

THE ROLE OF INNATE GRAMMAR AND INPUT IN THE ACQUISITION OF
CHINESE RELATIVE CLAUSES

by

Hui-Yu Catherine Huang

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Hui-Yu Catherine Huang, titled The Role of Innate Grammar and Input in the Acquisition of Chinese Relative Clauses and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

_____ Date: 04/18/2016
Heidi Harley

_____ Date: 04/18/2016
Simin Karimi

_____ Date: 04/18/2016
Sandiway Fong

_____ Date: 04/18/2016
LouAnn Gerken

_____ Date: 04/18/2016
Rudolph Troike

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: 04/18/2016
Dissertation Director: Heidi Harley

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Hui-Yu Catherine Huang

ACKNOWLEDGEMENTS

Writing the acknowledgements is probably the most nervous moment for me since I started working on my PhD research in linguistics at University of Arizona for all sorts of reasons. That being said, I actually can't wait to say "*Thank you a gazillion times!*" to all those who I always want to thank but maybe forgot to thank timely and probably will still accidentally miss to thank here, which, you know I never meant to. And as always, all errors are mine with respect to this dissertation.

First, I would like to thank my advisor, mentor and friend Heidi Harley. Ever since I met her, I have been deeply impressed again and again by her penetrating thoughts (of almost everything), her constant encouragements to students and her warm happy smiles. Whenever I had questions and went to her office, I always left the office being equipped with more confidence and smiles that allow me to pursue new and enlightening ideas without fear any more. I wish I could find the best word or phrase to describe how I thank her. But I remember that the time that I have spent with her, even on discussing hard-to-explain linguistic phenomena, is always always full of enjoyment. It has been such a delightful pleasant great experience to work with her!! Simin Karimi has been a wonderful mentor and friend as well. Her support for students of the entire department is just amazing. For example, she even let us stay in her house to have a seminar with Noam Chomsky to the exclusion of all faculty members, which I always think amusing and secretly giggled at a lot (with a grateful attitude). But because of her endless support and guidance, I got so nourished both academically and mentally in this department. Sandiway Fong plays a vital role in leading me to the realm of computational linguistics. I still remember when I first took his course, I felt so frightened because I had so many difficulties understanding the course materials. But the first time when I entered his office to ask questions, he was unbelievably friendly and even brewed a cup of coffee or tea for me and just started answering my questions (with smiles!). He never frowned when I asked questions (including very stupid ones), so that really encourages me a lot (to ask more stupid questions? JK.), and I really really thank him for guiding me throughout these years. Now, were it not LouAnn Gerken I would not have started thinking about some of the important issues in this dissertation. She probably didn't believe it, but taking her class has changed my view of language acquisition a lot (in a good way of course). I really enjoyed the discussions with her, because those were very enlightening and interesting! Last but not least, being able to meet Rudy Troike in Tucson and to become good friend with him is an amazing experience in life that probably only he and I could understand. I so much appreciate all the time that he has spent with me and all the linguistics resources that he has shared with me. I am so so grateful for his constant encouragements, support, warmth and love.

I also want to take this opportunity to thank Cecile McKee, who kindly took me under her wing for my first three years in the program. I have always admired her intelligence and diligence to work on both research and public outreach. Were it not because of her guidance, I would not have been able to accomplish any experimental

study that later on became an important part of this dissertation. I owe so much to her, and I only wish I could work as hard as she does and match her standards one day.

The other faculty members are to be equally thanked since they are as knowledgeable, supportive, passionate, and approachable as my lovely committee members. Amy Fountain is one of the most selfless people that I have ever met. She is always there whenever we need her. Being able to work with her as a TA also helps me to gain sufficient confidence and experience to be able to stand in front of many students who are ready to challenge. I also like to hang out with Andy Wedel and Adam Ussishkin (and their cat), especially when there is a party at their house (Don't get us wrong; it is usually a party for prospective students and our department always has a good reputation for being an amazing host!). I also enjoyed the occasional discussions on syntax and semantics with Massimo Piattelli-Palmarini, Tom Bever, Andrew Carnie, Tyler Peterson, Robert Henderson, Andy Barss in SynSalon meetings, whose insights and knowledge always filled me with awe. I want to thank Natasha Warner as well. Her statistics course has helped me to embark on a research journey that is more fruitful than I could have ever imagined. I also thank Diana Archangeli, Diane Ohala, Mike Hammond, Janet Nicol, Stacey Oberly for their timely guidance and care on different occasions.

