn* (跟, and), *yě* (也, also). As a conjunction, “and” is used to connect words, phrases, clauses, or sentences in English; its correspondent in Chinese, *hé* (和, and), however, is only good at the word or phrasal level but not clausal or sentential level. (77) and (78) below were both the output of Group Low participants, in which the clausal level “and” in English was transferred as *hé* (和, and) to connect clauses. However, (77) and (78) are not grammatical because Chinese as a parataxis-prominent language does not use *hé* (和, and) to connect clauses or sentences. Such L1 transfer does not work for *yě* (也, also) at the clausal and sentential level, either. “Also” in English can be used to connect words, phrases, clauses, or sentences. However, its semantic equivalent *yě* (也, also) in Chinese is an adverb that precedes the verb within a clause. In (79), the output of a Group Low participant, *yě* (也, also) was ungrammatically used to connect clauses in Chinese.

(77) *Tā jiāo wǒ zěnmē shuō Spanish hé yìn xiàng wǒ zuò zhōngwén.*

她教我怎麼說 spanish 和印象 {影響} 我做中文。

[She teach me how say Spanish **and** influence me do Chinese.]

She taught me how to speak Spanish and influenced me to do Chinese.

(78) *Tā lái xiàwēiyí 2006 nián hé wǒ lái 2007 nián.*

他来夏威夷 2006 年和我来 2007 年。

[He come Hawaii 2006 year **and** I come 2007 year.]

He came to Hawaii in 2006 and I came in 2007.

(79) *Wǒmen xǐhuān bù yíyàng de yīnyuè, yě bù yíyàng de diànyǐng.*

我們喜歡不一樣的音樂，也不一樣的電影³。

[We like not the same music, **also** not the same movie.]

We like different kinds of music and also different kinds of movies.

As a hypotaxis-prominent language, English prefers employing overt conjunctions to connect sentences. As a result, overt conjunctions of the second type were also kept in the Chinese output of English speaking L2 Chinese speakers. As a parataxis-prominent language, Chinese does not exclude the use of overt conjunctions in clauses or sentence connecting, however, covert conjunctions are favored when producing more concise and complex information units. The following written outputs (80), (81), and (82) were produced in the FW task respectively by participants of Group Low, Group High, and Group Native. In a developmental diagonal, (80), (81), and (82) can be viewed as snapshots of three stages along the interlanguage developmental

³ When completing TW&ST, participants could type in either simplified or traditional Chinese script. The output quoted in this dissertation was all kept in their original form.

trajectory of an English speaking L2 Chinese speaker. At the low proficiency level, the sentences mostly take the forms of redundant complete sentences and are not well connected, as shown in (80), where the sentences were relatively independent and isolated. As an L2 Chinese speaker's proficiency level increased, sentences were more closely connected or layered. However, as in English the hypotactic syntagmatic mechanism remained dominant at this stage and therefore various overt conjunctions were employed to achieve higher Chinese complexity. (81) was produced by a Group High participant. Similarly, (80) and (81) were also talking about relationships with the parent, but a variety of conjunctions including *kěshì* (可是, but), *yīnwèi* (因為, because), *rúguǒ* (如果, if), *suǒyǐ* (所以, so), and *chúle* (除了, except) were used to connect sentences. These overt conjunctions did connect the clauses and resulted in more cohesion even between the redundant complete sentences. It is also grammatical to apply overt conjunctions in Chinese. However, as Chinese language proficiency increases, instead of employing the preferred English hypotactic mechanism in order to achieve higher syntactic complexity, speakers choose to employ a clause in juxtaposition without explicitly indicating its structural relationships. (82) was the output of a Group Native participant which consisted of 8 terminable TC-units as well as 18 dependent single TC-units (STCU). In (82) there were an average of 2.3 dependent single TC-units paratactically connected in each terminable TC-units with only one overt conjunction, *dànshì* (但是, but), used once in the whole paragraph.

(80) *Wǒ gēn wǒ de bàba yǒu yí gè hěn hǎo de guānxi. Tā huì gěi wǒ hěn hǎo de jiànyì.*

Tā hěn yònggōng, kěshì wǒ méiyǒu tā nàme yònggōng. Wǒ néng kànshū liǎng gè

xiǎoshí, tā néng kànshū liǎng gè xīngqī. Wǒ yào yìtiān zhǎng dà gēn wǒ de bàba

yìmóyiyàng. Wǒ gēn wǒ de bàba xǐhuān zài diànshì shàng kàn lánqiú, kěshì xiànzài
wǒmen de diànshì méiyǒu rén yòng.

我跟我的爸爸有一个很好的关系。他回{会}给我很好的建议。他很用工
{功}，可是我没有他那吗{么}用工{功}。我能看书两个小时，他能看书两个星
期。我要一天站{长}大跟我的爸爸一模一样。我跟我的爸爸喜欢在电视上看蓝筹
{篮球}，可是现在我们的电视没有人用。

[I with my father have a very good relationship. He will give me very good
suggestions. He very work hard, but I not him that work hard. I can read book two
hours. He can read book two weeks. I want one day grow up with my father same. I
with my father like on the TV watch basket, but now our TV nobody use.]

I have a very good relationship with my father. He will give me very good
suggestions. He is very hard working, but I am not as hard working as him. I can read
for two hours, while he can read for two weeks. I want to be exactly like my father
when I grow up. I like watching basketball on TV with my father, but nobody is using
my TV now.

(81) Wǒ gēn wǒ māma de guānxì shì hěn hǎo. Kěshì tā hěnduō shíhòu huì ràng wǒ hěn
shēngqì, yīnwèi tā chángcháng dōu huì jiào wǒ bāng tā zuò hěnduō dōngxī. Rúguǒ wǒ
míngtiān yào kǎoshì tā dōuhuì jiào wǒ bāng tā. Suǒyǐ xiànzài yīnwèi wǒ yǒu hěnduō
kǎoshì, rúguǒ tā jiào wǒ bāng tā, wǒ dūhuì bùguǎn tā. Chùle tā chángcháng jiào wǒ
bāng tā de shíhòu. Tā shì yí gè hěn hǎo de māma. Wǒ shénme dōu néng gǒu gēn tā
shuō. Tā chángcháng

我根{跟}我媽媽的關係是很好。可是他{她}很多時候會讓我很生氣，因為他{她}常常都會叫我幫他{她}做很多東西。如果我明天要考試他{她}都會叫我幫他{她}。所以現在因為我有很多考試，如果他{她}叫我幫他{她}，我都會不管他{她}。除了他{她}常常叫我幫他{她}的時候。他{她}是一個很好的媽媽。我甚麼都能夠{够}根{跟}他{她}說。他{她}常常

[I with my mother DE relationship is very good. But she often times will let me very angry, because she always will ask me help her do many things. If I tomorrow need to exam she all will ask me help her. So now because I have many exams, if she ask me to help her, I all will not care her. Besides she always ask me help her De time. She is a very good mother. I even all can with her speak. She always]

The relationship between my mother and I is good. But she will often make me angry, because she will always ask me to help her with many things. Even if I have an exam the next day she will still ask me to help her. Therefore, since I have many exams now, whenever she asks me to help her, I ignore her. Aside from her always calling me for help, she is a great mother. I can tell her anything. She always...

(82) *Wǒ de mèimei jīnnián niàn chūzhōng yì niánjí. Tā zài dì yī xuéqī jiéshù de shíhou, qǔdéle quán bān dì yī míng de chéngjī, quánjiā rén dōu hěn wèi tā gāoxìng, gǔlì tā jìxù nǚlì. Tā xiǎndé hěn bù yǐwéirán, rènwéi dì yī míng méiyǒu shéme liǎobùqǐ de, zhòngyào de shì tā xǐhuān zìjǐ qǔdé hǎochéngjī de guòchéng. Tā yuánběn shì yí gè bú ài xuéxí de hái zi, xǐhuān kàn dònghuà piān, xǐhuān kàn nǎo cǎn de liánxùjù, zài xiǎoxué liù niánjí de shíhòu hái mǐliàn qīngchūn ǒuxiàng diànyǐng, yí dù hěn bèi wō bǐshì. Dàn shì zhuǎnbiàn jiù xiàng zài shùnjiān fāshēng, tā shàngle chūzhōng yǐhòu jiù wánquán*

*miànmào yì xīn, xǐhuān tiānwén kēxué, xǐhuān yīnyuè, xǐhuān yùndòng, zhènggè rén
dōu sànfā chū huóli.*

我的妹妹₁今年念初中一年级。她₂在第一学期结束的时候，取得了全班第一名的成绩，全家人₃都很为她高兴， \emptyset_3 鼓励她继续努力。她₄显得很不以为然， \emptyset_4 认为第一名没有什么了不起的，重要的是₅她喜欢自己取得好成绩的过程。她₆原本是一个不爱学习的孩子， \emptyset_6 喜欢看动画片， \emptyset_6 喜欢看脑残的连续剧， \emptyset_6 在小学六年级的时候还迷恋青春偶像电影， \emptyset_6 一度很被窝{我}鄙视。但是转变₇就像在瞬间发生，她₈上了初中以后就全面貌一新， \emptyset_8 喜欢天文学， \emptyset_8 喜欢音乐， \emptyset_8 喜欢运动， \emptyset_8 整个人都散发出活力。

[My younger sister₁ this year read junior high first year. She₂ when the first semester ended, obtained whole class the first place DE grade, the whole family₃ all very for her happy, \emptyset_3 encourage her continue effort. She₄ seems very not regarding it as right, \emptyset_4 thought the first place does not have something amazing, what is important₅ is she like herself gain good grade DE process. She₆ originally was a not-love study DE child, \emptyset_6 like watching cartoons, \emptyset_6 like watching brain-handicapped dramas. \emptyset_6 During the sixth years in elementary school still obsess is youth movie, \emptyset_6 for a while by me was looked down upon. However, the change₇ just happed within one day, she₈ after going to junior high school her appearance all new, \emptyset_8 likes astronomy, \emptyset_8 like music, \emptyset_8 like exercise, \emptyset_8 the whole person sending forth vitality.

My younger sister is in her first year of junior high this year. At the end of her first semester, she placed first in her class. The whole family was happy for her, and encouraged her to continue working hard. She appeared to disagree and did not think

that first place was a big deal. What mattered to her was the process by which she achieved good grades. In the past, she was a child who did not like studying. She liked watching cartoons, and brain-damaging soap operas. In the 6th grade, she was even obsessed with youth movie, which was something I had always disdained. However, the change seemed to have taken place in a flash. When she attended junior high, she totally changed to another person. She liked astronomy, music, and exercise. Her entire being was full of vitality.

Such transition from applying more hypotactic syntagma to more paratactic syntagma in terms of transition from applying more overt connectives to more covert connectives, however, does not necessarily exert a positive increase on the length measure of Chinese syntactic complexity, especially on the clausal complexity level. As covert connectives are phonologically unpronounced, more application of such paratactic syntagma causes a decrease in the mean length of terminable TC-units as well as the single TC-units. It might cause a trade-off with the increase on length measures contributed by other complexity composition mechanisms. This is also in line with the descriptive statistics in Section 5.2.1, on measure ③ MLSTCU where the score increase from Group High to Group Native was not as much as the increase from Group Low to Group High but the overall increase on measure ① MLTTCU along with proficiency increase was compensated by more dependent single TC-units in one terminable TC-unit, as well as by more complex terminable TC-units in total. This comes back to advocating for an organic approach to syntactic complexity investigation. Both length and ratio measures should be included in order to detect syntactic complexity on varied levels of

global, clausal, and subclausal complexity. While Chinese speakers of lower proficiency do not form complex terminable TC-units or mostly form only the typical topic chain, Chinese speakers of higher proficiency do form complex terminable TC-units in the form of varied types of topic chains including the telescope chain. For more advanced Chinese output, the ratio for different types of terminable TC-units should also be included as measures to provide insights into syntactic complexity.

6.1.4 Typical topic chain and varied types of topic chain

Using a database of 24,000 characters including works of Lao She, Bing Xin and Chen Rong, Li (2005) found the typical topic chain was the most frequently used pattern. The database identified a total of 1,158 (74.8%) tokens of typical topic chain and a total of 390 (25.2%) tokens of all the other types of topic chain. Such high frequency in use suggested that this typical topic chain is a common and easily accessible structure. Li (2005) also observed that the major factors influencing the frequency ranking are the styles of writing and the subject matter of the texts.

The distribution of the ten types of topic chain was found related to the language proficiency in this dissertation. No topic chains or some typical topic chains were found in the output of speakers of lower Chinese proficiency. However, Chinese speakers of higher proficiency levels were found to be able to produce more varied types of topic chains. In other words, complex terminable TC-units of more varied inner-structures were produced in their language output.

For example, in the GR task, provided with the same sentences, more varied types of terminable TC-units were composed by higher proficiency participants compared with

the participants of lower level. As we introduced in 4.2.2, seven semantically coherent but formally incohesive sentences (in both simplified and traditional Chinese scripts) were provided for the participants to connect into one coherent paragraph by manipulating the sentences, changing the word order, or omitting words but without leaving out any of the given information (See Appendix B or C for the complete task with seven sentences). Below are the English translations of these seven sentences in Table 30.

Table 30

Sentence provided for the GR task in TW&ST and their English translations

| Sentence provided in TW&ST | English translations |
|---|--|
| <ul style="list-style-type: none"> • 王朋在学校的宿舍住了两个学期了 | <ul style="list-style-type: none"> • Wang Peng has lived in a dorm on campus for two months already |
| <ul style="list-style-type: none"> • 王朋觉得宿舍太吵 | <ul style="list-style-type: none"> • Wang Peng feels the dorm is too noisy |
| <ul style="list-style-type: none"> • 王朋觉得房间太小 | <ul style="list-style-type: none"> • Wang Peng feels the room is too small |
| <ul style="list-style-type: none"> • 王朋觉得宿舍不能做饭，很不方便 | <ul style="list-style-type: none"> • Wang Peng feels it is very inconvenient to not be able to cook in the dorm |
| <ul style="list-style-type: none"> • 王朋准备下个学期搬出去 | <ul style="list-style-type: none"> • Wang Peng is prepared to move out next semester |
| <ul style="list-style-type: none"> • 王朋找房子找了一个多月 | <ul style="list-style-type: none"> • Wang Peng has been looking for a place for more than a month. |
| <ul style="list-style-type: none"> • 王朋没有找到合适的 | <ul style="list-style-type: none"> • Wang Peng has not found anything suitable yet. |

When connecting formally incohesive sentences to be a coherent paragraph, L1 and L2 Chinese speakers of varied proficiency levels applied different strategies. Below (83), (84), and (85) are output collected respectively from participants of Group Low, High, and Native. In (83) the low proficiency L2 Chinese speaker added overt conjunctions to connect simple sentences into more cohesive complex and longer sentences instead of unconnected individual redundant complete sentences. Such conjunctions included *kěshì* (可是, but), *yīnwèi* (因為, because), *háiyǒu* (還有, also; in addition), and *suǒyǐ* (所以, so, therefore). However, such manipulation by the Group Low participant remained within each single TC-unit; there was no manipulation made to combine single TC-units into complex terminable units. The subject of each single TC-unit provided was redundantly kept, preventing independent TC-units from being combined to form a complex terminable TC-unit. In (83), there were 8 single TC-units which composed 7 terminable TC-units. Each terminable TC-unit consisted of 1 independent single TC-unit, with the exception of the third terminable TC-unit in (83), which consisted of 2 dependent single TC-units: *Yīnwèi sùshè₃ bùnéng zuò fàn, Ø₃ hěn bù fāngbiàn* (因為宿舍₃不能做飯, Ø₃很不方便。Because one cannot cook in the dorm, it is very inconvenient). However, this 3rd terminable TC-unit was not produced by the participant. It was one of the sentences provided in the GW task.

As Chinese proficiency increases, topic repetition takes more varied forms. Instead of repeating the same topic in its full form all the time, pronouns and correferential zeros come into play and there is higher integration of TC-units. Paragraph (84) was rewritten by a Group High participant who received an EI score of 80.5/120. In (84) there were a total of 5 terminable TC-units, 3 out of which (TTCU-2, 3, 5) consisted

of more than 2 dependent single TC-units (DSTCU) connected by correferential zeros. MLTTCU of paragraph (84) was 14.4 characters. The same topic *Wáng Péng* (王朋, Wang Peng) was repeated twice equally in its full form, the third person pronoun *tā* (他, he), and correferential zeros. However, these two terminable TC-units in (84) were both typical topic chains according to the *Ten types of Chinese topic chains* listed in Appendix A, in which both of the two correferential zeros were taking the position as subjects of the clauses. As the L2 Chinese speakers' language proficiency increased, more varied types of topic chain were included in their output.

Paragraph (85) was produced by a Group High participant. This L2 Chinese speaker received 115/120 in the EI test that showed a native like Chinese language proficiency. In (85), *Wáng Péng* (王朋, Wang Peng) was not repeated in its full form at all. Instead, it was replaced with a pronoun *tā* (他, he) or a correferential zero. There were a total of 5 terminable TC-units that were composed of 14 dependent single TC-units in (85). TTCU-1 and TTCU-3 each consisted of only 1 independent single terminable TC-unit. Both TTCU-2 and TTCU-4 were composed in the form of typical topic chains. In contrast, TTCU-5 was a much more complex telescope chain which was aforementioned as (54) in Section 2.2.3. In this terminable TC-unit, there were two intertwined topics: *tā* (他, he) and *sùshè* (宿舍, dorm). Each led one strand of this complex telescope chain TTCU-5. Topic *tā* (他, he)_I and Topic *sùshè* (宿舍, dorm)_J both appeared in the first dependent single TC-unit. Then Topic *tā* (他, he)_I was repeated in the form of correferential zeros in DSTCU-4 and DSTCU-6, which formed the Topic *I* strand. Topic *sùshè* (宿舍, dorm)_J was repeated in the form of the coreferential zero in

DSTCU-2, DSTCU-3 and DSTCU-5, which formed the Topic *j* strand. As we can see in Figure 9, Topic *tā* (他, he)_I strand and Topic *sùshè* (宿舍, dorm)_j strand were intertwined in one telescope chain and could not be separated. Topic *I* strand takes the form of the typical topic chain, while Topic *j* strand I a mixture of typical topic chain (DSTCU-1 and DSTCU-2), covert double topic chain (DSTCU-1 and DSTCU-3, DSTCU-1 and DSTCU-5). Such a topic intertwined chain structure can be considered one of the most complex terminable TC-units. From a developmental perspective, being able to produce such a complex terminable TC-unit reveals an advanced degree of language proficiency of the speaker.