I would also like to acknowledge The Affiliated Children's Day Care Center of Taipei Mingde Elementary School, YingXi Kindergarten in Hsinchu, and The Affiliated Children's Day Care Center of Tainan WenNei Elementary School for allowing me to work with them and to conduct the experiments with their lovely students. I also thank Mia Vento, Sihwei Chen, Trina Kao, Estelle Chieh Chih Cheng, Xiumei Chen and Xue Du for their assistance with the experiments. I am very grateful for all their help!

The acknowledgement section would be seriously incomplete without mentioning the friends and colleagues that I have been indebted to so much because of the joy, happiness, and support that they have brought to me. I want to thank Elly Zimmer, Yan Chen, Maureen Hoffmann, Patrick Thomas and David Ellison for their encouragements, help, and friendship, especially during my most difficult times. I also want to thank Becky Sharp, Gus Hahn-Powell, Samuel Johnston, Kristen Howe and, especially David Ellison, for helping me with the programming assignments; without their help, I could have died (no kidding!). I also enjoyed discussing Korean syntax with Jaehoon Choi and Hyun Kyoung Jung in the Hive cubicle. HK has been the best mentor ever! I will miss my mentee a lot, Rachel Brown, because we have had so much fun chatting with each other in The Scented Leaf. I also love the time that I spent with Chen-Chun E with coffee from Caffè Luce (my favorite!), because we could talk about life and syntax forever! I will miss Becky Sharp and Samuel Johnston a lot for their never-ending love and friendship; I will miss Rolando Coto's enthusiasm about everything; I will miss Samantha Wray's bravery to explore things that are untouched; I will miss the time when I played badminton with Zechy Wong, Jianrong Yu, and Grace (We need to play together again so keep in touch!). I will miss the time that I spent together with Ryan Smith and Colleen Patton in our syntax

lab, where we discussed syntax, language acquisition (Wait, isn't it a "syntax" lab? Whatever.), language revitalization, Colleen's cat, life, mysteries, music, etc. Of course I will also miss the lovely and passionate Megan Figueroa, the smart Shiloh Drake, the sweet and warm Nick Kloehn, the funny Josh Meyer (BTW., I'll remember you and *some other guys* "broke" my durian), the tall calm Dane Bell and the hilarious super talented artist Emily Monken Bell, the cool elegant Mary-Caitlyn Valentinsson, the amazing Elise Bell, the curious Stanley Donahoo, the shy Bill Cotter, the "what's up" Adam King, the caring Megan Stone, the warm-hearted Alex Trueman, the most mysterious BJB (i.e., Bryan Gordon), the very considerate Julia Fisher, and so many wonderful others for whom I'm sorry I'm now running out of proper adjectives to describe you — Lio Mathieu, Deniz Tat, Sylvia Reed, David Medeiros, Dan Brenner, Jae-Hyun Sung, Brecht Welch, Jeffrey Punske, Kevin Schluter, Andrea Davis, Michael Sayle, Keisha Josephs, Priscilla Shi, Patricia Lee, Mohsen Mahdavi Mazdeh, Ben Martin, Isabel McKay, Rana Nabors, Noah Nelson, Ammon Pike, Trevor Sullivan, Joseph Dupris, Mia Vento, Margarita Beltran Tatyana G Slobodchikoff, Amy LaCross, Ryan Nelson, Ethan Dickinson, Yuan-Lu Chen, Mandy Liu, Jennifer Chen, Mohammed Al-Sammak, Skye Anderson, Gregory Anderson, Zach Chapman, Monica Larcom, Ana Gonzalez, Christina Newhall, Jonathan Geary, Shannon Grippando, Alan Hogue, Sunjing Ji, Keisha Josephs, Jaime Parchment, Megan Kittleson, George-Michael Pescaru, Jessica Ray, Leah Rice, Kara Hawthorne, Jessamyn Schertz ...Ahh...why do you all those that I have mentioned above make it so hard for me to not miss you? (Ok, I agree that it's because you all are awesome, including those that I don't know why I forgot to mention...).

I will not forgive myself if I do not thank my dear staff in the department as well. Thank you Marian Wiseley, Kimberley Young, Shayna Walker! There would be no chance for me to survive in the program without the help and support from all of you. You all rock!

I have also benefited from discussions with scholars on several different occasions. My work and academic progress has been influenced by their help/comments/suggestions/insights/criticisms/support. I am extremely grateful to Dustin Chacón because the discussions with him at the final stage of writing this dissertation have helped me to clarify my thoughts and make progress a lot. The other scholars that I would like to thank also include: Jeffrey Lidz, Barbara Pearson, Laura Wagner, Colin Phillips, Noam Chomsky, Norbert Hornstein, Aya Meltzer-Asscher, Jidong Chen, Yi-Ching Su, Chun-Chieh Hsu, James Huang, Dana McDaniel, Jaklin Kornfilt, Jason Ginsburg, Bodo Winter, Hadas Kotek, Michael Yoshitaka Erlewine, Maryia Fedzechkina, Gregory Scontras, etc...

I also want to thank those faculty members who had guided me before I entered the PhD program at UA. I want to thank Anjum Saleemi, because he is the key person who has inspired me the most to want to study linguistics when I was an undergraduate. I thank Lyih-Peir Luo for introducing me to syntax trees; at first the trees really confused me. I thank Loren Billings for encouraging me to pursue advanced study in linguistics. Furthermore, I want to thank Chen-Sheng Luther Liu,

Tzong-Hong Jonah Lin, and Hui-Chuan Hsu in particular for their guidance in linguistics when I was studying in the MA program in Taiwan. Without their nourishment and encouragement during those years, I would not have been able to come to this amazing UA linguistics program to study, let long getting a PhD. Thank you thank you all!!

I also want to thank all my dear friends who have played an important role in my life. Thank you Crystal Hsu, Hazel Chen, Jessamine Lee, Trina Kao, Tina Chou, WanWan Chuang, Michelle Chang, Hsin-Fen Chang, Luis Carlos Davis, Pei-Yi Katherine Hsiao, Siwei Chen, Ching-Yu Helen Yang, Hui-Chin Tsai, Chaokai Shi, and Yi-Hsun Chen. It is because of your love and support that have made this achievement possible. Thank you and I love you all.

I thank my cousins and my brothers whom I have bombarded with tons of Chinese examples to inquire thy intuitions throughout these years. Thank you especially to my dear cousin Cheng-Ying Huang to let me stay with him in Portland to breathe (and to drink boba tea and to eat hot pot, etc.) when I got so tired of writing. Thank you my best cousin Chia-Chi Wang for sending me warm words timely when I was frustrated by my slow progress in research.

Finally, to my parents and my brothers, thank you for always loving me and I love you always. And I thank my Lord for being my best companion and best friend throughout these years.

TABLE OF CONTENTS

LIST OF TABLES	11
LIST OF FIGURES	12
LIST OF ABBREVIATIONS	13
ABSTRACT	14
Chapter 1: Overview	15
1 Innate Grammar and Input in Language Acquisition	15
2 Relative Clauses and Language-Specific Properties	16
3 Hypotheses	19
4 Conclusion	21
Chapter 2: Relative Clauses in the UG-Based Framework	25
1 Introduction	25
2 Gapped Relative Clauses	31
2.1 The Distribution of Gapped Relative Clauses.....	31
2.2 Previous Analyses of Gapped Relative Clauses	35
3 Resumptive Relative Clauses	44
3.1 The Distribution of Resumptive Relative Clauses	45
3.2 Previous Analyses of Resumptive Relative Clauses	57
4 Chinese Relative Clauses	59
4.1 Chinese Gapped Relative Clauses.....	60
4.2 Chinese Resumptive Relative Clauses	67
4.3 Aoun & Li (2003)	72
5 Conclusion	77
Chapter 3: Chinese-Speaking Children’s Comprehension of Gapped and Resumptive Relative Clauses	78
1 Introduction	78
2 Past Research on the Acquisition of RCs	80
2.1 Lee (1992)	82
2.2 Su (2004).....	87
2.3 Interim Summary	90
3 Experimental Research	92
3.1 Subjects	92
3.2 Tasks and Materials	93
3.3 Data Coding	102
4 Results and Analyses	103
5 General Discussion	107
6 Conclusion	109
Chapter 4: Non-UG Approaches to Language Acquisition	111
1 Introduction	111
2 Frequency in the Acquisition of Syntax	115
2.1 Kidd et al. (2007)	117