(83) Group low (EI Score: 39.5/120): *Wáng péng₁ zài xuéxiào de sùshè zhùle liǎng gè xuéqīle. Kěshì wáng péng₂ juéde sùshè tài chǎo, yīnwèi sùshè₃ bùnéng zuò fàn, Ø₃ hěn bù fāngbiàn. Hái yǒu wáng péng₄ juéde fángjiān tài xiǎo, suǒyǐ wáng péng₅ zhǔnbèi xià gè xuéqī bān chūqù. Kěshì wáng péng₆ zhǎo fángzi zhǎole yí gè duō yuè. Wáng péng₇ méiyǒu zhǎodào héshì de.*

王朋₁在學校的宿舍住了兩個學期了//(TTCU-1)。可是王朋₂覺得宿舍太吵//(TTCU-2)，因為宿舍₃不能做飯，Ø₃很不方便//(TTCU-3)。還有王朋₄覺得房間太小//(TTCU-4)，所以王朋₅準備下個學期搬出去//(TTCU-5)。可是王朋₆找房子找了一個多月//(TTCU-6)，。王朋₇沒有找到合適的//(TTCU-7)。

[Wang Peng₁ has lived at a school dorm for two semesters//(TTCU-1). But Wang Peng₂ felt that the dorm too noisy//(TTCU-2). Because dorm₃ could not cook, very inconvenient//(TTCU-3). What's more, Wang Peng₄ felt the room too small//(TTCU-4), so Wang Peng₅ prepared next semester move out//(TTCU-5). But Wang Peng₆

looked for house look LE more than one month//(TTCU-6). Wang Peng₇ did not find suitable//(TTCU-7).]

Wang Peng has lived in the dorm on campus for two months. But Wang Peng felt that the dorm was too noisy. Because one cannot cook in the dorm, he felt it was really inconvenient. Also Wang Peng felt the room was too small, so Wang Peng prepared to move off campus next semester. But Wang Peng has been looking for more than a month; he could not find a suitable place yet.

(84) Group High (EI Score: 80.5/120): *Wáng péng zài xuéxiào de sùshè zhùle liǎng gè xuéqīle. Wáng péng juéde sùshè tài chǎo, yě juéde fángjiān tài xiǎo. Tā juéde sùshè bùnéng zuò fàn, hěn bù fāngbiàn. Wáng péng zhǔnbèi xià gè xuéqī bān chūqù. Tā zhǎo fángzi zhǎole yí gè duō yuè, yě méiyǒu zhǎodào héshì de.*

王朋₁在學校的宿舍住了兩個學期了//(TTCU-1)。王朋₂覺得宿舍太吵， \emptyset_2 也覺得房間太小//(TTCU-2)。他覺得宿舍₃不能做飯， \emptyset_3 很不方便//(TTCU-3)。王朋₄準備下個學期搬出去 //(TTCU-4)。他₅找房子找了一個多月， \emptyset_5 也沒有找到合適的//(TTCU-5)。

[Wang Peng₁ at a school dorm lived for two semesters//(TTCU-1). Wang Peng₂ feels the dorm too noisy, \emptyset_2 also feels the room too small//(TTCU-2). He feels the dorm₃ cannot cook, \emptyset_3 very inconvenient//(TTCU-3). Wang Peng₄ prepared next semester move out//(TTCU-4). He₅ look for a place look LE for more than a month, \emptyset_5 yet not find a fit//(TTCU-5).]

Wang Peng has been living in the dorm on campus for two semesters. Wang Peng felt the dorm was too noisy; the room was also too small. He felt it was very

inconvenient that he could not cook in the dorm. Wang Peng planned to move out next semester. He has been looking for a place for more than a month, and has yet to find a suitable place.

(85) Group High (EI Score: 115/120): *Wáng péng zài xuéxiào de sùshè zhùle liǎng gè xuéqīle. Tā juéde sùshè tài chǎo, fángjiān tài xiǎo, yòu bùnéng zuò fàn. Tài bù fāngbiàn le. Suǒyǐ tā zhǔnbèi xià gè xuéqī bān chūqù. Jiéguǒ, tā zhǎo fángzi zhǎole yí gè duō yuè, kěshì méiyǒu zhǎodào héshì de. Tā jiù juéding hái shì xiǎng bàn fǎ xíguàn sùshè ba! Tài chǎo kěyǐ yòng ěrsāi, fángjiān tài xiǎo méiguānxi, kěyǐ sòng diào yìxiē tā de shū hé yīfu. Méi dìfāng zuò fàn zuì bùchéng wèntí, zài zhōngguó shàng jiē chīfàn kě fāngbiànliǎo!*

王朋₁ 在学校的宿舍住了两个学期了//(TTCU-1)。他觉得宿舍₂ 太吵、 \emptyset_2 房间太小， \emptyset_2 又不能做饭。 \emptyset_2 太不方便了//(TTCU-2)。所以他₃ 准备下个学期搬出去 //(TTCU-3)。结果，他₄ 找房子找了一个多月， \emptyset_4 可是没有找到合适的 //(TTCU-4)。他_{5i} 就决定还是想办法习惯宿舍_{5j} 吧/(DSTCU-1, TTCU-5)! \emptyset_{5j} 太吵可以用耳塞/(DSTCU-2, TTCU-5)， \emptyset_{5j} 房间太小没关系/(DSTCU-3, TTCU-5)， \emptyset_{5i} 可以送掉一些他的书和衣服/(DSTCU-4, TTCU-5)。 \emptyset_{5j} 没地方做饭最不成问题/(DSTCU-5, TTCU-5)， \emptyset_{5i} 在中国上街吃饭可方便了/(DSTCU-6, TTCU-5)!

[Wang Peng₁ at school DE dorm LE two semester//(TTCU-1). He felt dorm₂ too noisy, \emptyset_2 room too small, \emptyset_2 plus cannot cook. \emptyset_2 Too inconvenient//(TTCU-2). So he₃ prepared next semester move out//(TTCU-3). As a result, he₄ looked for room look LE more than one month, \emptyset_4 but no found suitable// (TTCU-4). He_{5i} then decided still think ways to get used to the dorm_{5j}/(DSTCU-1, TTCU-5)! \emptyset_{5j} Too noisy can use ear

buds/(DSTCU-2, TTCU-5), \emptyset_{5j} room too small does not matter/(DSTCU-3, TTCU-5),
 \emptyset_{5i} can give away some of his books and clothes/(DSTCU-4, TTCU-5). \emptyset_{5j} No place
make food most not become problem/(DSTCU-5, TTCU-5), \emptyset_{5i} in China go to the
street eat very convenient// (DSTCU-6, TTCU-5)!]

Wang Peng has been living in a dorm on campus for two semesters. He felt the dorm was too noisy and the room too small. What was worse was that you could not cook in the dorm, which was really inconvenient. Therefore, he had been looking for a place for more than a month. But he had yet to find a suitable place. Consequently, he decided to try to get used to the dorm. If it got too noisy, he could use on ear plugs. If the room got too small, it didn't matter, he could just give away some of his books and clothes. If there was no place to cook, it was the least of his problems since it is so convenient to eat out in China!

As stated in Section 3.3, it was proposed that the ratio of different types of terminable TC-unit be included as one way of estimating global level Chinese syntactic complexity. Scoring patterns by proficiency groups on this measure will likely be of high value in future studies for providing provide more insights on complexity development as well as insights with regards to length and ratio measures.

6.2 Chinese syntactic complexity in proficiency guidelines

In the field of foreign language learning, proficiency guidelines are provided to help define foreign language competence and describe proficiency at different levels of development. The ACTFL Proficiency Guidelines and ILR Scale are currently two of the

most widely applied proficiency guidelines used in the United States. In terms of the four skills of listening, speaking, reading, and writing, such proficiency guidelines provide developmental descriptions of general language use while allowing different details for varied languages. The ACTFL Proficiency Guidelines (2012) “identify five major levels of proficiency: Distinguished, Superior, Advanced, Intermediate, and Novice. The major levels Advanced, Intermediate, and Novice are subdivided into High, Mid, and Low sublevels” (See more at: <http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>). The ILR Scale describes language use in six “base-levels” designated 0, 1, 2, 3, 4, and 5; and in “plus-level” designated 0+, 1+, 2+, 3+, 4+ “when proficiency substantially exceeds one skill level and does not fully meet the criteria for the next level” (see more at: <http://www.govtilr.org/Skills/ILRscale1.htm>).

Although the ACTFL Guidelines and ILR Scale are not completely interchangeable with each other, they nevertheless both include syntactic complexity in terms of structure, pattern, and grammatical relation, etc. in the description for each language proficiency level. For the speaking and writing skill, “simple structures and grammatical relations”, “basic syntactic patterns”, “common structures”, “repetitive structure”, “frequently used syntactic structures”, “recombinations of learned vocabulary and structures”, and “foreign style” are included in referring to lower syntactic complexity levels, while the descriptions for higher syntactic complexity included “complex syntactic patterns”, “a wide range of structures”, “a full range of structures”, “complex high-frequency structures”, “low frequency complex structures”, “a sophisticated control of vocabulary and phrasing”, and “a wide variety of cohesive devices”. In the ILR Scale, examples are provided for complex structures, such as

“tense/aspect usage, case morphology, passive constructions, word order, and embedding (relative clauses)”, and “complex modification and embedding in Indo-European languages”. Examples also cover cohesive devices, “relative constructions, object pronouns, connectors, etc.” as examples of basic cohesive elements of discourse, while “ellipses and parallelisms, and subordinates in a variety of ways” are examples of “a wide variety of cohesive devices”.

Provided with the quantitative and qualitative analysis assessing Chinese syntactic complexity via TC-units, this dissertation further depicts and restructures the outline of Chinese syntactic complexity development based on the three stages proposed by Jin (2006, p. 134-5) (translated as): *threshold*, *growth*, and *leap*. As the unit of analysis, the TC-unit proposed in this dissertation should be included in the description of Chinese syntactic complexity corresponding to different proficiency levels. At the elementary level, L2 Chinese learners lack a comprehensive understanding of Chinese typological features in terms of topic-prominence, parataxis-prominence, and discourse-oriented. Without knowledge of how a terminable TC-unit is composed via correferential zero, learners are able to compose Chinese output in the form of complete sentences with redundant repetition of the same subject in its full form or with pronouns. The single TC-units that elementary level learners produce are mostly short and independent at the threshold stage. Learners start to lengthen single TC-units by adding more attributive and adverbial modifiers, forming more serial verb construction within one intonation contour, adding post-verbal arguments within one intonation contour, and using more sophisticated phrases or structures, etc. Middle level learners continue such practice since lower level to lengthen single TC-units at subclausal level. In addition, overt connectives

play an important role as cohesive devices at this stage. However, intermediate level learners start to combine single TC-units via coreferential zero as well as to apply overt conjunctions to better connect single TC-units in a complex terminable TC-unit. The complex terminable TC-units intermediate level learners produced mainly take the form of typical topic chains. At advanced level, learners are able to produce more complex terminable TC-units out of all the TC-units. Such complex terminable TC-units take the forms of varied topic chains including telescope chains. Such terminable TC-units consist of well knit single TC-units via coreferential zeros and overt conjunctions. Chinese L2 speakers finally get to the topic-prominent stage regardless of the typological difference of their L1.

Therefore, for general foreign language proficiency guidelines, instead of word order which mostly refers to the accuracy of syntactic structures, the example provided for complex structures might also need to include single TC-units, and terminable TC-units in the form of varied types of topic chains, as indicators of even more sophisticated structures. In addition, the example of a variety of cohesive devices can also include covert connectives for languages of parataxis-prominence.

6.3 Pedagogical implications

6.3.1 Complexity developmental stages

As mentioned in the Chapter 1 of Introduction, as an overlooked dimension in the triad of CAF, complexity development calls for long-overdue attention in the field of Chinese second language teaching and research. As mentioned in Section 1.1, the Chinese nationwide *Standards for Mandarin Chinese Proficiency* (Hanban, 1995)

provides specific descriptive and quantifiable requirements for accuracy and fluency with regards to listening, speaking, reading, and writing for each proficiency level, but not for complexity development. Considering the examples in ILR Scale discussed in Section 6.2, a better understanding of what complex structures look like in Chinese language development is essential. With further clarification and confirmation of the conceptualization and operationalization of Chinese syntactic complexity in this dissertation, it is hoped that the complexity dimension at the macro level may be introduced into Chinese language learning and assessment.

In proficiency guidelines that define Chinese language competence and describe proficiency at different levels of development, more operationalized Chinese complexity developmental requirements in the form of a rubric should be included. Such a rubric can then be adapted and incorporated into the pedagogical syllabus in stages for L2 Chinese learning and teaching. The three developmental stages of Chinese syntactic complexity were depicted in terms of the development of TC-unit as proposed in this dissertation (see Section 6.2). Corresponding teaching emphases at different stages therefore should be integrated into Chinese language teaching. At the beginning of Chinese learning, it is very crucial to have teachers' awareness to help learners distinguish Chinese sentential structures from sentences of Indo-European languages. This is especially true for native English speaking L2 Chinese learners. Chinese is a topic-prominent, parataxis-prominent, and discourse-oriented language in contrast to English as a subject-prominent, hypotaxis-prominent, and sentence-oriented language. First is to have learners understand the arbitrary boundary of a Chinese sentence. As discussed in Section 2.2.3, the conventional sentence in Chinese actually enjoys arbitrary punctuational boundaries. Teaching material

at an elementary level can present such an arbitrary boundary of a Chinese sentence in contrast to the English sentence boundary in order to prepare learners to perceive that TC-units instead of a sentence better fit the analysis of Chinese syntactic structure. A terminable TC-unit can be shorter or longer than a punctuational sentence. Learners can be provided exercises to add punctuation marks for a short punctuation free paragraph and compare their punctuation marks with each other as well as the punctuation marks provided in the original text. Second, to help learners perceive Chinese text in the unit of TC-unit, some typical topic-comment structure without verb like (1) *Jīntiān xīngqīwǔ*. (今天星期五。 Today Friday) can be provided for learners to get a better sense of the topic-prominence of Chinese language. To lengthen a single TC-unit, examples can also be provided by adding more attributive and adverbial modifiers, forming more serial verb construction within one intonation contour, adding more complex arguments after a verb, and using more sophisticated phrases or structures, etc. Second, with an understanding of the arbitrary boundary of a Chinese sentence, learners can then be provided with exercises to identify the beginning and end points of a complex terminable TC-unit in the form of a typical topic chain. Learners can then be guided by the teacher to notice how the dependent single TC-units are connected. Therefore, learners can get a better sense of the topic-prominence and discourse-prominence of Chinese language. Examples of L1 transfer such as the overuse of complete sentences, as well as redundant repetition of the same subject can be presented to help learners achieve a more comprehensive understanding regarding the different unit of analysis and syntagma of Chinese from English. The inner-structure of a complex terminable TC-unit can then be introduced. In a complex terminable TC-unit that consists of more than one single TC-unit, such

dependent single TC-units share one topic that is repeated in the form of coreferential zero.

As their proficiency increases, intermediate level learners who have become more aware of the topic-prominence and discourse-prominence of Chinese should be first provided ample exercises to produce typical topic chains by applying coreferential zero. More illustration and analysis of varied types of complex terminable TC-units can be provided. Different types of topic chains categorized by Li (2005) according to the different thematic roles of the omitted topic as well as the different positions that the coreferential zero take can be presented to intermediate level learners. Second, by contrasting the output of speakers of different proficiency levels, the instructor can help learners understand how single TC-unit lengthening and combining is applied at the different stages of Chinese complexity development. Exercises should be provided to help learners lengthen a single TC-unit at subclausal level as well as combine multiple single TC-units into a complex terminable TC-unit. Third, the difference of paratactic and hypotactic syntagma in terms of covert and overt conjunctions can be emphasized at this stage. Teachers bring learners to the awareness that provided sufficient contextual clue, the juxtaposition of dependent single TC-units without explicitly indicating their structural relationships by overt conjunction is considered more cohesive in a meaning-driven language as Chinese. How conciseness is preferred in Chinese while preciseness is preferred in English is contrasted. Comparison of two similar complex terminable TC-units with covert or overt conjunctions, such as (51) and (52) (See Section 2.2.2.4), can be presented for learners to understand the nuances in information density.

As learners' language proficiency stages progress to advanced levels, with single TC-units lengthened, the teaching focus should then be placed at an increased global Chinese syntactic complexity by combining more single TC-units into a complex terminable TC-unit and by producing a higher proportion of such complex terminable TC-units out of all Chinese output. The form of such complex terminable TC-units is not limited to a typical topic chain. With the continuous input of more varied types of topic chains, learners at this stage should be provided with more opportunities to produce varied types of topic chains, including telescope chains. In addition, students are expected to proactively apply more covert conjunctions in combining single TC-units into a complex terminable TC-unit.

6.3.2 Composing Terminable TC-unit

As clarified in Section 2.1.3, the breadth and depth of the syntactic complexity construct, corresponding to declarative and procedural knowledge internalization, refers to the internalization of different types of topic chain structure and how to assemble and restructure such topic chains. From a macro perspective of TC-unit learning, at the early stage of Chinese language learning, a terminable complex TC-unit in the form of a typical topic chain sentence can be introduced to prepare L2 Chinese learners with declarative knowledge of the different syntagmatic mechanisms of Chinese from English. Along the different levels of Chinese language teaching, ten types of topic chains can be introduced gradually based on the pedagogical text: typical topic chain, cataphoric topic chain, patient-agent topic chain, patient-patient topic chain, theme-patient topic chain, preposed topic chain, presented topic chain, montage topic chain, overt double topic

chain, and covert double topic chain (See Appendix A for illustrations). Given examples, even telescope chains can be introduced at a superior level. L2 Chinese learners are then exposed with varied forms of complex terminable TC-units used in varied contexts and genres.

From a micro perspective of the composition of an individual topic chain, Chu (2006) illustrated three steps of topic chain development: *introduction*, *pick-up*, and *continuation*. Two examples were given and analyzed in Chu (2006) to show the development. The complexity measures proposed in this dissertation are here applied below to describe the structures of these topic chains.

(86) *Luòyáng yǒu gè míng gēnǚ, jiào yángníngluó, cōnghuìguò rén, yǐ yǔyán jiān qiǎo guān jí yìshí.*

洛阳有个名歌女，叫杨苧罗，聪慧过人，以语言尖巧冠极一时。

[Luoyang have a famous singing girl_i (DSTCU-1), \emptyset_i call Yang Ningluo (DSTCU-2), \emptyset_i smart than people (DSTCU-3), \emptyset_i by language pungent ingenious champion extreme one time (DSTCU-4).]

There is a famous song girl in Luoyang whose name is Yang Ningluo. She was the most famous in her time for her intelligent and artful use of language.

(87) *Luòyáng yǒu gè míng gēnǚ, hái yǒu gè wǔnǚ, cái mào jīngrén, yě yíyàng yǒumíng.*

洛阳有个名歌女，还有个舞女，才貌惊人，也一样有名。

[Luoyang_i have a famous singing girl (DSTCU-1), \emptyset_i also has a dancing girl_j (DSTCU-2), \emptyset_j talent appearance stunning (DSTCU-3), \emptyset_j also same famous (DSTCU-4).] (p. 203-211)

There is a famous song girl in Luoyang. In addition, there is a dancer who has great talent and stunning beauty. She is equally famous.

The first single TC-unit of (86) is an existential clause, and it is coded as DSTCU-1 here. This DSTCU-1 of (86) includes two potential topics: *Luòyáng* (洛阳, a place name), and *gēnǚ* (歌女, song girl). The following DSTCU-2 picked up *gēnǚ* (歌女, singing girl) as the topic and repeated it in the form of a coreferential zero, which combined DSTCU-1 and DSTCU-2 in the form of a presented topic chain. DSTCU-3 and DSTCU-4 extended this presented topic chain by repeating the same topic via coreferential zero in the form of a typical topic chain. Therefore, DSTCU-1, 2, 3, and 4 form one terminable TC-unit since with one topic, *gēnǚ* (歌女, singing girl), picked up and repeated in the form of coreferential zero.

Example (87) illustrates another way to form a topic chain. The same DSTCU-1 introduces the same two potential topics. Instead of *gēnǚ* (歌女, singing girl), DSTCU-2 in (87) picked *Luòyáng* (洛阳, a place name) as the topic and repeated it in the form of a coreferential zero. DSTCU-1 and 2 were combined to form a typical topic chain. However, DSTCU-3 and 4 do not continuously repeat *Luòyáng* (洛阳, a place name) as their topic. DSTCU-3 picked up the theme in DSTCU-2, *wǔnǚ* (舞女, dancing girl), as the topic and DSTCU-4 repeated this same topic in the form of coreferential zero. With the topic of DSTCU-1 repeated in coreferential zero as well as the full form of the topic of DSTCU-3 and 4 included, DSTCU-2 overarched two parts of a telescope chain. DSTCU-2 is part of the complex terminable TC-unit consists of DSTCU-1 and 2, but also

is part of the complex terminable TC-unit consists of DSTCU-2, 3, and 4. Therefore, DSTCU-1, 2, 3, and 4 in (85) form a terminable TC-unit in the form of a telescope chain.

Setting off with the three main steps of topic chain composing outlined in Chu (2006), *introduction*, *pick-up*, and *continuation*, the ten patterns of topic chains Li (2005) categorized can be applied as more possibilities to elaborate and diversify the composing of dependent single TC-units in each terminable TC-unit. For example, the topic of the terminable TC-unit does not have to be presented in the first dependent single TC-unit as shown in (86) and (87). Instead, the topic can take a position in its full form in the subsequent dependent single TC-unit and form a cataphoric topic chain. The topic can take the thematic role of agent, theme, or patient in each dependent single TC-unit. There can also be two or more strands intertwined in a telescope chain terminable TC-unit. As in the telescope chain (54) introduced in Section 2.2.3, both the agent, *tā* (他, he), and the patient, *sùshè* (宿舍, dorm), introduced in DSTCU-1, were picked up as topics and continues two strands to form a telescope chain in the following 5 continuous DSTCUs. Topic *tā* (他, he) was repeated in the form of coreferential zero in DSTCU-2, 3, and 5, while Topic *sùshè* (宿舍, dorm) was continued in DSTCU-4, and 6.

6.3.3 Teaching TC-unit in class

Before introducing TC-unit in class, it is important to raise learners' awareness of the arbitrariness of Chinese sentence boundary. As discussed in Section 6.3.1, learners can be provided exercises to add punctuation marks for a short punctuation free paragraph and compare their added punctuation marks with other students as well as the with original text. Differentiating the Chinese sentence from the sentence in English

enables learners to perceive and develop Chinese syntactic complexity in the unit of TC-units.

Classroom teaching assignments for complexity development should include both speaking and writing tasks. Such tasks for English-speaking L2 Chinese speakers can have two emphases: composing and contrasting. While conducting such tasks in a classroom teaching context, immediate task repetition with pedagogical interventions corresponding to different concerns of different proficiency groups can facilitate the achievement of a better outcome.

6.3.3.1 Task focusing on composing and contrasting

Considering the syntagmatic difference between Chinese and English, for English-speaking L2 Chinese speakers, classroom teaching design focusing on complexity development can have two emphases: composing and contrasting. Composing tasks help learners learn how to assemble and lengthen terminable TC-units; contrasting tasks provide learners metalinguistics awareness while developing terminable TC-units. The examples below illustrate different types of composing tasks: (i) extending, (ii) sequencing, (iii) combining, (iv) inserting, (v) chain forming; and contrasting tasks: (vi) punctuation marking, (vii) conjunction converting, (viii) translating. The literal English translations provided here in this dissertation will not be included in the actual exercise for L2 Chinese learners.

Following the acquisition order from input, intake, to output, learners can first be exposed to a typical single TC-unit, such as example (1) *Jīntiān xīngqīwǔ*. (今天星期五。 Today Friday) listed in Section 2.2.1.1. The teacher can then lead learners to

conduct a contrastive analysis between such topic-comment structure and its English translation “It’s Friday today”, which with more comparisons introduces learners to a differentiation of the topic-prominence of Chinese from the subject-prominence of English.

With more awareness of the difference between topic-prominence and subject-prominence, a typical topic chain consisting of two or more dependent single TC-units can then be followed. Li (2005, p.178) suggested introducing the typical topic chain as it was used in initiating communicative dialogues with an example (88) used to start an introduction.

(88) *Tājiào wáng zhōng, shì zhōngguó rén.*

他叫王中，是中国人。(p.178)

[He_i be-called Wang Zhong, Ø_i be Chinese.]

His name is Wang Zhong. He is Chinese.

With this typical topic chain presented, composing tasks can then be designed in varied forms. Several are listed below:

(i) Extending.

Exercise (89) uses a typical topic chain as the starter and asks learners to introduce more information about 王中 by adding more dependent single TC-units. The sample answer given in (89b) is an example with two more dependent single TC-units (“在北京出生、长大，现在在美国工作”) added, extending the original chain provided (“他叫王中，是中国人”). The same topic must be repeated in the form of coreferential

zero in order to extend the original topic chain. If complete sentences with overt subject instead of dependent TC-units with referential zero is added by students, the teacher can then guide students to drop the repetitive overt subject. Students can be scaffolded to mark the topic in the topic chain starter provided as well as the coreferential zero in the dependent single TC-units they added. Such practice can raise their meta-linguistic awareness of the syntagma of Chinese in order to prepare them proactively to lengthen topic chains to increase their Chinese syntactic complexity.

(89) *You are introducing your friend 王中 to another friend. Below is the starter of your introduction. Add more information using more dependent single TC-units:*

(89a) 他叫王中，是中国人。

[He_i be-called Wang Zhong, \emptyset_i is Chinese.]

(Sample answer: (89b) 他叫王中，是中国人，在北京出生、长大，现在在美国工作。

[He_i be-called Wang Zhong, \emptyset_i is Chinese, \emptyset_i at Beijing born, grow up, \emptyset_i now at America work.]

(ii) Sequencing.

In (90), a list of disorganized dependent single TC-units that a terminable TC-unit consists of are provided for learners to put back in the right order. Such dependent single TC-units are ready to be sequenced without any revision needed.

(90) *You friend noticed a new friend at the party and asked you about him. You are telling your friend about this new person 王中. Put the single dependent TC-units provided below back into the right order in the form of one terminable TC-unit.*

| | |
|--|--|
| <ul style="list-style-type: none">• 是中国人• 他叫王中• 现在在美国工作• 在北京出生、长大 | <ul style="list-style-type: none">• [is Chinese]• [he be-called Wang Zhong]• [now at America work]• [at Beijing be born, grow up] |
|--|--|

(Sample answer: (90a) 他叫王中，是中国人，在北京出生、长大，现在在美国工作。

(iii) Combining.

Similar with the GR task in TW&ST, combining tasks require learners to put a list of independent single TC-units into a cohesive topic chain. In order to put the independent single TC-units into a topic-chain, the learners have to first identify the shared topic of these independent single TC-units, and then save its complete form in one single TC-unit while repeating it in the form of coreferential zero for the rest. Such exercises can again start with a typical topic chain, to be followed with other types of chains.

(91) *You are telling your friend about 王中. Combine the following sentences provided, so you can connect the information within one terminable TC-unit instead of four independent sentences.*

| | |
|--|---|
| <ul style="list-style-type: none"> • 他叫王中。 • 他是中国人。 • 他在北京出生、长大。 • 他现在在美国工作。 | <ul style="list-style-type: none"> • [He be-called Wang Zhong.] • [He is Chinese.] • [He at Beijing born, grew up.] • [He now at America work.] |
|--|---|

(Sample answer: (91a) 他叫王中，是中国人，在北京出生、长大，现在在美国工作。

(iv) Inserting.

As shown in (92), a single TC-unit can be provided to learners for them to insert into a terminable TC-unit. Such a terminable TC-unit can either be an independent or an otherwise incomplete one. Such inserting tasks require learners not only to be aware of the syntagmic mechanism of the topic chain, but also to match the form along the stream of meaning.

(92) *Your friend is introducing you to another friend 王中. Part of the information is missing. Where should you insert “在北京出生、长大” back to this terminable TC-unit below?*

他叫王中，是中国人，现在在美国工作。

[He be-called Wang Zhong, be Chinese, now at America work.]

(Sample answer: (92a) 他叫王中，是中国人，在北京出生、长大，现在在美国工作。

(v) Chain forming.

Instead of extending a topic chain, learners are expected to be able to compose a complex TC-unit themselves with a single TC-unit provided. Taking Chu's (2006) three steps of topic chain development: *introduction*, *pick-up*, and *continuation*, learners have to first pick up a topic from the single TC-unit provide. As we discussed in 6.3.2, there might be different possibilities of topic picking up, followed with the composition of different types of topic chains. For example, with a single TC-unit provided in (93a), there are two potential topics “我”, or “中文”. If learners pick up “我” and repeat it in the form of coreferential zero as in (93b), a typical topic chain is then composed. If “中文” is instead picked up as the topic and repeated in the form of coreferential zero as in (93c), a cataphoric chain is then composed.

(93) *You are introducing yourself to a new friend. Below is the starter of your self-introduction. Add more information using more dependent single TC-units.*

(93a) 我喜欢学中文, _____

[I like study Chinese, _____]

(Sample answer 1: (93b) 我_i喜欢学中文, ∅_i也喜欢学日文。

[I_i like study Chinese, ∅_i also like study Japanese.]

(Sample answer 2: (93c) 我喜欢学中文_i, 因为∅_i很有意思。

[I like study Chinese_i, because ∅_i very interesting.]

(94) *Describe your room. Start your description with the first dependent TC-unit provided below. Try to elaborate your description using different types of complex TC-units.*

(94a) 我的房间里有一张很大的书桌, _____

[I DE room inside has a CL very big desk, _____]

(Sample answer 1: (94b) 我的房间里 i 有一张很大的书桌, \emptyset_i 还有一张最舒服的床
和一个不大不小的衣柜。

[I DE room inside $_i$ has a CL very big desk, \emptyset_i in addition has a CL most
comfortable bed and a CL not-big-not-small closet.]

(Sample answer 2: (94c) 我的房间里有一张很大的书桌 i , \emptyset_i 上面摆满了沁香的鲜
花 j , \emptyset_j 让人心情舒畅。

[I DE room inside has a CL very big desk $_i$ (DSTTU-1), \emptyset_i on top displayed full
LE fragrant flowers $_j$ (DSTTU-2), \emptyset_j made people mood (DSTTU-3).]

With the same first dependent TC-unit provided, (94b) composed a typical chain, while (94c) is a telescope chain consisting of a montage topic chain and presented topic chain. In (94b), “我的房间” was picked up as the topic and continued in the form of coreferential zero. In contrast, (94c) first picked up “书桌” as the topic and repeated the topic in the form of coreferential zero in DSTTU-2. However, “书桌” was not continuously chosen as the topic of DSTTU-3. Instead, “鲜花” was repeated in the form of a coreferential zero. Therefore, DSTTU-2 and DSTTU-3 formed a presented topic chain. DSTTU-2 as the shared dependent single TC-unit in two types of topic chain functioned as the joint and forged the telescope chain.

A starter of the topic chain can be provided for learners to continue a complex terminable TC-unit. However, learners can also be required to complete a complex TC-

unit with the its starter omitted. Below is an example asking learners to complete the omitted starter of a comple terminable TC-unit.

(95) *You are leaving your friend a voice message about your plans after your classes are over today. Fill in the blank to compose a Cataphoric topic chain.*

(95a) _____, 我准备去超市逛逛。

(Sample answer: (95b) \emptyset_i 上完了中文课, 我 i 准备去超市逛逛。

[\emptyset_i Take finishe LE Chinese class, I $_i$ prepare to supermarket stroll-stroll]

In addition to the typical topic chain which is most frequently used, other types of topic chain should also be included in such exercises on chain forming as learners Chinese proficiency increase.

In addition to these composing tasks as discussed: (i) extending, (ii) sequencing, (iii) combining, (iv) inserting, and (v) chain forming; the other task category for classroom TC-unit teaching is contrasting tasks: (vi) punctuation marking, (vii) conjunction converting, and (viii) translating. Teachers can conduct such tasks independently or as a subline along the composing tasks.

(vi) Punctuation marking.

As suggested above, learners can be asked to add punctuation to a short punctuation free paragraph. As a meaning-driven language, punctuation marks in Chinese can be chosen more subjectively according to the author's intended length of the pause, without being necessarily form-consistent as they are in form-driven languages like English. Learners can then compare their added punctuation marks with other students'

marks and against the original text. Teachers can then lead a discussion covering the arbitrariness of Chinese sentence boundary.

(96) Below is a Chinese paragraph with the original punctuation marks excluded. Add appropriate punctuation marks yourself.

(96a) 王朋在学校的宿舍住了两个学期了觉得宿舍太吵房间太小又不能做饭很不方便他准备下个学期搬出去可是找了一个多月还没有找到合适的房子

(Sample answer 1: (96b) 王朋在学校的宿舍住了两个学期了。觉得宿舍太吵，房间太小，又不能做饭，很不方便。他准备下个学期搬出去，可是找了一个多月，还没有找到合适的房子。

(Sample answer 2: (96c) 王朋在学校的宿舍住了两个学期了，觉得宿舍太吵房间太小又不能做饭，很不方便。他准备下个学期搬出去，可是找了一个多月还没有找到合适的房子。

Compare your punctuation marks with your classmates' and the original text.

With your partner consider the most correct version of this assignment or whether many versions are acceptable.

Different punctuated written formats of a complex terminable TC-unit can also be provided to learners for contrastive analysis. Such exercises of punctuating complex terminable TC-unit can be conducted as a subline along the single TC-unit sequencing exercise like (90). In exercise (97) below, both punctuation markings in (97b) and (97c) are acceptable, which shows the arbitrary of sentence boundaries in Chinese. Teachers can apply such punctuation marking tasks to raise learners' awareness that the length of a

terminable TC-unit does not necessarily align with the sentence boundary, which is arbitrary. With punctuation marks formalizing the sentence boundary, a terminable TC-unit can be part of a sentence, a complete sentence, or can go beyond a sentence.

Teachers can then guide learners to identify the beginning and end points of a terminable TC-unit by coreferential zero. Such contrast analysis help enhance learners' understanding of Chinese as a topic-prominent language, differentiating it from English as a sentence-prominent language.

(97) Below is a complex terminable TC-unit with the original punctuation marks excluded. Add appropriate punctuation marks.

(97a) 他叫王中是中国人在北京出生长大现在在美国工作。

(Sample answer 1: (97b) 他叫王中，是中国人，在北京出生、长大，现在在美国工作。

(Sample answer 2: (97c) 他叫王中，是中国人。在北京出生、长大。现在在美国工作。

Compare your punctuation marks with your classmates' and the original text.

Discuss with your partner whether one complex terminable TC-unit has to be a single sentence or can it also be two sentences according to the punctuation marks.

(vii) Conjunction coverting.

When combining single independent TC-units into a complex TC-unit, in addition to raising learners' awareness of the topic-prominence realized as coreferential zero, contrasting tasks can also pin on the preference of covert conjunctions in TC-unit combining.

(98) You are introducing 王中 to another friend. Include all the information provided in the following four sentences coherently into one terminable TC-unit. Pay attention to the use of correferential zero and conjunctions.

| | |
|---|---|
| <ul style="list-style-type: none"> • 这位是王中。 • 因为他喜欢中国文化。 • 所以他在中国住了五年。 • 所以他的中文说得特别棒。 | <ul style="list-style-type: none"> • This CL is Wang Zhong. • Because he likes speaking Chinese. • So he at China live LE five years. • So he DE Chinese speak DE especially awesome. |
|---|---|

(Sample answer 1: (98a) 这位是王中，因为喜欢中国文化，所以在中国住了五年，所以中文说得特别棒。

(Sample answer 2: (98b) 这位是王中，因为喜欢中国文化，所以在中国住了五年，中文说得特别棒。

(Sample answer 3: (98c) 这位是王中 (DSTCU-1)，喜欢中国文化 (DSTCU-2)，在中国住了五年 (DSTCU-3)，中文说得特别棒 (DSTCU-4)。

Compare the three provided possible answers. Discuss with your partner which one shows the highest coherence in Chinese and why.

Possible answers like (98a), (98b), and (98c) can be listed for contrastive analysis. (98a) represents the redundant use of conjunctions, with all three conjunctions, “因为”, “所以”, and “所以” in the provided sentences kept in their original form. In (98b), the second “所以” was replaced by its covert form. Furthermore, the preference of pair conjunctions used in Chinese is realized by using “因为……所以……” in (98b).

Possible answer (98c) connected all the single TC-units via covert conjunction. With the four dependent single TC-units presented in the form of juxtaposition, (98c) realized higher information density in terms by encoding all the information within one terminable TC-unit. (98c) formed a telescope chain jointing a presented topic chain (DSTCU-1, 2), a typical topic chain (DSTCU-2, 3), and a covert double topic chain (DSTCU-3, 4).

(viii) Translating.