3 Bayesian Models of Language Learning	120
3.1 Bayes' Theorem	122
3.2 The Background of Bayesian Models of Cognition.....	122
3.3 Bayesian Models of Language Learning	128
4 Conclusion	136
Chapter 5: A Corpus Analysis of Chinese Relative Clauses and Their Implications for Non-UG Statistical Approaches	138
1 Introduction	138
2 Previous Corpus Analyses of Chinese Relative Clauses	139
2.1 Corpus Studies of the Chinese Treebank	139
2.2 Child Speech Corpus Analysis.....	145
3 The Current Corpus Study: Chinese Treebank	149
3.1 The Corpus: The Penn Chinese Treebank and Relative Clauses	149
3.2 The Tool for Querying Tree Data Structures: Tregex.....	153
3.3 Extracting Relative Clauses with Tregex in the Treebank.....	157
3.4 Types of Relative Clauses.....	160
4 Results	162
5 Discussion	168
5.1 Data Within Frequency-Based Approaches	168
5.2 Data Within Bayesian Approaches	174
6 Conclusion	179
Chapter 6: The Analysis of Resumptive Pronoun and Gap Structures in Chinese Relative Clauses	181
1 Introduction	181
2 Gaps in Relative Clauses	188
2.1 Gaps and Island Conditions	192
2.2 A'-Gaps and the Weak Crossover Condition.....	206
2.3 Asymmetries Between Subject and Object Gaps in Chinese.....	209
3 Overt Resumption	216
3.1 Aoun & Li (2003): Overt Resumption in Chinese Relative Clauses	218
3.2 Overt Resumption and Island Conditions	222
3.3 Resumption and the Weak Crossover Condition	235
3.4 Overt Resumption and Reconstruction	237
4 When Gaps and Resumption in Chinese RCs Surprise Us	244
4.1 Coordination of Gapped and Resumptive Relative Clauses	244
4.2 Gaps and Reconstruction Within Islands	247
5 The Revised Proposal	250
5.1 The Analysis of Resumptive Pronoun and Gap Structures.....	250
5.2 Guillot & Malkawi (2006, 2007): Reconstruction	254
6 Conclusion	260
Chapter 7: Toward a UG-Based Processing Account of the Acquisition Pattern	263
1 Introduction	263
2 Parsing Relatives and Language-Specific Characteristics	271

2.1	Parsing Grammaticized Resumption.....	274
2.2	Parsing Ungrammatical Resumption	275
2.3	Length of Dependency and Parsing Ungrammatical Resumption	277
3	Parsing Chinese Relatives	279
4	A UG-Based Processing Proposal of Comprehension Patterns.....	286
5	Conclusions.....	290
	Appendices.....	292
	References.....	298