Translation tasks can be applied as a comprehensive contrastive exercise covering the contrast of topic-prominence versus sentence-prominence, parataxis prominence versus hypotaxis-prominence, and sentence-oriented versus discourse-oriented. Such tasks can include translation at the sentence or paragraph level, translation from English to Chinese or vice versa, and comparison of different versions of translation. In other words, it can be translation from one Chinese complex terminable TC-unit to English sentence(s), or vice versa; or translation between a series of Chinese terminable TC-units and English sentences, or vice versa. In addition, such translation can be either inappropriate or very well-acknowledged version. Learners can be asked to actually translate the text provided before analyzing, or both the original text and the translated version can be provided for learners to perform contrastive analysis.

Exercise (99) below is a contrasting task with two possible Chinese translations of one English sentence. The Chinese translation (99b) is perfectly grammatical. However, it literally followed the word order of the English sentence (99a), the subject “那幅画” with a VP modifier therefore sounds overlong and straggling. In contrast, (99c) broke the

English word order in (99a) and used a catatrophic topic chain in translation. Such a catatrophic topic chain linked two actions, laying eyes on the painting in DSTCU-1 “ \emptyset_i 看到墙上的那幅画” and recalling my childhood in DSTCU-2 “我_i总会想起自己的童年”, in a dynamic and smooth rhythm. Teachers can then lead the learners to differentiate and appreciate such nuance difference in different translations.

(99) Below (99b) and (99c) are two Chinese translations for the same English sentence (99a). Pick one that better fits the typological features of Chinese and support your pick by contrasting the different versions of translation.

(99a) *The painting hanging there always reminds me of my days in Hawaii.*

(99b) 挂在墙上的那幅画总是让我想起自己的童年。

(99c) 看到墙上的那幅画，我总会想起自己的童年。

As learners' Chinese proficiency advance, sentences with a more complex structure can be included in translating tasks. Instead of providing translations, exercise (100) below asks learners to translate the English sentence into Chinese, which provides an opportunity for learners to comprehensively apply their meta-linguistic knowledge of the Chinese-English syntagmatic mechanism difference.

(100) *Translate the following English sentence (100a) into Chinese. Then compare the difference between the English and your Chinese translation in terms of syntactic structure.*

(100a) “Quite elaborate ceramics points to a long tradition of pottery making whose origins are still to be found.”

(Sample answer 1: (100b) (中美洲) 相当精致的制陶技术表明, 该地陶器制造有悠久的传统, 虽然尚没发现其源自何处。(Wu, 1995)

(Sample answer 2: (100c) (中美洲) 相当精致的陶艺作品, 显示出该地悠久的陶器制造传统, 待人追溯其源头。

With ample practice at the sentence level, such translation exercises can then be given at paragraph level. (101a) below is an excerpt from a Chinese novel *Wéichéng* (围城, *Fortress Besieged*) written by a reputable Chinese scholar and writer Zhongshu Qian (1947). (101b) is (101a)'s corresponding English translation by Kelly and Mao (1980). Teachers can then lead learners to observe TTCU-3 in (101a), a telescope chain consisting of two connected typical topic chain. While such a telescope chain structure functioned as a camera, the scene described in (101a) smoothly advanced from one character to another. Though (101b) tried best to keep the original punctuation of (101a), it only managed to stay at clause level; the topic-prominence was not transferred in English. As one topic chain, TTCU-3 in (101a) was translated into three sentences in (101b).

(101) *Below (101a) is an excerpt from a Chinese novel Wéichéng (围城, Fortress Besieged). (101a) is the corresponding English translation of (101a) published.*

Compare TTCU-3 in (101a) with Sentence 2-4 in (101b) and discuss on how Chinese topic chain structure functions differently from the English sentence in terms of narrating the successive actions.

(101a) (Original Chinese excerpt: 鸿渐一路上赔小心 //(TTCU-1), 鲍小姐只无精打采 //(TTCU-2)。送她回舱后, 鸿渐也睡了两个钟点。一起身就去鲍小姐舱外弹壁唤她名字, 问她好了没有, 想不到门帘开处, 苏小姐出来, 说鲍小姐病了, 吐过两次, 刚睡着呢 //(TTCU-3)。鸿渐又羞又窘, 敷衍一句, 急忙跳走 (TTCU-4)。

(101b) (Translation: All the way back he was very apologetic, but she remained in low spirits //(Sentence-1). After seeing her to her cabin, he slept for two hours himself //(Sentence-2). As soon as he got up he went to her cabin, tapped on the partition, and called her name, asking if she felt any better //(Sentence-3). To his surprise, the curtain opened and Miss Su came out saying Miss Pao was sick, had thrown up twice, and had just fallen asleep //(Sentence-4). He was at once chagrined and embarrassed; he said something lamely and beat a hasty retreat (Sentence-5). (Kelly & Mao, 1980)

Overall, the above illustrated eight types of tasks can take a form of combination. Combinations of composing tasks and contrasting tasks are easily conducted, such as adding (viii) translating after (ii) sequencing. A contrasting step added onto the composing task can further enhance a learner's understanding of the different typological features between Chinese and English. Combining different types of composing tasks adds variety as well. For example, a combination of (ii) sequencing and (iii) combining will provide learners independent single TC-units in a disorganized format. While trying to identify the topic and connecting the sentences provided, learners additionally have to comprehend and sequence the independent single TC-units to make sure they are picking up the right topic. Another example is to combine (iv) inserting and (v) chain forming. In addition to providing the first single TC-unit for learners to add more dependent single

TC-unit to, such tasks will also provide another single TC-unit and require learners to include this one single TC-unit into the complex terminable TC-unit they compose.

Speaking tasks for topic chain composing

Such topic chain composing exercises can not only be conducted in written modality, but spoken as well. For example, in (i) extending and (v) chain forming exercises, learners can be given an aural cue and asked to extend the first complex TC-unit or form the topic chain orally. Such an aural cue can be the first sentence of a topic or a story, which can be given only or accompanied with other cues such as a picture series to help learners develop their TTCUs.

The CS task in TW&ST can be revised into such a speaking exercise as illustrated below in Exercise (102).

(102) *Narrate a complete story based on the following cartoon strip. Start your story with the first sentence as: “晚饭时间，妈妈在厨房做好了饭。”*

(Sample answer in spoken modality : (102a) “晚饭时间，妈妈在厨房做好了饭。这时候，孩子却不在厨房。孩子在哪儿呢？妈妈让爸爸赶紧去教孩子来吃饭。……”

6.3.3.2 Immediate task repetition for different proficiency levels

When conduct such tasks, immediate task repetition with pedagogical inventions in between two-time enactments can be applied. Take the translation tasks as example, after learners finish the first draft, different forms of pedagogical invention can be issued. Teachers can advise learners to exchange and critique each other's translation. Teachers

can also take one student's translation as the example and review it together with the class, or some key patterns can be provided for students to better structure their translation. Afterwards, an immediate second time translation can provide students to produce a better translation in a short period of class time. Such immediate task repetition is as well, if not more, applicable for speaking tasks in class.

As discussed in Section 5.2.4.2, given the participants' positive self-perception as well as well as different concerns depending on their levels, the immediate task repetition can be modified to include pedagogical interventions to maximize L2 Chinese learners' language complexity. Corresponding scaffolding between twice enactments of the same task can be applied for Chinese language courses at different levels. For lower proficiency learners who are more concerned about retrieving linguistic items, greater accessibility to vocabulary and patterns can be provided before the second retelling. A mini vocabulary and patterns lesson or a search session can be included to achieve better task repetition. If the lower proficiency L2 speakers are allowed to take notes while watching the video, this additional step should be helpful as well. For Chinese learners of higher proficiency, their concern was less on linguistic aspects including language complexity. To help higher level Chinese L2 speakers get better at task planning to retell the narration, allowing a second viewing of the video before the second retelling can help them include more content from the video. The retelling will then be more complete, including more details and complex language within the time limit. To guide higher proficiency level Chinese learners to achieve more native like performance, more focus can be directed to language quality in terms of story organization, fluency, language genre, and vividness. For example, comparison of different versions of story telling can

be provided to increase awareness of the higher proficiency level Chinese learner's to more sophisticated language use.

6.3.4 Interim summary

As analyzed in Section 5.2.3, this dissertation applied four of the proposed Chinese complexity measures as predictors and generated satisfying accuracy at group membership classification according to proficiency differences. These four Chinese complexity measures consisted of two global complexity measures ① MLTTCU and ② CTTCU/ATTCU, as well as two clausal level complexity measures ③ MLSTCU and ④ STCU/TTCU. For spoken Chinese, measure ① MLTTCU was confirmed in this dissertation as the most predictive measure, and therefore the most efficient, for Chinese syntactic complexity assessment. For written Chinese, as single TC-unit combining play a more important role in syntactic complexity making compared with spoken Chinese as proficiency increases, applying both ① MLTTCU and ④ STCU/TTCU as predictor variables generated more accurate group membership classification. Observing both the spoken and written output, this dissertation seemed to confirm Jin's (2006) three stages of Chinese complexity development, *threshold*, *growth*, and *leap*. The four TC-unit-based measures observed an outline of Chinese syntactic complexity development in terms of the length and inner-structure of the terminable TC-units. Generally, at the stage of lower syntactic complexity, shorter simple terminable TC-units consisting of only one independent single TC-unit or shorter complex terminable TC-units consisting of less dependent single TC-units are produced. In order to achieve higher Chinese syntactic complexity, longer terminable TC-units consisting of more dependent single TC-units are

to be produced. Additional qualitative individual case analyses proved that while lower proficiency speakers rely more on lengthening single TC-units, applying paratactic syntagma, and forming typical topic chains to produce more syntactically complex Chinese, higher proficiency Chinese speakers preferred combining single TC-units, applying hypotactic syntagma, and forming varied types of other topic chain structures.

Another factor affecting language complexity, cognitive task complexity, was tested as one manipulable factor of task design, along the lines of both resource-directing and resource-dispersing. Higher Chinese syntactic complexity was produced in more cognitively complex tasks along the line of resource-directing variables in terms of more involved element amounts, longer required performance time, and higher demands on cognitive reasoning. On the other hand, no statistically significant difference was found in the immediate task repetition, which was presumed to be less cognitively complex along the line of resource-dispersing in terms of longer and more in-depth task planning. However, participants' retrospective surveys showed overall positive perceptions regarding their own performance and language strategies in the repetition condition. In addition, different concerns of different groups further revealed the different major challenges for speakers of different proficiency levels, which shed light upon language teaching for different L2 Chinese proficiency levels.

With a clearer picture of *what* Chinese syntactic complexity is and *how* it is developed provided, pedagogical implications were provided from both a macro as well as a micro perspective for developing Chinese syntactic complexity in Chinese as a second language teaching and learning. Corresponding teaching emphases should be integrated into Chinese language teaching alongside the three Chinese complexity

development stages. Both individual topic chain composing steps as well as classroom teaching task designs were illustrated. With regards to composing and contrasting, this dissertation provided some samples of classroom teaching task designs: (i) extending, (ii) sequencing, (iii) combining, (iv) inserting, (v) chain forming, (vi) punctuation marking, (vii) conjunction converting, and (viii) translating. Such task design can take forms of both speaking and writing modality. While conduct such tasks in a classroom teaching context, immediate task repetition with pedagogical inventions corresponding to different concerns of different proficiency groups were suggested to push out better outcome.

CHAPTER 7

CONCLUSIONS

7.1 Significance of the present study

7.1.1 Findings of this dissertation

Conceptualization of Chinese syntactic complexity

As discussed above, the current complexity measures are mostly developed in a global fashion with little attention being paid to tailoring typological differences. Such measures have proved not to be as valid for Chinese complexity measures as they are for Indo-European languages (Jin, 2006; Yuan, 2009). Departing from such indiscriminately developed but not as applicable measures, this dissertation traced the global conceptualization of the multifaceted construct of complexity as adopted from Bulté and Housen (2012, p. 22). In doing so it localized their conceptualization in order to define Chinese syntactic complexity as the number and the nature of the single TC-units that a terminable TC-unit consists of, as well as the number and the nature of the relationships between single TC-units. Scrutinizing Chinese syntactic complexity via a contrastive perspective and without looking through the lens of Indo-European language studies, this dissertation reviewed the topic-prominence, parataxis-prominence, and sentence-oriented nature of the Chinese language, in contrast to the subject-prominence, hypotaxis-prominence, and discourse-oriented nature of the English language. By bringing the insight of typological difference to bear when viewing the syntagmatic mechanism of languages, this dissertation reformed the layout of clause complexing which originally subsumed only coordination and subordination and instead, proposed a taxonomy of

clause complexing (as shown in Figure 4 in Section 2.2.2.3) by giving consideration to the different clause combining mechanisms for topic-prominent languages like Chinese. In an updated taxonomy such as this, the clause combining mechanisms include, but are not limited to: subordination, coordination, and the topic chain. Other possible forms of clause combining may be included according to the typological differences existing in other languages.

Operationalization of Chinese syntactic complexity

By including the *Terminal Topic-Comment Unit* and the *empty category* as indices for Chinese syntactic complexity assessment, Jin's (2006) proposal was pointed in a promising direction. However, as previously discussed in Section 3.3, many questions remained unaddressed in order to propose an operationalizable unit of analysis that clarifies the identification of a topic, as well as the inner-structure and demarcation of a topic chain. First, there was no comprehensible operationalization on how to identify the topic or the beginning and end points of a topic chain. The boundary of a topic chain was confused with the arbitrary sentence boundary labeled by punctuation marks. Relying on punctuation marks for topic chain segmentation not only resulted in subjective coding for spoken Chinese output, but also risked the reliability of the written Chinese syntactic complexity analysis. Of the very few Chinese syntactic complexity studies currently available, Jin's (2006) and Jiang's (2013) were on written Chinese and bypassed the problem of the arbitrary nature of Chinese sentence boundaries. Yuan (2009) followed Chu (1998)'s definition to segment clauses: "minimally consisting of a predicate of various forms" (p. 354), yet sentence level segmentation was not specified.

Second, neither the inner-structure of a topic chain nor the relationship between topic chains and other non-topic-chain output were clarified which prevented research from fully applying topic chain as the unit of analysis for Chinese syntactic complexity. Third, rather than conducting one-at-a-time t-tests between each L2 Chinese speaker group and the native Chinese speaker group on each Chinese syntactic measures, what was needed instead were more powerful quantitative analyses on data elicited from tasks of varied complexity in order to provide a more comprehensive picture for the multi-faceted construct of syntactic complexity.

This dissertation for the first time proposed a taxonomy of TC-units as illustrated in Figure 6 (see Section 3.3). A *terminable TC-Unit* can be categorized as a *simple terminable TC-Unit* or a *complex terminable TC-Unit*. A simple terminable TC-Unit consists of one independent *single TC-unit* in the form of a topic-comment structure. For the independent single TC-unit, its topic is not repeated in the preceding or subsequent topic-comment structure. A complex terminable TC-Unit consists of more than one dependent single TC-unit which takes the form of a topic chain. In a complex terminable TC-Unit, the dependent single TC-units share the same topic which takes its full form once and is repeated in the form(s) of *coreferential zero(s)*. A *coreferential zero* refers to an element that does not have any phonological content and is unpronounced but corefers to the topic mentioned in preceding or subsequent clause(s). Coreferential zero is used to identify the beginning and end points of a terminable TC-unit. If several successive topic-comment structures share the same topic however the topic is not repeated in the form of coreferential zero but the full forms, a pronoun, or a demonstrative, it is then not a topic chain but several sequenced independent topic-comment structures. Whenever repetition

of the same topic takes the form of its full form, a pronoun, or a demonstrative, a new topic-comment structure or topic chain is then activated. Whenever a different topic is introduced in its full form, a pronoun, or a demonstrative, or anaphorically repeated in coreferential zero, a new topic-comment structure or a topic chain is as well activated.

By taking an organic and sustainable approach, this dissertation proposed a series of TC-unit based measures to comprehensively measure Chinese syntactic complexity at varied levels (as listed in Table 4, Section 3.3). Global level Chinese syntactic complexity measures include but are not limited to: the mean length of terminable TC-unit (MLTTCU), complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU), and the ratio of different types of terminable TC-unit. Clausal level complexity Chinese syntactic measures include but are not limited to: the mean length of single TC-unit (both independent and dependent) (MLSTCU), and the single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU). Phrasal level Chinese syntactic complexity measure can be dependents per head. The frequency of a specific form, such as the frequency of a unique topic-comment structure, can also be listed as a measure for Chinese syntactic complexity measure.

Validation of the TC-unit based Chinese syntactic complexity measures

To validate the proposed measures, this dissertation designed a *Chinese Timed Writing and Speaking Test* (TW&ST) to enable and standardize the written and spoken output elicitation from L1 and L2 Chinese speaker participants. Two versions of the TW&ST were designed and utilized: TW&ST (English instructions) for L2 Chinese

speaker participants and TW&ST (Chinese instructions) for L1 Chinese speaker participants. It generally took about 45 minutes to complete *TW&ST (English instructions)* and about 35 minutes for *TW&ST (Chinese instructions)*. TW&ST (English instructions) for L2 Chinese speaker participants consisted of nine sequential parts: (a) a background information survey; (b) a preparation session; (c) a comic strip description task (CS); (d) a video story retelling task (V1); (e) an immediately repeated video story retelling (V2); (f) a retrospective survey; (g) a free writing task (FW); (h) a guided re-writing task (GR); and (i) a Mandarin elicited imitation (EI) test (Zhou & Wu, 2009). In TW&ST (Chinese instructions), all the guidelines, instructions, and questions were translated into Chinese, different background information survey was included, and the Mandarin EI test was excluded.

Four of the TC-unit based measures proposed in this dissertation were applied in order to code and score the elicited complete spoken data sets ($N=115$) and complete written data sets ($N=116$). These four measures were: ① mean length of terminable TC-unit (MLTTCU), ② complex terminable TC-unit/all the terminable TC-units (both simple and complex) (CTTCU/ATTCU), ③ mean length of single TC-unit (both independent and dependent) (MLSTCU), and ④ single TC-units (independent or dependent) per terminable TC-unit (both simple and complex) (STCU/TTCU).