LIST OF TABLES

TABLE 1: RESUMPTIVE PRONOUNS IN RCS IN DIFFERENT LANGUAGES	51
TABLE 2: RESUMPTIVE PRONOUNS IN RCS IN CHINESE	70
TABLE 3: NUMBER OF PARTICIPANTS BY CATEGORY	93
TABLE 4: PERCENTAGE OF CORRECT RESPONSES BY RC TYPE AND AGE GROUP	103
TABLE 5: INFORMATION OF EACH CHILD GROUP	104
TABLE 6: FREQUENCY PATTERNS OF CHINESE RCS IN THE ADULTS' CORPORA..	143
TABLE 7: RELATIVE FREQUENCY OF RC BY RELATIVIZED POSITION IN CHILDREN AND CAREGIVERS	146
TABLE 8: RELATIVE FREQUENCY OF RC BY MODIFIED SYNTACTIC FUNCTION IN CHILDREN AND CAREGIVERS	147
TABLE 9: FREQUENCY PATTERNS OF RCS IN CHILDREN AND CAREGIVERS' SPONTANEOUS SPEECH	147
TABLE 10: TREGEX SEARCH PATTERNS FOR SUBJECT, OBJECT AND INDIRECT OBJECT RCS	162
TABLE 11: RAW COUNTS AND DISTRIBUTIONAL FREQUENCY OF 3 TYPES OF RCS	163
TABLE 12: EXAMPLES OF RCS FROM THE CHINESE TREEBANK	163
TABLE 13: RESULTS OF THE SEQUENCE “ <i>DUI</i> ‘TO’ + PRONOUN”	167
TABLE 14: RAW COUNTS AND FREQUENCY OF EXPERIMENTAL VERBS IN SINICA CORPUS	173
TABLE 15: EIGHT POTENTIAL GRAMMARS IN THE SAME HYPOTHESIS SPACE FOR RCS	178
TABLE 16: GAPS IN CHINESE ARGUMENT RCS	182
TABLE 17: RESUMPTIVE PRONOUNS IN CHINESE ARGUMENT RCS	182
TABLE 18: LEAST EMBEDDED GAPS AND ISLAND CONFIGURATIONS	205
TABLE 19: FURTHER EMBEDDED GAPS AND ISLAND CONFIGURATIONS	206
TABLE 20: GAPS AND WEAK CROSSOVER EFFECTS	209
TABLE 21: TYPES OF NPS	210
TABLE 22: LEAST EMBEDDED RESUMPTIVE PRONOUNS AND ISLAND CONFIGURATIONS	234
TABLE 23: FURTHER EMBEDDED RESUMPTIVE PRONOUNS AND ISLAND CONFIGURATIONS	235
TABLE 24: RESUMPTIVE PRONOUNS AND WEAK CROSSOVER EFFECTS	237
TABLE 25: RESUMPTIVE PRONOUNS (RPs) AND VARIABLE-BINDING (VB)	243
TABLE 26: DISTRIBUTION OF GAPS AND RESUMPTIVE PRONOUNS (RPs) IN CHINESE ARGUMENT RCS	286

LIST OF FIGURES

FIGURE 1: SAMPLE EXPERIMENTAL LAYOUT	96
FIGURE 2: MEANS FOR GROUP BY RC TYPE FOR CORRECT COMPREHENSION PERCENTAGE	104
FIGURE 3: DISTRIBUTION/TENDENCY OF CHINESE GAP/RP STRATEGY IN THE HIERARCHICAL STRUCTURE.....	287

LIST OF ABBREVIATIONS

1	First Person
ACC	Accusative Case
ASP	Aspectual Marker
CL	Classifier
COMP	Complementizer
COND	Conditional Marker
DEM	Demonstrative
FEM	Feminine
FUT	Future Marker
HAB	Habitual Aspect
INF	Infinitive
IR	Irrealis
MASC	Masculine
NEG	Negation
P	Plural
PF	Perfective Aspect
PNM	Prenominal Modifier
POSS	Possessor Marker
PROG	Progressive Marker
PTC	Particle
Q	Question Particle
REL	Relativizer
SFP	Sentence Final Particle

ABSTRACT

This dissertation concerns one of the long-term debates about whether an innate grammar, i.e. Universal Grammar, is necessary for a child to acquire a language. The researchers who do not agree that an innate grammar needs to exist argue that the input or the child's experience of language is sufficient to acquire a language. In my work, I examine the relative clause, which is a complex structure cross-linguistically, as a case study to investigate the controversy between the role of innate grammar and the role of input in language acquisition. Based on both the comprehension data of Chinese-speaking children that I collected and the statistical analyses of corpora that I completed, I suggest that a satisfactory theory of language acquisition needs to still incorporate an innate grammar in order to accommodate the various and sometimes inconsistent input properties I document. Nevertheless, the acquisition pattern cannot be explained well without a good understanding of language-specific properties and of their potential interactions with language-external factors such as the comprehension system that may be at work in acquisition.

CHAPTER 1: OVERVIEW

1 Innate Grammar and Input in Language Acquisition

This thesis concerns one of the most debated questions of how children acquire language. On the one hand, proponents of Universal Grammar (UG) argue that some innate specifications or biases of grammar must exist because the input does not provide enough information to account for the speed and accuracy with which children acquire their target grammar, including knowledge of the grammaticality and, crucially, the ungrammaticality of various complex structures — the “poverty of the stimulus argument” (see, e.g. Chomsky 1965, 1986). On the other hand, some non-UG, domain-general approaches argue that information in the input such as relative frequencies of syntactic constructions, in combination with the communicative functions they express are sufficient for children to master their target grammar (e.g., Goldberg et al. 2004; Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007; Huttenlocher et al. 2002; Theakston et al. 2004; also see Diessel 2007 for a more comprehensive survey of frequency effects in language use and development). More recently, the Bayesian inference approach to language acquisition shows that a learner can learn a grammar by using the Bayesian inference even given limited linguistic data, thus challenging the need for postulating UG (e.g., Xu & Tenenbaum 2007; Kemp et al. 2007; Perfors et al. 2006, 2011, among others).