To further investigate and confirm the validity of the proposed measures, this dissertation then conducted discriminant function analyses by task correlating the participants' Chinese proficiency level with their syntactic complexity level. The four proposed Chinese syntactic complexity measures were confirmed with high efficiency with a satisfying accuracy (61.2%~76.5%) at group membership classification. For

spoken Chinese, the measure MLTTCU only was confirmed with an approximate accuracy (67.8%~77.4%) at group membership classification. The measure MLTTCU only can be utilized as one of the most efficient measures for Chinese syntactic complexity assessment. For the speaking task, CS, V1 and V2, respectively, employing only complexity measure MLTTCU correctly predicted 78 (or 67.8%) (Wilks' lambda = .459, $\chi^2(2, N=115) = 87.312, p = .000$), 85 (or 73.9%) (Wilks' lambda = .332, $\chi^2(2, N=115) = 123.369, p = .000$), and 89 (or 77.4%) (Wilks' lambda = .370, $\chi^2(2, N=115) = 111.409, p = .000$) of the total cases' group membership. For written Chinese, applying both MLTTCU and STCU/TTCU measures as predictor variables generated an approximately accurate group membership classification (62.9%~68.1%) since single TC-unit combining plays a more important role in syntactic complexing compared with spoken Chinese as proficiency increases. For the FW and GR writing tasks, it can be seen that applying both MLTTCU and STCU/TTCU measures as group membership predictors reaches higher accuracy through SPSS CLASSIFY. A total of 79 (or 68.1%) (Wilks' lambda = .434, $\chi^2(4, N=116) = 93.933, p = .000$) and 73 (or 62.9%) (Wilks' lambda = .554, $\chi^2(4, N=116) = 66.354, p = .000$) of cases were correctly classified for the FW and GR tasks respectively.

Development of Chinese syntactic complexity

Connect the quantitative against qualitative analyses on Chinese syntactic complexity analyses, this dissertation suggested to confirm the three stages of Chinese complexity development as outlined and described by Jin (2006): *threshold*, *growth*, and *leap*. Instead of describing the three stages via coordination, subordination, and topic chain,

this dissertation consistently depicted the three developmental stages via the length and compositionality of TC-units in learners' spoken and written output. In both speaking and writing tasks, participants of lower Chinese proficiency produced shorter terminable TC-units consisting of fewer dependent single TC-units, whereas participants of higher Chinese proficiency produced longer terminable TC-units consisting of more dependent single TC-units. At a lower proficiency level, terminable TC-units mostly took the form of simple terminable TC-units or short complex terminable TC-units in the form of the typical topic chain. At a higher proficiency level, terminable TC-units more oftenly took the form of complex terminable TC-units. Such complex terminable TC-units were composed of varied types of topic chains including telescope chains. In addition, it was also found that with the output of low proficiency English speaking L2 Chinese speakers at one end and the output L1 Chinese speakers at the other end, the Chinese syntactic complexity development along proficiency increase displayed a transitional development from more reliance on lengthening each single TC-units to combining more single TC-units. Lower proficiency English speaking L2 Chinese speakers relied more on lengthening single TC-units in order to produce more complex Chinese output. By contrast, higher proficiency English speaking L2 Chinese speakers and L1 Chinese speakers relied more on combining single TC-units in order to produce more complex Chinese output. Such transitional development was confirmed by the descriptive statistics in terms of the mean score gaps on the four measures between different proficiency groups (See Section 5.2.1). Take the FW task as an example, the mean scores on measure ① MLTTCU were evenly distributed between three proficiency groups, but on measure ③ MLSTCU, there occurred a drop-down for Group Native compared along the proficiency increase. However, such drop-

down on measure ③ MLSTCU was compensated for by the increasing amount of dependent single TC-units being combined as the extended gaps between Group Native and Group High on the ratio measure ④ STCU/TTCU indicated, as well as by the increasing amount of complex terminable TC-units as the extended gaps between Group Native and Group High on the ratio measure ② CTTCU/ATTCU indicated. Though the mean length of single TC-units of Group Native might be even shorter than those of Group High, more combining of single TC-units contributed to the overall longer terminable TC-units for Group Native compared with Group High. The transitional development was also confirmed by the different patterns of correlation at global and clausal complexity level between participants' Chinese proficiency scores respectively on the length and ratio syntactic complexity measures (See Section 5.2.2). At the global level, length measure ① MLTTCU overall displayed higher correlations than ratio ② CTTCU/ATTCU; while at the clausal level by contrast, ratio measure ④ STCU/TTCU overall displayed higher correlations than length measure ③ MLSTCU.

Interpretation of Chinese complex structures in proficiency guidelines

Without a comprehensive understanding of the construct of Chinese syntactic complexity, the requirement of *sophisticated structures* on proficiency guidelines was not provided with an operationalized definition or was deviated to word order which mostly refers to the accuracy. Since the four TC-unit based measures were confirmed with a high efficiency (61.2%~76.5%) at proficiency group membership classification, this dissertation proposed to apply the taxonomy of TC-units in Chinese syntactic complexity analysis. *Sophisticated structures* in Chinese language can be interpreted as longer

terminable TC-units which consist of more dependent single TC-units in the forms of more varied types of topic chains including telescope chains.

Pedagogical implications for Chinese syntactic complexity developing

This dissertation calls for an official introduction of the complexity dimension into the current, widely-adopted accuracy-fluency dyad pedagogical model in order to apply the triad dimensions of CAF (complexity-accuracy-fluency) in Chinese second language teaching and learning. It can begin with including the dimension of Chinese syntactic complexity into Chinese language proficiency guidelines and the development of assessment rubrics that can then be applied to guide textbook compilation and classroom teaching design. TC-unit proposed in this dissertation can be potentially applied as a pedagogical unit instead of punctuational sentence in Chinese language learning and teaching.

To foster both declarative and procedural knowledge internalization, this dissertation calls for an introduction of the ten types topic chain classified in Li (2005) to L2 Chinese learners. With declarative knowledge of varied types of topic chain, the internalization of proceduralized knowledge of assembling and restructuring such topic chains can be enabled via two emphases: topic chain composing and contrasting. This dissertation provided a series of classroom teaching task design for complexity development: (i) extending, (ii) sequencing, (iii) combining, (iv) inserting, (v) chain forming, (vi) punctuation marking, (vii) conjunction converting, and (viii) translating. These eight types of tasks can be conducted separately or in the form of combination, not only in the written but in spoken format as well. Additionally, this dissertation proposed

immediate task repetition levels as a form of task planning in the classroom teaching context. Such immediate task repetition can be applied to lower L2 learners' communication anxiety and improve their self-perceived performance, as a result increase their willingness to communicate inside or outside the classroom. Between the two enactments of a task, this dissertation suggested incorporation of different pedagogical inventions corresponding to the concern of learners of different proficiency levels: greater accessibility to vocabulary and patterns for lower proficiency learners who are more concerned about retrieving linguistic items, more content information and focus on language organization, fluency, genre, and lexical choice for higher level Chinese L2 speakers.

7.1.2 Methodological implication

Employing the notion of *GlobaLocality*, this dissertation mixed a top-down approach, which starts from clarifying the theoretical construct of syntactic complexity, with a bottom-up approach, which takes a particular feature of the investigated language Chinese. While detecting the construct of Chinese syntactic complexity, an organic approach as advocated by Norris and Ortega (2009), was used to investigate Chinese syntactic complexity via varied levels subsumed: global complexity, clausal complexity and subclausal complexity. As pointed out by Norris and Ortega (2009), when investigating multilayer, multifaceted, and dynamic constructs in second language acquisition, one major threat to validity that occurs during behavior identification is construct underrepresentation. Such construct underrepresentation takes place when the complex links between a theoretical interpretation and its required behavioral evidence

are inadequately understood and/or incorrectly conveyed into practice. In order to warrant such adequate links between a theoretical interpretation and its required behavioral evidence and conduct research organically, in addition to the quantity of such links in this case, takes the form of varied levels of syntactic complexity, this dissertation also found that the quality of such links requires careful deliberation such as the differences grounded in the behavioral evidence of different nature. To warrant such a quality of the links, this dissertation suggests taking into account the different language typological features for defining and measuring Chinese syntactic complexity.

The contrast between Egyptian and Greek architectures illustrates the importance of taking an organic approach in research. For the geometric pyramid which is imposingly constructed on the ground flattened out, its view from different directions can remain rigidly similar. In contrast, some architectures are organically adapted to their geographic surroundings in terms of multiple plateau and terraces along the slope of mountain. In the archaeological area of Delphi, in Ancient Greece, spectators are invited to multifaceted views from different perspectives or different standpoints from the buildings and sculptures. Via one single perspective, one might perceive a full picture of a pyramid, but certainly not for the archaeological area of Delphi where the sculptures were even consciously designed for views from different standpoints. While language learning is certainly not one-dimensional, an organic approach is therefore crucial to obtaining a comprehensive and in-depth picture of multilayer, multifaceted, and dynamic constructs in second language acquisition.

A *GlobaLocality* notion practiced in an organic approach is not only applicable to the CAF investigation, but should be promoted for any dynamic, multilevel, and

multifaceted construct study in SLA. Construct underrepresentation between the theoretical construct and its required behavioral evidence in the form of quantitatively inadequate and/or qualitatively inorganic links also exist in research of other dynamic, multilevel, and multifaceted constructs in SLA.

Practicing a *GlobaLocality* notion through an organic approach brings about unavoidable challenges. From the top-down perspective, misconducting commonly takes the form of construct underrepresentation. As discussed, during the measurement process as illustrated in Figure 2, the links between a theoretical interpretation and its required behavioral evidence can be only partially or selectively studied. There can be a mismatch between the conceptual and operational part of the measurement process as categorized and shown in Figure 2. When only part of the multidimensional theoretic construct is operationalized, caution should be used to avoid generalizing the claim or result for comprehensive application based on a selective part of the construct. The conclusions of the study should be carefully defined but not overgeneralized. For example, in the studies of Chinese syntactic complexity, as discussed in Section 3.1, Jiang's (2013) proposed working definition defined a T-unit as any self-standing clause in order to apply the globally developed measures in Chinese syntactic complexity analysis. However, without looking into the typological features of the Chinese language, such a simplistic adoption of global measures limited Chinese syntactic complexity analysis at the clausal level. It failed to detect the global level Chinese syntactic complexity due to the lack of clarifications on the number and nature of the relationships between these constituent T-units. In addition, only at the clausal level was the mean length of a T-unit applied as the only measure and thus, making it questionable if such analysis was able to gain a

comprehensive picture of the multidimensional construct of Chinese syntactic complexity. If the measure of clause(s) per T-unit was added into analysis, then the application of such a working definition could possibly generalize the same complexity level as a constant 1 for any Chinese output.

Practical issues arise among the communications between academic fields of different dominant research languages. The field of Chinese as a second language acquisition has been holding an open eye to be introduced to and has adopted numerous research findings from the field of English as a second language acquisition. However, there is usually a communication lag between these two fields, though it has been diminishing, due to causes such as language barrier and database accessibility. A language barrier here refers to the communicative difficulties or delays caused by the fact that different languages dominate the two research related fields. Publications in the field of English as a second language acquisition are published in English, however, the majority of publications in the field of Chinese as a second language acquisition are published in Chinese, despite publications in English done by scholars working overseas. Though English literature has been increasingly included in studies conducted in China, for Chinese researchers the publications in Chinese are still the most convenient to search. It demands a lot of time to translate the most updated and influential English publications into Chinese. This situation not only applies to the field of Chinese as a second language acquisition but it also applies to other disciplines where English literature has an edge over Chinese literature. Another reason for communication lag between the field of Chinese as a second language acquisition and the field of English as a second language acquisition might be related to the limited accessibility to international

journals or databases at universities and research institutes in China. Although more databases have become available at universities and research institutions in China, database accessibility in China is still limited or less convenient when compared with U.S. libraries where one can more conveniently see the full text of publications in English. Such language barrier and limited database accessibility challenges also apply in reverse. General second language acquisition studies most commonly use English as the default language and do not necessarily include other languages such as Chinese when it comes to typing in key words to search for currently available literature. As for database accessibility, universities and research institutions in the U.S. may have limited access to the body of Chinese literature.

The challenges from a bottom-up perspective lie in the process of localization. Since the time when western linguistic theories were first introduced to traditional Chinese linguistic studies in *Mashi Wentong* (1898), there has been on-going mismatches and disputes, such as the debate on word-orientation or character-orientation of language, the noncorrespondence between Chinese word class and syntactic component, and the debate on if subject-predicate structure or topic-comment analysis better fits Chinese syntactic analysis, etc. It has been long criticized given the metaphor that mechanical application is like wearing the Indo-European lens in Chinese linguistics study (Zhu, 1985, 1994). Such mismatches and debates in Chinese linguistics has, as a matter of fact, broadened the linguistics study originated from Indo-European language studies and thus has contributed more diversity to the linguistics research in general.

Recognizing the need for introducing the CAF model in order to foster within learners not only an “effective communicative problem solver, but also a longer-term

linguistic development” (Skehan, 1996, p. 21), this dissertation advocates the introduction of the CAF model into Chinese second language teaching and research. From a bottom-up perspective, however, when adopting studies based on one language family to another, cautions can be used to review the typological features of both language families using a contrastive perspective. To keep from being merely mechanical in application, this dissertation steps into the shoes of the Chinese language itself and defines and assesses the Chinese syntactic complexity construct by examining and contrasting its own typological features to the English language. As aforementioned, the dimension of complexity was not officially included in the Chinese nationwide Standards for Mandarin Chinese Proficiency (Hanban, 1995). The practice of Chinese foreign language teaching in the U.S. widely follows the ACTFL/ILR proficiency guidelines. In the ACTFL/ILR proficiency guidelines, though requirement of different range and variety of structures is included in the descriptions of different proficiency levels, an applicable definition of Chinese syntactic complexity is not provided. As discussed in Section 6.2, there lacked research to provide a systemic, comprehensive and accessible interpretation of Chinese complex structures. With the assumed English clause complexing syntagma in the form of subordination and coordination found not accurately applicable to Chinese clause complexing syntagma due to the arbitrary sentence boundary, paratactic connectives, and mismatched coordination and subordination subcategorization, this dissertation proposed to analyze Chinese syntactic complexity based on the taxonomy of TC-units.

By localizing a theoretical construct designed in a global fashion, the findings also contribute to a more accessible operationalization as well as a greater comprehensive

understanding of the global construct. By identifying the difference in Chinese syntactic complexity operationalization from Indo-European languages, this dissertation broadened our understanding of language complexity in a global fashion. The validation of TC-unit based Chinese syntactic complexity measures operationalization confirmed the possibility of more research that will broaden our vision and understanding of the syntactic complexity construct.

7.2 Limitation of the study and suggestions for future studies

More time for the writing task

As mentioned in Section 5.2.2, due to the overall time restriction of TW&ST, the two writing tasks, FW and GR, were allocated with restricted time, 7 minutes and 5 minutes respectively. Though the time designed for the writing tasks was considerably longer than the speaking tasks, the writing modality requires longer time for syntactic complexity composition, especially for higher literary written Chinese of which the lexical choice and complex topic chain trigger might need longer time for a polished piece of work.

The lower correlation in writing tasks comparing with speaking tasks between the scores on four complexity measures and participants' EI scores might have suggested a roof effect for higher proficiency participants producing correspondent complexity level in the speaking tasks. With the TC-unit based measures confirmed with high efficiency at proficiency group membership classification in this dissertation, future research can separate focus on the syntactic complexity of spoken Chinese and written Chinese in order to get more specialized pictures.

Longitudinal study in a classroom context

This dissertation designed and conducted TW&ST to collect spoken and written output from Chinese speakers of varied proficiency levels for complexity measure validation. In addition to cross-sectional study, to fully apply the organic and sustainable approach, more qualitative analysis on the longitudinal development of subsystems of complexity in line with L2 proficiency development is called up. For a multifaceted construct as complexity, a comprehensive investigation cannot be achieved without more studies being conducted in various contexts.

As addressed in the field of task-based language teaching (TBLT), longitudinal studies situating in the classroom context is called for research to “get real”. Norris (2011) called out to direct the TBLT research to fully realize its “potential to combine good theoretical ideas and empirical understandings towards refinement of effective language education”. Applying TBLT at the instructional level, task can be taken not only the unit of language practice activities, but also the unit for input and assessment with Chinese syntactic complexity included. Situating in the authentic environment of classroom learning and teaching, longitudinal studies will be able to provide more comprehensive and sophisticated insights in the dynamic interaction among the development of accuracy, fluency, and complexity, as well as among the development of the subsystems of complexity. For example, when employing more idiomatic expressions in written Chinese, the interaction between complexity at subclausal and clausal level, or even the interaction between syntactic complexity and lexical complexity can potentially show great difference from complexity of spoken language. Lexical complexity covers the density, diversity, sophistication, and compositionality of lexical elements. Another

aspect calls for further study is the emergence of the different types of topic chain along learners' Chinese proficiency development. The frequencies of the input and output of varied types of topic chain can provide more usage-based insight for the teaching of different types of topic chains in terms of pedagogical order and corresponding foci at different developmental stages. Provided with systemic guidelines, complexity development in Chinese as a second language learning and teaching can then be expected to be more effective.

More complexity measures at all levels

Instead of covering all three aspects of CAF, this dissertation focused on one part of the complexity construct: syntactic complexity. To exclude the factor of accuracy, all participants' output was transcribed and coded without filtering out the ungrammatical output to be error free. Four of the proposed Chinese syntactic complexity measures were empirically validated. These four measures, MLTTCU, CTTCU/ATTTCU, MLSTCU, and STCU/TTCU are length and ratio measures that tackle complexity at global and clausal level. Considering the multifaceted nature of complexity, more measures detecting global, clausal, and subclausal/phrase level complexity can be applied and validated such as ratio of different types of terminable TC-unit, dependents per head, and frequency of a specific form proposed in Section 3.3.

In addition to the linear correlation between complexity measure scores and proficiency level or time for acquisition, further investigation of the interaction among the subsystems of complexity needs to be addressed. Considering the multi-dimensional trait of the complexity construct, of different levels of complexity, global, clausal,

subclausal/phrase level complexity, and specific form, the interaction can be studied situating in different proficiency levels, language of different syntagma, cognitive task complexity, and task modality, etc.

Appendix A. Ten types of Chinese topic chains

Table 33 (cited from Li, 2005, p. 25)

Chinese Topic chain patterns

| # | Pattern | Structure | Simple illustration |
|---|-----------------------------|---|--|
| 1 | Typical Topic Chain | T1 (Agent/Theme) — C1 ↓↑ T2 (∅, Agent/Theme) — C2 | (100) <i>Tā zuò wán shǒushù, ∅ zǒuliǎo.</i> 她做完手术, ∅ 走了。 She finished surgery, ∅ left. |
| 2 | Cataphoric Topic Chain | T1 (∅, Agent) — C1 ↓↑ T2 (Agent) — C2 | (101) <i>∅ Shuōzhe, tā xiào qǐlái.</i> ∅ 说着, 他笑起来。 ∅ Saying this, he started laughing. |
| 3 | Patient-Agent Topic Chain | T1 (Agent) — C1 [... V NP(Patient)] ↓ T2 (∅, Theme) — C2 | (102) <i>Wǒ wènle tāmen, ∅ dōu bù zhīdào.</i> 我问了他们, ∅ 都不知道。 I asked them, ∅ didn't know |
| 4 | Patient-Patient Topic Chain | T1 (Agent) — C1 [... NP(Patient)] ↓↑ T2 (∅, Patient) — C2 | (103) <i>Māma gěi wǒ qián, wǒ bù huā ∅.</i> 妈妈给我钱, 我不花 ∅。 Mom gave me money, I didn't spend ∅. |


| | | | |
|---|------------------------------|---|---|
| 5 | Theme-Patient Topic Chain | T1 (Theme) — C1 ↓↑ T2 (∅, Patient) — C2 | (104) <i>Fàn hǎole, nǐ chī ∅ ba.</i> 饭好了，你吃 ∅ 吧。 Dinner is ready, you can eat ∅. |
| 6 | Preposed Topic Chain | T1 (Patient) — C1 ↓↑ T2 (∅, Patient) — C2 | (105) <i>Shū, wǒ bú mǎi ∅, yě bù kàn ∅.</i> 书，我不买 ∅，也不看 ∅。 Books, I don't buy ∅ and don't read ∅. |
| 7 | Presented Topic Chain | T1 (Location/Time) — C1 [...NP] ↓↑ T2 (∅) — C2 | (106) <i>Qiáng shàng yǒu fú huà, ∅ hěn piàoliang.</i> 墙上有幅画，∅ 很漂亮。 On the wall is a picture, ∅ very pretty. |
| 8 | Montage Topic Chain | T1 (Locative) — C1 [...NP] ↓↑ T2 (∅+position word, Locative) — C2 | (107) <i>Zhuō shàng yǒu gè hú, ∅ lǐ miàn yǒu chá.</i> 桌上有个壶，∅ 里面有茶。 On the table there is a pot, ∅ inside there is tea. |
| 9 | Overt Double Topic Chain | T1 — C1 [T' — C'] ↓↑ T2 (∅) — C2 [T'' — C''] | (108) <i>Tā rén lǎo, ∅ xīn bù lǎo.</i> 他人老，∅ 心不老。 He, age is old, ∅ heart is not old. |

| | | | |
|----|------------------------------|--|--|
| 10 | Covert Double Topic Chain | $\begin{array}{c} T1 - C1 \\ \downarrow \uparrow \\ T2 (\emptyset) - C2 [T' - C'] \end{array}$ | <p>(109) <i>Wǒ kūle, ∅ lèi liú mǎnmiàn.</i></p> <p>我哭了，∅ 泪流满面。</p> <p>I cried, ∅ tears streaming down.</p> |
|----|------------------------------|--|--|

Appendix B. TW&ST (English instruction) screenshots

Screen 1 (Not timed)

Language Interaction Labs: Center for Language & Technology <http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language Interaction Labs

Prompter

00twist
logout

Check My Mic

Screen 1

TWIST (1/22)
(help with this question)

help
Seconds Remaining

**Chinese Timed Writing and Speaking Test
(TW&ST)**

Next

home

© 2013 Center for Language & Technology

1 of 1 3/25/15, 9:44 AM

Screen 2 (Not timed)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (2/22)

(help with this question)

help

Seconds Remaining

Screen 2

Guidelines

1. This test contains five sections: (a) comic strip description; (b) video story retelling; (c) free writing; (d) guided rewriting; and (e) Mandarin elicited imitation test.
2. Each section will display its own time limit; once the limit is reached, the program will proceed to the next section. Once you proceed to the next page, you **CANNOT** go back to the previous one. **DO NOT** refresh the page. (Notice that it may take quite some time for the computer to upload your recording after you finish.)
3. You will have a total of **50 minutes** to finish this test. After 50 minutes the program will automatically close. The computer will start timing once you click "**Next**".
4. Raise your hand to call for assistance anytime during the test.

Next

home

© 2013 Center for Language & Technology

1 of 2

3/13/15, 3:45 PM



Answer the following questions without missing any.

1. Your name *

First Name
Last Name

2. Your Participant No. (on your consent form) *

3. Your gender *

 Male Female

4. Your age *

5. Your dominant language *

English, Cantonese, Mandarin, or others?

* Are you fluent in other language(s) besides your dominant language? *

6. Your class standing *

 Undergraduate Masters Ph.D. Other

7. Your major/field of study *

8. How many years have you been studying Chinese? *

9. What Chinese course(s) are you currently taking? *

10. What Chinese language course(s) have you taken in UH or other college, high school, and other institution or language program? When and where did you take the course(s)? *

E.g. CHN101-----UH-----Spring 2012

11. Does anyone in your family speak Mandarin, Cantonese, or another Chinese dialect? *

Screen 4 (Not timed)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (4/22)
(help with this question)

help

Seconds Remaining

Screen 4

Trial

To help familiarize you with the interface of this test, we will now conduct a brief trial. This trial will not be counted into your grade.

On the next slide you will see a comic strip. Based on the comic strip, narrate a brief story in Chinese. You will have **10 seconds** to prepare before recording begins. Your recording should be up to **10 seconds**.

Next

home

© 2013 Center for Language & Technology

1 of 1

3/14/15, 11:08 AM

Screen 5 (Timed for 10 seconds)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

[Check My Mic](#)

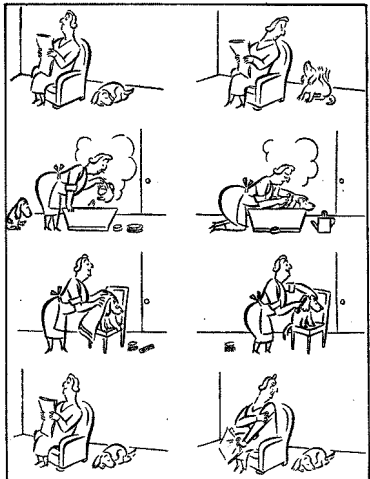
TWIST (5/22)
(help with this question)

[help](#)

Screen 5

7
Seconds Remaining

Trial

| | |
|--|---|
|  | <p>Recording will begin in 10 seconds.</p> |
|--|---|

[home](#)

© 2013 Center for Language & Technology

1 of 1

3/14/15, 11:09 AM

Screen 6 (Timed for 10 seconds)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (6/22)
(help with this question)

help

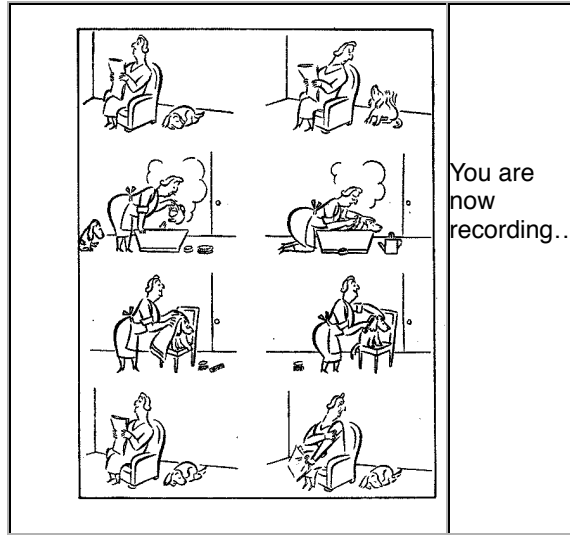
Screen 6

6
Seconds Remaining

00:08 **Ready.**

Mic Check

Trial



home

© 2013 Center for Language & Technology

Screen 7 (Not timed)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (7/22)

(help with this question)

help

Seconds Remaining

Screen 7

End of Trial

Notice the trial is just to help you familiarize with the computer settings. It is not counted into your total score.

Click "**Next**" to begin the test and timing.

Next

[home](#)

© 2013 Center for Language & Technology

1 of 1

3/14/15, 11:16 AM

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

Screen 8

TWIST (8/22)
(help with this question)

help
Seconds Remaining

Comic Strip Description

You have up to **2 minutes** to finish this test section.

On the next slide you will see a comic strip. Based on the comic strip, narrate a brief story.

You will have **30 seconds** to prepare before recording begins. Your recording should be up to **1.5 minutes**.

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

Screen 9 (Timed for 30 seconds)

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST (9/22)

(help with this question)

help

Screen 9

28
Seconds Remaining

Comic Strip Description

Recording will begin in **30 seconds**.

[home](#)

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/1

Screen 10 (Timed for 90 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

ootwist
logout
Check My Mic

TWIST Chinese (10/20)

(help with this question)

help

Screen 10

72
Seconds Remaining

01:12 Ready.
Mic CHECK 

Comic Strip Description



You are now recording...

home

© 2013 Center for Language & Technology

Screen 11 (Not timed)

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (11/22)

(help with this question)

help

Seconds Remaining

Screen 11

Video Story Retelling

Watch the following short English video clip about the Chinese traditional festival Nian年, and then retell the story in Chinese. **DO NOT translate the story literally, word-for-word, but instead, retell the story.** You have **30 seconds** to prepare. Your recording should be **2-3 minutes**.

Time for the video watching is limited for only once play without any pausing. So **DO NOT pause the video** at any point; otherwise you will not be able to finish watching all of the video.

Note taking is **NOT** allowed.

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/1

Screen 12 (Timed as the length of the video: 3 minute and 40 seconds)

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST (12/22)
(help with this question)

help
Seconds Remaining

Screen 12

A Story of Chinese New Year 年的故事 Nián de Gùshi

Right click mouse and choose "show controls" if the video does not show



home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

Screen 13 (Timed for 30 seconds)

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

Screen 13

TWIST (13/22)
(help with this question)

help
26
Seconds Remaining

Video Story Retelling

Recording will begin in **30 seconds**.

Vocabulary you might need: (will also be provided in next page when you are recording)

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

[home](#)

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

Screen 14 (Timed for 180 seconds)

3/14/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (14/22)

(help with this question)

help

Screen 14

176
Seconds Remaining

02:57 **Ready.**

Mic **CHECK**

Video Story Retelling

You are recording now...

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

Notice that it takes quite some time to upload your recording after you finish.

[home](#)

© 2013 Center for Language & Technology

Screen 15 (Timed for 30 seconds)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

[Check My Mic](#)

TWIST (15/22)
(help with this question)

[help](#)

Screen 15

26
Seconds Remaining

Video Story Retelling (Repeat)

You've got **a second chance!**

Repeat your recording, retelling the story of Nian年 to **IMPROVE** on your performance.

Again, your recording should be **2-3 minutes.**

The computer will automatically move to the next page and start recording when it reaches **30 seconds.**

[home](#)

© 2013 Center for Language & Technology

Screen 16 (Timed for 180 seconds)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (16/22)

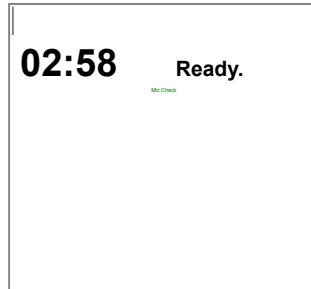
(help with this question)

help

175

Seconds Remaining

Screen 16



Video Story Retelling (Repeat) You are now recording...

Vocabulary you might need:

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

Notice that it takes quite some time to upload your recording after you finish.

[home](#)

© 2013 Center for Language & Technology

Screen 17 (Not timed)

3/25/2015



Language
Interaction
Labs

Language Interaction Labs: Center for Language & Technology

Prompter

00twist
logout

Check My Mic

TWIST (17/22)

(help with this question)

help

Seconds Remaining

Screen 17

Recall your two video story retellings to answer the following questions. Check ONE of the three boxes for each question.

Your Participant No. (find it on your consent form) *

Section A *

| | the 1st time | the 2nd time | same for both times |
|--|-----------------------|-----------------------|-----------------------|
| 1. Overall I performed better | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. I spoke more fluently | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. I used more accurate language | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. My language was more complex/sophisticated | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. My organization of the story was better | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. I felt more ease and confidence | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. I was more focused when I was retelling | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. I remembered more of the story content | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. I deliberately tried harder to speak without pause/hesitation | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. I deliberately tried harder to use the right words/grammar | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. I deliberately tried harder to use more advanced words/structures | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. I deliberately tried harder to use more diverse words/structures | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 13. I deliberately paid more attention to the organization of my story | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

Section B *

| | true | false | not sure |
|--|-----------------------|-----------------------|-----------------------|
| 14. I am familiar with this story about Nián before watching this video | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. As a result of the 1st time speaking, I spoke more fluently the 2nd time | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. As a result of the 1st time speaking, I spoke more accurate the 2nd time | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. As a result of the 1st time speaking, I better structured words/expressions the 2nd time | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. As a result of the 1st time speaking, I got tired and impatient the 2nd time | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 19. I could perform better if I had a 3rd retelling opportunity | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

20. Comparing your performance and strategies used for the two retellings, are there any other thoughts or comments you would like to share?

<http://cit.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

Submit

Next
Home

© 2013 Center for Language & Technology

Screen 18 (Timed for 420 seconds)

Language Interaction Labs: Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>



Language
Interaction
Labs

Prompter

00twist

[logout](#)

[Check My Mic](#)

TWIST (18/22)

(help with this question)

[help](#)

416

Seconds Remaining

Screen 18

Free Writing

You have up to **7 minutes** to finish this test section. See your seconds remaining at top right.

Write a well-organized composition describing your relationship with your father/mother/brother/sister/friend (choose any one of them). You may talk about their profession or personality, and the relationship between the two of you, etc.. Use details to help elaborate your writing. Type your composition in Chinese characters. Change the language setting above right to allow you type in Chinese. (Or raise your hand to call for assistance.) Type *pinyin* if you don't know the character.

The computer will automatically move to the next slide when it reaches **7 minutes**.

Please enter your response here.

[home](#)

© 2013 Center for Language & Technology

Screen 19 (Timed for 420 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (19/22)

(help with this question)

help

294

Seconds Remaining

Screen 19

Guided Rewriting

You have up to **5 minutes** to finish this test section.

Based on the following sentences, turn it into a coherent story. Construct a coherent paragraph in a way that is better in terms of sentence structures. You may manipulate sentences, change word order, omit words, and add connectors, etc., but try not to leave out any information or add any information.

The **COMPUTER** will automatically move to the next slide when it reaches **5 minutes**.

【simplified characters】

1. 王朋在学校的宿舍住了两个学期了
2. 王朋觉得宿舍太吵
3. 王朋觉得房间太小
4. 王朋觉得宿舍不能做饭，很不方便
5. 王朋准备下个学期搬出去
6. 王朋找房子找了一个多月
7. 王朋没有找到合适的

【traditional characters】

1. 王朋在學校的宿舍住了兩個學期了
2. 王朋覺得宿舍太吵
3. 王朋覺得房間太小
4. 王朋覺得宿舍不能做飯，很不方便
5. 王朋準備下個學期搬出去
6. 王朋找房子找了一個多月
7. 王朋沒有找到合適的

Copy the sentences from the box above and **paste** them in your answer to save time.

Please enter your response here.

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST (20/22)

(help with this question)

help

Seconds Remaining

Screen 20

Mandarin Elicited Imitation Test

This section of the test will take **9 minutes 40 seconds**.

For this last part, you will hear 30 Chinese sentences, one by one, of various length and complexity.

After hearing each sentence, you will repeat it as exactly as possible in the time provided.

Prior to the Chinese part, you will first do a practice round using sentences in English.

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>

1/2

Screen 21 (Timed for 580 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST (21/22)

(help with this question)

help

581
Seconds Remaining

Screen 21

Record while listening

09:42

Ready.

Mic Check

Mandarin Elicited Imitation Test

This section of the test will take **9 minutes 40 seconds**.
Prior to the Chinese repetition part, you will first do a practice round
using sentences in English.

You are recording now...
Notice that it takes quite some time to upload your recording after
you finish.

[home](#)

© 2013 Center for Language & Technology

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

Screen 22

TWIST (21/22)
(help with this question)

help
Seconds Remaining

This is the end of the test.
Xiexie and congratulations! :)

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/1/>


1/2

Appendix C. TW&ST (Chinese instruction) screenshots

Screen 1 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology

 Language Interaction Labs

Prompter

00twist
logout
Check My Mic

Screen 1

TWIST Chinese (1/20)
(help with this question)

help
Seconds Remaining

Chinese Timed Writing and Speaking Test
(TW&ST)
中文版

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/> 1/2

Screen 2 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (2/20)
(help with this question)

help

Seconds Remaining

Screen 2

考试说明

1. 本测试主要包括四个部分：（一）看卡通讲故事；（二）看短片复述故事；（三）写作；（四）改写句子。
2. 测试的每一部分各有**时间限制**，每一部分的剩余时间都显示于电脑**右上方**。在录音结束后，请**耐心等待**电脑上传你的录音。时间到后电脑会自动跳至下一页。请**注意**，进入下一页以后，您无法再返回上一页。请**不要**刷新页面。
3. 本测试共计**35分钟**时间。**35分钟**后电脑会自动结束测试。测试期间如有问题，请随时告知。
4. 现在，请您点击"**Next**"开始考试。

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 3 (Not timed)



Language Interaction Labs

Language & Technology

Prompter

00twist
logout
Check My Mic

TWIST Chinese (3/20)
(help with this question)

help

Seconds Remaining

Screen 3

请你回答以下几个问题。请不要遗漏任何问题。

1. 姓名

2. 序号 (参见知情同意单)

3. 性别 男 女

4. 年龄

5. 您现在所在的年级; 如已毕业, 请选择您的最高学历水平。

- 大学本科
- 硕士研究生
- 博士研究生

6. 专业

7a. 除了普通话以外, 您是否能够流利使用某一(些)方言?

- 不是
- 是

7b. 如果你在第7a题选择了“是”, 请在下面具体说明您能流利使用哪种方言?

- 8a. 请选择能够代表你英语水平的考试和成绩
- 大学英语四级
 - 大学英语六级
 - 大学英语专业四级
 - 大学英语专业八级
 - TOEFL成绩 80分以上
 - IELTS成绩5.5分以上
 - 硕士入学英语考试及格
 - 博士入学英语考试及格
 - 见下面8b

8b. 如果以上考试你都不曾参加过, 请在下面写出能反映您的英文水平考试及成绩:

9. 你是否在主要说英文的国家和地区居住过一个月以上的时间? 在哪些国家和地区? 多长时间?

Screen 4 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST Chinese (4/20)
(help with this question)

help
Seconds Remaining

Screen 4

试验部分

试验部分将帮助您尽快适应该测试的电脑界面，不计入总成绩。

在下一页，你会看到一组卡通图片。请您根据这组图片，描述一个完整的故事。您有10秒钟的时间准备，然后电脑会自动跳至下一页进行录音。您同样有10秒钟的时间进行录音。

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 5 (Timed for 10 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

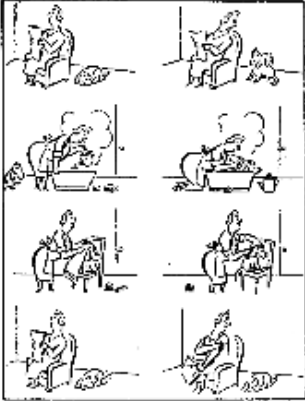
TWIST Chinese (5/20)
(help with this question)

help

Screen 5

2
Seconds Remaining

试验部分

| | |
|--|----------------|
|  | 录音会在10秒钟后自动开始。 |
|--|----------------|

[home](#)

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 6 (Timed for 10 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (6/20)


(help with this question)

help






Screen 6

8
Seconds Remaining

00:08 Ready.

Mic CHECK 

试验部分

| | |
|---|--|
|  |  |
|  |  |
|  |  |
|  |  |

您正在录音。

home

© 2013 Center for Language & Technology

Screen 7 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST Chinese (7/20)
(help with this question)

help
Seconds Remaining

Screen 7

试验部分结束

请注意刚才的试验部分只是为了帮助您熟悉电脑界面，不计入考试成绩。

点击 **"Next"** 开始正式考试部分，并开始计时。

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 8 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

Screen 8

TWIST Chinese (8/20)
(help with this question)

help
Seconds Remaining

看卡通说故事

共计时间：2分钟

在下一页您将看到一组卡通图片。请根据这组卡通图片说一个完整的故事。
在开始录音之前，你有 **30 秒** 进行准备。你的录音最长应不超过 **1.5 分钟**。

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 9 (Timed for 30 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST Chinese (9/20)
(help with this question)

help

Screen 9

27
Seconds Remaining

看卡通说故事

| | | |
|---|---|----------------|
|  |  | 30秒后电脑将自动开始录音。 |
|  |  | |
|  |  | |

[home](#)

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/1

Screen 10 (Timed for 90 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

Screen 10


TWIST Chinese (10/20)

(help with this question)

help

72
Seconds Remaining

01:12 **Ready.**

Mic **CHECK** 

看卡通说故事



您正在录音。

home

© 2013 Center for Language & Technology

Screen 11 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST Chinese (11/20)
(help with this question)

help
Seconds Remaining

Screen 11

看短片复述故事

下面您将看到一则关于中国的传统节日“年”的英文短片。看完后，请您用中文复述这个故事。
请注意：不要逐字翻译这个故事，而是用您自己的语言完整地复述这个故事。看完短片后您有30秒钟准备。你的录音时间为2-3分钟。

短片只能播放一次，不能暂停或重放。

观看短片时，**不允许**记笔记。

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 12 (Timed as the length of the video: 3 minute and 40 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (12/20)

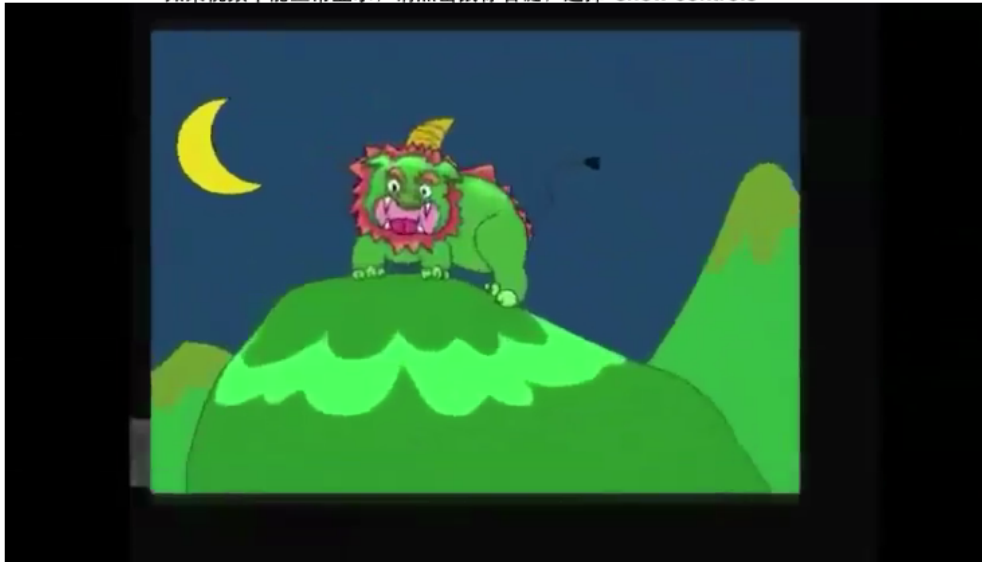
(help with this question)

help

Seconds Remaining

Screen 12

A Story of Chinese New Year 年的故事
如果视频不能正常显示，请点击鼠标右键，选择"show controls"



home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/1

Screen 13 (Timed for 30 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (13/20)
(help with this question)

help

Screen 13

27
Seconds Remaining

看短片复述故事

您现在有**30秒**准备，录音在**30秒**后自动开始。

请用中文复述这个故事。不要逐字翻译。

词语提示

- monster: 怪物
- bamboo: 竹子

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 14 (Timed for 180 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic


TWIST Chinese (14/20)

(help with this question)

help

Screen 14

178
Seconds Remaining

02:58 **Ready.**
Mic **CHECK** 

看短片复述故事

您正在录音。

词语提示:

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

录音完成后，请耐心等待电脑上传您的录音。

[home](#)

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 15 (Timed for 30 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (15/20)
(help with this question)

help

Screen 15

27
Seconds Remaining

看短片复述故事 (第二次)

恭喜您! 您得到了第二次机会!

请您再复述一次“年”的故事, 使您的故事复述得更好, 努力得到更高的分数。您的录音时间应控制在 **2 3 分钟**。

30 秒准备时间后, 电脑会自动跳至下一页。

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 16 (Timed for 180 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (16/20)

(help with this question)

help

Screen 16

177
Seconds Remaining

02:58 **Ready.**

Mic **CHECK**

看短片复述故事（第二次）

您正在录音。

词语提示：

- monster: 怪物 guàiwù
- bamboo: 竹子 zhúzi

录音完成后，请耐心等待电脑上传您的录音。

[home](#)

© 2013 Center for Language & Technology

Screen 17 (Not timed)

3/25/2015



Language Interaction Labs: Center for Language & Technology

Prompter

00twist
logout
Check My Mic

TWIST Chinese (17/20)
(help with this question)

help
Seconds Remaining

Screen 17

回顾你对《“年”的故事》的两次复述，回答下面的问题：

序号（见于知情同意单） *

ex: 23

第一部分 *

| | 在第 1 次复述时 | 在第 2 次复述时 | 两次都一样 |
|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. 总体来说，我表现得更好 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. 我说得更加流利 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. 我使用的语言更加准确 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. 我使用的语言更加复杂 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. 我说的故事结构更好 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. 我觉得更加放松和自信 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. 我复述的时候更加专注 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. 我记得更多故事的内容 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. 我有意识地想减少我的停顿和拖延 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 10. 我有意识地想使用更准确的词汇和结构 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 11. 我有意识地想使用更高级的词汇和结构 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 12. 我有意识地想使用更多样的词汇和结构 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 13. 我有意识地想组织更好的故事结构 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

第二部分

| | 对的 | 不对 | 不确定 |
|------------------------------------|-----------------------|-----------------------|-----------------------|
| 14. 我以前就知道这个视频所讲的“年”的故事 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 15. 因为有了第一次复述，我在第二次复述时更流利 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 16. 因为有了第一次复述，我在第二次复述时更准确 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 17. 因为有了第一次复述，我在第二次复述时能更好地组织我的语言结构 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 18. 因为有了第一次复述，我在第二次复述时失去了耐心 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

19. 如果有第三次机会, 我会复述得更好

20. 比较你在两次复述中的语言表现和使用策略, 你还有什么体会?

*

提交问卷

点击上面的按钮“提交问卷”后, 请点击页面底部的“Next”继续。

Next
home

© 2013 Center for Language & Technology

Screen 18 (Timed for 420 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (18/20)
(help with this question)

help

Screen 18

413
Seconds Remaining

写作

时间限制：7 分钟。电脑右上角显示您的剩余时间。

请以“我和我的父亲/母亲/哥哥/弟弟/姐姐/妹妹/朋友（选择其中任何一个）”为题，写一篇短文。可以说一说他/她的职业、性格，你和他/她之间的关系等等。请用具体的例子充实、扩展您的写作。

您有7分钟的写作时间。7分钟后，电脑会自动跳至下一页。

Please enter your response here.

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/1

Screen 19 (Timed for 300 seconds)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout

Check My Mic

TWIST Chinese (19/20)
(help with this question)

help

Screen 19

286
Seconds Remaining

改写句子 时间限制: 5分钟

请改写下面的7个句子, 成为一个连贯的段落。你可以改变句子结构, 变换词序, 改写、删除或添加词语使段落连贯。但是, 请不要添加或者删除这些句子组成的故事内容及所提供的信息。

1. 王朋在学校的宿舍住了两个学期了
2. 王朋觉得宿舍太吵
3. 王朋觉得房间太小
4. 王朋觉得宿舍不能做饭, 很不方便
5. 王朋准备下个学期搬出去
6. 王朋找房子找了一个多月
7. 王朋没有找到合适的

5分钟后, 电脑会自动跳至下一页。请复制和粘贴上面7个句子来节省打字的时间。

Please enter your response here.

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Screen 20 (Not timed)

3/25/2015

Language Interaction Labs: Center for Language & Technology



Language
Interaction
Labs

Prompter

00twist
logout
Check My Mic

TWIST Chinese (20/20)
(help with this question)

help
Seconds Remaining

Screen 20

恭喜您完成测试!
再次感谢您的参与:)

Next

home

© 2013 Center for Language & Technology

<http://clt.manoa.hawaii.edu/langlabs/prompter/collection/4/>

1/2

Appendix D. Mandarin Chinese EI test (cited from APPENDIX C MANDARIN REPETITION TASK in Zhou, 2012, p. 188)

You are going to hear several sentences in English. After each sentence, there will be a short pause, followed by a tone sound {TONE}. Your task is to try to repeat exactly what you hear. You will be given sufficient time after the tone to repeat the sentence. Repeat as much as you can. Remember, DON'T START REPEATING THE SENTENCE UNTIL YOU HEAR THE TONE SOUND {TONE}. Now let's begin.

I like flowers. I am writing a letter. I don't think I need a big car. As it is raining, I don't go out. The little girl hurt herself and started to cry. As soon as I returned home, I watched TV with my sister.

That was the last English sentence

Now, you are going to hear a number of sentences in Mandarin. Once again, after each sentence, there will be a short pause, followed by a tone sound {TONE}. Your task is to try to repeat exactly what you hear in Mandarin. You will be given sufficient time after the tone to repeat the sentence. Repeat as much as you can. Remember, DON'T START REPEATING THE SENTENCE UNTIL YOU HEAR THE TONE SOUND {TONE}.

Now let's begin.

- 1 我得去剪头发了。
- 2 红色的书在桌子上。
- 3 这个城市街道很宽。
- 4 他每天早上都要洗澡。
- 5 我听说明天可能会下雨。
- 6 你刚才说你今天在做什么？

- 7 我不觉得他开车开得很好。
- 8 晚饭以后,我好好地睡了一觉。
- 9 我喜欢看有快乐结局的电影。
- 10 这些房子好是好,就是太贵了。
- 11 昨天死了小猫的小男孩很伤心。
- 12 那家饭馆的中国菜应该很不错。
- 13 你真的很喜欢听流行音乐,对不对?
- 14 她刚把公寓所有的房间都漆完了。
- 15 在红绿灯那过马路,然后一直往前走。
- 16 我希望别墅能便宜一些,我才买得起。
- 17 我现在交往的那个人非常有幽默感。
- 18 我想要一个我的宠物可以住的大房子。
- 19 我希望今年的天气会比去年暖和一点。
- 20 我的一个好朋友老帮他的邻居看孩子。
- 21 他得先把房间打扫干净才可以出去玩。
- 22 我最快乐的回忆就是那次跟你去看戏。
- 23 那个被警察抓到的小偷长得又高又瘦。
- 24 根据统计,每年吸烟的人数越来越多了。
- 25 这次考试根本没有你跟我说的那么难。
- 26 她点菜的时候只点有肉的,从来不点青菜。
- 27 你昨天喂的那只黑猫就是被狗追的那只。

28 可以麻烦您把桌子上的那本书递给我吗?

29 我不知道十点半的火车是不是已经开走了。

30 为什么还是有很多人早上什么东西都不吃呢?

This is the end of the repetition task. Thank you.

Appendix E. EI task scoring rubric (cited from APPENDIX D MANDARIN
 REPETITION TASK in Zhou, 2012, p. 190)

SCORE 0

| Criteria | Examples |
|---|--|
| <ul style="list-style-type: none"> Nothing (Silence) | |
| <ul style="list-style-type: none"> Garbled (unintelligible, usually transcribed as XXX) | |
| <ul style="list-style-type: none"> Minimal repetition, then item abandoned: <ul style="list-style-type: none"> - Only 1 word repeated - Only 1 content word plus function word(s) - Only 1 content word plus function word(s) plus extraneous words that weren't in the original stimulus - Only function word(s) repeated <p>NOTE: with only, just, yet (meaningful adverbs), score 1</p> | <ul style="list-style-type: none"> - 我喜欢..... (12/#9) - 你昨天挨摸...什么什么.. (01/#27) - 昨天..校劳 (36/#11) - 我.... (113/#1) - 她刚把...做完了...(116/#14) <p>(score 1)</p> |

SCORE 1

| Criteria | Examples |
|---|---|
| <ul style="list-style-type: none"> When only about half of idea units are represented in the string but a lot of | <ul style="list-style-type: none"> - 住的大房子. (88/#18) - 晚饭以后,我... (56/#8) |

| | |
|---|---|
| <p>important information in the original stimulus is left out</p> <ul style="list-style-type: none"> When barely half of lexical words get repeated and meaningful content results that is unrelated (or opposed) to stimulus, frequently with hesitation markers | <ul style="list-style-type: none"> - 在红绿灯...(61/#15) - 那个被警察...(61/#23) - 她点菜的时候...(107/#26) - 我...想快...电影 (71/#9) - 我希望...便宜...(107/#16) |
| <ul style="list-style-type: none"> Or when string doesn't in itself constitute a self-standing sentence with some (targetlike or nontargetlike) meaning (This may happen more often with shorter items, where if only 2 of 3 content words are repeated and no grammatical relation between them is attempted, then score 1) Also when half of a long stimulus is left out, and the sentence produced is incomplete | <ul style="list-style-type: none"> - 那家饭....还的不错(30/#12) - 我听说的夏天下雨 (66/#5) - 你真的...对不对?(112/#13) - 可是好是好,快了 (102/#10) - 这个东市道很还宽 (97/#3) - 我能希望...便宜...(55/#16) - 考试真多的...那么很难(55/#25) - 考试...那么难.(102/#25) - 你昨天...那黑猫...(92#27) |

SCORE 2

| Criteria | Examples |
|----------|----------|
|----------|----------|

| | |
|---|---|
| <ul style="list-style-type: none"> When content of string preserves at least more than half of the idea units in the original stimulus; string is meaningful, and the meaning is close or related to original, but it departs from it in some slight changes in content, which makes content inexact, incomplete, or ambiguous | <p>-我想我的宠物住的一个大房子 (51/#18) (<left out “要,可以”; changed the position of “一个”)</p> <p>-这房子好的好,太贵了 (6/#10) (<left out the measure word “些” and cohesive “就是”; incorrect use of the pattern “好是好”)</p> <p>-我不觉得..开得很好 (97/#7) (<left out “他开车” and made the sentence ambiguous)</p> |
|---|---|

SCORE 3

| Criteria | Examples |
|--|--|
| <ul style="list-style-type: none"> Original, complete meaning is preserved as in the stimulus. Strings which are quite ungrammatical can get a 3 score, as long as exact meaning is preserved. Some synonymous substitutions are acceptable. Examples of acceptable substitutions (SCORE 3): 好好=好好地;桌子=桌 | <p>-我不觉得他开车得开很好(1/#7)</p> <p>-他每天都早上都要洗澡(56/#4)</p> <p>-晚饭以后,我好好睡了一觉 (11/#8)(Score 3) -红色的书在桌上 (59/#2)(Score 3) -昨天死了小猫的小男孩好伤心 (57/#11)(Score 3)</p> |

| | |
|--|--|
| <ul style="list-style-type: none"> Anything with ‘很’ can be substituted with ‘好’ Examples of unacceptable substitutions or omissions (SCORE 2): <ul style="list-style-type: none"> - 个>些 - 在>想 - 听说>听到 - 很>最 - 才>然后 - 递给>给 Changes in grammar that affect meaning should be scored as 3. For example, a present progressive tense repeated as past or as future should be scored as meaning change (score 2). Similarly, singular/plural differences between stimulus and repeated string | <p>-这些考试根本没有你跟我说的那么难 (41/#25)(Score 2)</p> <p>-你刚才说你今天想做什么 (11/#6)(Score 2)</p> <p>-我听到明天可能会下雨(56/#5)(Score 2)</p> <p>-我不觉得他开车开得最好 (104/#7)(Score 2)</p> <p>-他得先把房间打扫干净然后可以出去玩 (41/#21)(Score 2)</p> <p>-可以麻烦您把桌子上的那本书给我吗 (67/#28)(Score 2)</p> <p>-你刚才说你今天要做什 么.(41/#6)(Score 2)</p> <p>-晚饭以后,我好好地睡一觉. (41/#8)(Score 2)</p> |
|--|--|

| | |
|---|--|
| <p>change the meaning, not only the grammar (score 2).</p> <ul style="list-style-type: none"> • Ambiguous changes in grammar that COULD be interpreted as meaning changes from a NS perspective should be scored as 2. That is, as a general principle in case of doubt about whether meaning has changed or not, score 2. | <p>-这些考试根本没有你跟我说的那么难 (41/#25)(Score 2)</p> <p>-这些房子好是好,但太贵了 (14/#10)(Score 2) (但 refers to a turn in meaning, but 就是 not only refers to a turn in meaning, but also points out the only flaw.)</p> |
|---|--|

SCORE 4

| Criteria | Examples |
|--|-----------------|
| <ul style="list-style-type: none"> • Exact repetition: String matches stimulus exactly. Both form and meaning are correct without exception or doubt. | |

REFERENCES

- Anderson, J. R. (1995). *Learning and memory: An integrated approach*. New York: Wiley.
- Beijingdaxue Zhongwenxi Xiandaihanyu Jiaoyanshi [Peking University Modern Chinese Teaching and Research Section]. (eds.). (2004). *Xiandai Hanyu [Modern Chinese] (Recompiled)*. Beijing: The Commercial Press.
- Brindley, G. (1994). Task-centered language assessment in language learning: The promise and the challenge. In: N. Brid, P. Falvey, A. Tsui, D. Allison, & A. McNeil (Eds.). (1994). *Language and learning: Papers presented at the Annual International Language in Education Conference, Hong Kong* (pp. 73-94). Hong Kong: Hong Kong Education Department.
- Brown, H. D. (2001). *Teaching by Principles: An Interactive Approach to Language Pedagogy*. New York: Longman.
- Brown, J. D., Robson, G., & Rosenkjar, P. (2001). Personality, motivation, anxiety, strategies, and language proficiency of Japanese students. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 361-398). Honolulu, HI: Second Language Teaching & Curriculum Center, University of Hawai'i Press.
- Brumfit, D.J. (1984). *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.
- Bulté, B. & Housen, A. (2012). "Defining and Operationalising L2 Complexity". In: Housen, A., Kuiken, F., & Vedder, I. (eds.), *Dimensions of L2 Performance and*

- Proficiency: Investigating Complexity, Accuracy and Fluency in SLA* (pp. 21-46). Amsterdam/Philadelphia: John Benjamins.
- Bygate, M. (1996). Effect of task repetition: Appraising the development of second language learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-46). Oxford: Heinemann.
- 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: Longman.
- Byrnes, H, Crane, C., Maxim, H., and Sprang, K. (2006). Taking text to task: Issues and choices in curriculum construction. *ITL – International Journal of Applied Linguistics*, 152, 85-109.
- Chao, Y-R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chen, P. (1989). Hanyu de lingxing huizhi de huayu fenxi (Discourse analysis of zero anaphors in Chinese). *Zhongguo Yuwen (Chinese Linguistics)*, 87(5), 365-378.
- Chu, C. C. (1998). *A discourse grammar of Mandarin Chinese*. New York and Berne: Peter Lang Publishing.
- Chu, C. C. (2006). *Hanyu pianzhang yufa [A discourse Grammar of Mandarin Chinese]*. (Pan, W. Trans.) Beijing: Beijing Language and Culture University Press.
[Original work: Chu, C. C. (1998). *A discourse Grammar of Mandarin Chinese*. New York and Berne: Peter Lang Publishing.]
- Crookes, G. 1989. Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367-383.

- Curzan, A., & Adam M. (2009). *How English Works: A Linguistic Introduction* (2nd ed.). New York: Pearson Longman.
- De Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. In: Long, M. & Doughty, C. (Eds.). *The Handbook of Language Teaching* (pp. 726-755). Oxford: Blackwell Publishing.
- Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford: Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27(3), 464-491.
- Fang, M. (2006). Shumianyu de liangzhong beijinghua shouduan [Two ways of backgrounding in writtern language]. *The 14th Symposium on Modern Chinese Grammar*. Shanghai: Shanghai University of Finance and Economics.
- Foster, P. & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18, (3), 299-324.
- Foster, P, Tonkyn, A, & Wigglesworth G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*. 21, (3), 354-375.

- Gass, S. M., Mackey, A., Fernandez, M., & Alvarez-Torres, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49, 549–80.
- Givón, T. (1983). *Topic continuity in discourse: A quantitative crosslanguage study*. Philadelphia: John Benjamins.
- Givón, T. (1992). *English Grammar: A Function-based Introduction, 1/2*. Amsterdam: John Benjamins.
- González-Lloret, M. (2003). Designing task-based CALL to promote interaction-En busca de esmeraldas. *Language Learning & Technology*. 7 (1), 86-104.
- Hakstian, A. R., Roed, J. C., & Lind, J. C. (1979). *Two-sample T^2 procedure and the assumption for homogeneous covariance matrices*. *Psychological Bulletin*, 86, 1255-1263.
- Hanban. (eds.) (1995). Hanyu shuiping dengji biao zhun. *Yuyan Wenzhi Yingyong [Applied Linguistics]*. 4, 85-91.
- Housen, A., & Kuiken, F. (eds.). (2009). *Complexity, accuracy, and fluency in second language acquisition: Theoretical and methodological perspectives*. Special Issue of *Applied Linguistics*, 30 (4), 461-473.
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In: A. Housen, F. Kuiken, & I. Vedder, (Eds.), *Dimensions of L2 Performance and Proficiency: Investigating Complexity, Accuracy and Fluency in SLA* (pp. 1-20). Amsterdam/Philadelphia: John Benjamins.
- Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English*. London: Longman.

- Halliday, M.A.K., McIntosh, A. & Strevens, P. (1964). *The Linguistic Sciences and Language. Teaching*. London: Longmans.
- Halliday, M.A.K. & Hasan, R. (1985). *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Victoria: Deakin University Press.
- Halliday, M.A.K. & Matthiessen, C.M.I.M. (2004). *An Introduction to Functional Grammar*. London: Edward Arnold.
- Hopper, P. J. & Traugott, E. C., (2003). *Grammaticalization*. (2nd Edition). Cambridge: Cambridge University Press.
- Huang, C.-T. J. (1984). On the distribution and reference of empty pronouns. *Linguistics Inquiry*, 15, 531-544.
- Hunt, K.W. (1965). *Grammatical structures written at three grade levels*. Urbana, IL: The National Council of Teachers of English.
- Hunt, K.W. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Child Development* 53, 1 (Serial No. 134).
- Hunt, K.W. (1977). Early blooming and late blooming syntactic structures. In: C.R. Cooper, & L. Odell (Eds.). *Evaluating Writing* (pp.91-106). Urbana, Illinois: NCTE.
- Jin, H. (2004). The Role of Formulaic Speech in Teaching and Learning Patterned Chinese Structures, *Journal for Chinese Language Teachers Association (JCLTA)*, 39(1), 45-62.
- Jin, H. (2006). Cong Hanyu xiezuo guocheng kan CFL yuyan jiegou fuzadu de fazhan [The complexity development of Chinese as a foreign language: a writing

- perspective]. *Journal of Chinese Language Studies*, 2, 114-140, Beijing: Peking University Press.
- Jiang, W. (2013). Measurements of development in L2 written production: The case of Chinese L2. *Applied Linguistics*, 34 (1), 1-24.
- Kelly, J. & Mao, N. K. (1980). (tran.) *Fortress Besieged*. Bloomington and London: Indiana University Press.
- Kondo-Brown, K. (2005). Differences in language skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal*, 89, 563-581.
- Kormos, J. (2011) Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, 20, 148-161.
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12(4), 439-448.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590-619.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40(3), 387-417.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Li, C. N. & Thompson, S. (1976). Subject and topic: a new typology of language. In: Li, C. N. (Ed.). *Subject and Topic* (pp.457-489). New York: Academic Press.
- Li, C. N. & Thompson, S. (1981). *Mandarin Chinese: A function reference grammar*, Berkeley: University of California Press.

- Li, W. (2005). *Topic Chains in Chinese: A Discourse Analysis and Application in Language Teaching*. Deutsche: Lincom Enropa.
- Li, Y-C, Cheng, R. L., Foster, L., Ho, S H., Hou, J. Y. & Yip, M. (1989). *Mandarin Chinese: A practical reference grammar for students and teachers*. 2. Taiwan: The Crane Publishing Co. Chinese Materials Center Publications.
- Lian, S. (1993). *Contrastive Studies of English and Chinese*. China: Higher Education Press.
- Liu, D. (2004). Identical topics: a more characteristic property of topic prominent languages. *Journal of Chinese Linguistics*, 32(1): 20-64.
- Liu, Y., Yao, T-C., Shi, Y., Ge, L. & Bi, N-P. (2009). *Integrated Chinese, Level 2 Part 2, Textbook* (3rd ed.). Boston: Cheng & Tsui Company.
- Lu, J. & Guo, R. (1998). Hanyu yufa yanjiu suomianlin de tiaozhan [The challenges faced by Chinese grammatical studies]. *Chinese Teaching in the World*, 4, 3-11.
- Lynch, T. & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4, 221–50.
- Matthiessen, C., & Thompson, S.A. (1988). The Structure of Discourse and ‘Subordination’. In: Haiman J. & Thompson, S. A.eds., *Clause Combining in Grammar and Discourse*, Amsterdam: John Benjamins, 275-329.
- Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *SSLA*, 20, 83-108.
- Norris, J. M. (2009). Task-based teaching and testing. In: M. Long and C. Doughty (Eds.), *Handbook of language teaching* (pp. 578-594). Cambridge: Blackwell.

- Norris, J. M. (2011). *Classroom realities and their implications for TBLT research*.
Presentation at the 4th International Conference on Task-Based Language
Teaching, University of Auckland, New Zealand.
- Norris, J. M. (2015). Discriminant analysis. In L. Plonsky (Ed.), *Advancing quantitative
methods in second language research* (pp. 309-332). New York: Routledge.
- Norris, J. M. & Ortega, L. (2003). 'Defining and measuring SLA. In: C. J. Doughty & M.
H. Long (Eds.), *The Handbook of Second Language Acquisition*. Oxford:
Blackwell.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in
instructed SLA: The case of complexity. *Applied Linguistics* 30: 555-578.
- Norris, J. M., & Ortega, L. (2012). Assessing learner knowledge. In: S. M. Gass & A.
Mackey (Eds.), *Handbook of second language acquisition* (pp. 573-589). New
York: Routledge.
- O'Grady, W. & Dobrovolsky, M. (1989). *Contemporary Linguistics: An Introduction*.
New York: St. Martin's Press.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second
Language Acquisition*, 21, 109-148.
- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-
task planning, In: R. Ellis (Ed.), *Planning and task performance in a second
language*, (pp. 77-109). Amsterdam: John Benjamins.
- Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002). *An investigation of elicited
imitation tasks in crosslinguistic SLA research*. Paper presented at the Second
Language Research Forum, Toronto, October 4-6.

- Ortega, L., Iwashita, N., Rabie, S., & Norris, J. M. (in preparation). *A multilanguage comparison of measures of syntactic complexity*. Honolulu: University of Hawai'i at Mānoa, Second Language Teaching and Curriculum Center.
- Pallotti, G. (2009). CAF: Defining, Refining and Differentiating Constructs. *Applied Linguistics*, 30(4), 590-601. Oxford University Press.
- Pan, W. (2002). The contrast of Chinese and English in past 100 years. *Chinese Teaching in the World*. 59(1). 60-86.
- Picam T., L. Halliday, N. Lewis, & L. Morgenthaler. (1989). Comprehensible outputs as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*. 11(1). 63-90.
- Pu, M-M. (1997). Zero anaphora and grammatical relations in Mandarin. In: T. Givón (Ed.) *Grammatical Relations: A Functionalist Perspective* (pp. 281-321). Philadelphia: Benjamins.
- Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. UK: Longman.
- Robinson, P. (2001). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In Robinson, P. (Ed.) *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3-37). Amsterdam: John Benjamins.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching (IRAL)*, 45(3), 161-177.

- Robinson, P., Cadiemo, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics*, 30(4), 533-554.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.
- Shen, X. (1988). *Zhongguo juxing wenhua* [Chinese sentence Patterns] China: The Northeast Normal University Press.
- Shi, D. (1993). Discourse Topic Continuity and Syntactic Reduction. *BLS 19*, 313-322.
- Skehan, P. (1996). Second language acquisition and task-based instruction. In: J. Willis, & D. Willis (Eds.). *Challenge and change in language teaching* (pp. 17-30). Oxford: Heinemann.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task based instruction. *Language Teaching*, 36 (1), 1-14.
- Spilsbury, G., Stankov, L., & Roberts, R. (1990). The effects of test's difficulty on its correlation with intelligence. *Personality and Individual Differences*, 11(10), 1069–1077.
- Tabachnick, B. G., and Fidell, L. S. (2013). *Using Multivariate Statistics*, 6th ed. Boston: Pearson.
- 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 Publishing Company.

- Tai, J. H-Y. (1993). "Iconicity: Motivations in Chinese Grammar." In: M. Eid & G. Iverson (Eds.), *Principles and Prediction: The Analysis of Natural Language* (pp. 153-174). Amsterdam: John Benjamins Publishing Company,
- Tai, J. H-Y. (2002). Temporal Sequence in Chinese: A Rejoinder. In: L. Y. Su, Chinfa L. & K. Chui (Eds.), *Form and Function: Linguistic Studies in Honor of Shuanfan Huang* (pp. 331-351). Taipei: Crane Publishing Company.
- Thompson, S. A. & Longacre, R. E. (1985). Adverbial Clauses. In: T. Shopen, (Ed.), *Language Typology and Syntactic Description*, Cambridge: Cambridge University Press, 171-234.
- Towell, R. & Hawkins, R. (1994). *Approaches to second language acquisition*. Clevedon: Multilingual Matters.
- Tsao, F. (1979). *A Functional Study of Topic in Chinese: the First Step towards Discourse Analysis*. Taiwan: Student Book.
- Tsao, F. (1990). *Sentence and clause structure in Chinese: a functional perspective*. Taiwan: Student Book.
- Tomlin, R. (1987). Linguistic reflections of cognitive events. In: Rossell Tomlin (Ed.) *Coherence and Grounding in Discourse*. Pp. 455-480. Amsterdam: John Benjamins.
- Tomlin, R. (1989). *Focal attention, activated memory and the referential management of discourse production*. Paper presented at the Banff Annual Seminar in Cognitive Science, Baff, Alberta.
- Tomlin, R. & Pu, M. M. (1991). The management of reference in Mandarin discourse. *Cognitive Linguistics* 2(1): 65-93.

- Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal*, 92(2), 214-231.
- Wang, L. (1957). *Hanyu yufa gangyao* [*Chinese Mandarin Grammar Outline*]. China: Shanghai Xinzhishi Chubanshe [Shanghai New Affairs Press].
- Wang, Y. (1990). Hanying yuyan hongguan jiegou de qubie tezheng [the differentiating feature of Chinese and English macro structure]. China: *Journal of Foreign Language*, 6.
- Wang, Y. (1992). Hanying yuyan hongguan jiegou de qubie tezheng (xu) [the differentiating feature of Chinese and English macro structure (continued)]. China: *Journal of Foreign Language*, 5.
- Wang, H. (1994a). Cong zi he zizu kan ci he duanyu—ye tan hanyu zhong ci de huafen biao zhun [On word class categorizing: look at word and phrases based on character and characters]. *Chinese Language*, 2.
- Wang, H. (1994b). Hanyu de tedian yu yuyan de pubianxing—cong yuyan yanjiu de lizudian kan zhongxi yinxi lilun de fazhan. [The uniqueness of Chinese and generality of languages: viewing the Chinese and western phonological system development from the standpoint of linguistics research]. In: Department of Chinese Language and Literature, Peking University (eds.), *Zhuyi Erji*. [*Second collection of Jade*]. Beijing: Peking University Press.
- Wang, H. (2000). Hanyu yufa de jiben danwei yu yanjiu celve [The basic unit of Chinese and research strategy], *Language Teaching and Linguistic Studies*, 2, 10-18.

- Wang, L. (1957). *Hanyu yufa gangyao*. [Chinese Mandarin Grammar Outline]. Shanghai: Shanghai Xinzhishi Chubanshe [Shanghai New Knowledge Press].
- Wang, Y. (1990). Hanying yuyan hongguan jiegou de qubie tezheng [The differentiating feature of Chinese and English macro structure]. China: *Journal of Foreign Language*, 6.
- Wang, Y. (1992). Hanying yuyan hongguan jiegou de qubie tezheng (xu) [the differentiating feature of Chinese and English macro structure (continued)]. China: *Journal of Foreign Language*, 5.
- Wigglesworth, G. (1995). *The effect of planning time on oral test discourse*. Paper presented at the 17th Language Testing research colloquium, Baltimore, MD.
- Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14(1), 85-106.
- Wolfe-Quintero, K, Inagaki, S., & Kim. H-Y. (1998). *Second language development in writing: measures of fluency, accuracy and complexity*. Honolulu: University of Hawaii at Manoa, Second Language Teaching and Curriculum Center
- Wu, W. (1985). Yingyu zhong mouxie xingshi kending yisi fouding de juzi de fanyi [The translation of some formally negative but semantically positive English sentence]. *Chinese Translators Journal*, 12, 34-35, Beijing.
- Xu. T. (1991). Yuyi yufa chuyi [On sememic syntax]. *Language Teaching and Linguistic Studies*.3.
- Xu, T. (1994). Zi he hanyu yanjiu de fangfalun [Methodology on character and Chinese research]. *Chinese Teaching in the World*. 3.

- Xu, T. (1997). *Yuyanlun—Yuyixing yuyan de jiegou yuanli he yanjiu fangfa* [On Language: the structure rules and research methods on sememic language]. China: Northeast Normal University Press.
- Xu, T. (2005). Zi benwei he yuyan yanjiu [Character-oriented and Language research]. *Language Teaching and Linguistic Studies*, 6,1-11.
- Yuan, F. (2009). Measuring learner language in L2 Chinese in fluency, accuracy and complexity. *Journal of the Chinese Language Teacher Association (JCLTA)*, 44 (3): 109-130.
- Yuan, F. & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 24 (1), 1-28.
- Yuan, F. (2009). Measuring learner language in L2 Chinese in fluency, accuracy and complexity. *Journal of the Chinese Language Teacher Association*, 44(3), 109-130.
- Zhu, D. (1985). *Yufa Dawen*. [The Questions and Answers on Grammar]. Beijing: The Commercial Press.
- Zhou, Y. (2012). *Willingness to communicate in learning Mandarin as a foreign and heritage language*. (Doctoral dissertation). Retrieved from Ann Arbor, MI: ProQuest. (Identifier: ISBN9781267500793)
- Zhou, Y. and Wu, S.-L. (2009) *Can elicited imitation be used for the measurement of oral proficiency in L2 Chinese? A pilot study*. Unpublished manuscript, Department of East Asian Languages and Literatures, University of Hawaii at Manoa, Honolulu, U.S.A.

ⁱ Zhou and Wu kindly allow the EI test to be adopted as a part of TW&ST.

ⁱⁱ I am deeply indebted to Richard Medina, an IT Specialist at the Center for Language & Technology (CLT) in the College of Language, Linguistics, and Literature (LLL) of the University of Hawaii at Manoa, without whose great technological assistance this web-based test could not have been possibly developed.