In this work, I compare the role of innate grammar with the role of input in acquisition within the domain of relative clauses in Mandarin Chinese (henceforth Chinese). The UG-derived approach I focus on is Chomsky-style generative

grammar. I examine whether and how the proposed UG principles governing the building of hierarchical structure of sentences affect Chinese-speaking children's performance on different types of relative clauses. By contrast, I adopt a frequency-based proposal to assess the non-UG structural approach to language acquisition. If Chinese-speaking children's performance simply matches the frequency of associated constructions in the input, then the more frequently the construction is exemplified in the input, the better the child will perform on it. Furthermore, I explore how the language-specific properties, potentially part of UG, and their interactions with the processing system may contribute to Chinese-speaking children's acquisition pattern of relative clauses.

2 Relative Clauses and Language-Specific Properties

A relative clause is a clause that modifies a noun. For example, the bracketed portion in (1) is a relative clause, and the underlined part, *dog*, is the noun that it modifies. The entire noun phrase that contains a relative clause, e.g., *the dog that kissed the cat*, is referred to as a complex NP.

(1) the dog [that kissed the cat]

The relativized position can be empty or filled with a pronoun depending on the language of interest. In English, the relativized position in the relative clause is normally empty, as indicated by the underscore in (2a-b). The empty position is interpreted as if it contained the head noun. This blank is sometimes called a *gap*. As

(2a) shows, the gap can be in the subject position of a relative clause, which I call a subject RC. The gap can also be in the object position, as in (2b), i.e., an object RC.

- (2) a. the dog [that __ kissed the cat]
b. the cat [that the dog kissed __]

Other languages allow a pronoun to occur in the relativized position, such as the direct object RC in Hebrew (3). This pronoun is usually referred to as a *resumptive pronoun*.

- (3) tasimi madbeka al ha-safta [she ha-yalda menasheket
put sticker on the-granny REL the-girl kisses
ota]
her

‘Put a sticker on the granny that the girl kisses.’ (Lit. ‘Put a sticker on the granny that the girl kisses her.’)

(Arnon 2005: 44, ex. (4))

Chinese allows both the gap and the resumptive strategies in relative clause formation. The gap is used when the relativized position is a subject, as in (4a), or a direct object of a verb, as in (4b).

(4) Chinese: gapped RCs

- a. [___ xihuan Zhangsan de] na-ge nanhai
like Zhangsan REL that-CL boy
'the boy that likes Zhangsan'
- b. [Zhangsan xihuan ___ de] na-ge nanhai
Zhangsan like REL that-CL boy
'the boy that Zhangsan likes'

When the relativized position is an object of a preposition, as in (5a), or an indirect object of a verb, as in (5b), Chinese requires a resumptive pronoun.

(5) Chinese: resumptive RCs

- a. [Zhangsan dui ta changge de] na-ge nanhai
Zhangsan to him sing REL that-CL
boy
'the boy that Zhangsan sang for' (Lit. 'the boy that Zhangsan sang for him')
- b. [Zhangsan song ta shu de] na-ge nanhai
Zhangsan send him book REL that-CL boy
'the boy that Zhangsan sent books to' (Lit. 'the boy that Zhangsan sent books to him')

There are other language-specific properties such as using pre-nominal rather than post-nominal RCs that distinguish Chinese RCs from English RCs; however, I limit my discussion by concentrating primarily on the difference between the gap strategy and the resumptive strategy throughout this study.

3 Hypotheses

I tailor the propositions of different theories of language acquisition into two concrete hypotheses for testing. For the UG-derived approach to the acquisition of different types of relative clauses, I adapt Keenan and Comrie's (1977) Accessibility Hierarchy (AH) into a hypothesis for two main reasons. First, the sequence of relativized NP arguments in the hierarchy, as in (6), matches the embeddedness of different arguments in terms of the hierarchical structure, even though the AH is an implicational universal of relativizability of argument positions on the basis of typological data (where ">" originally means "more accessible to relativization than"):

(6) Accessibility Hierarchy (AH)

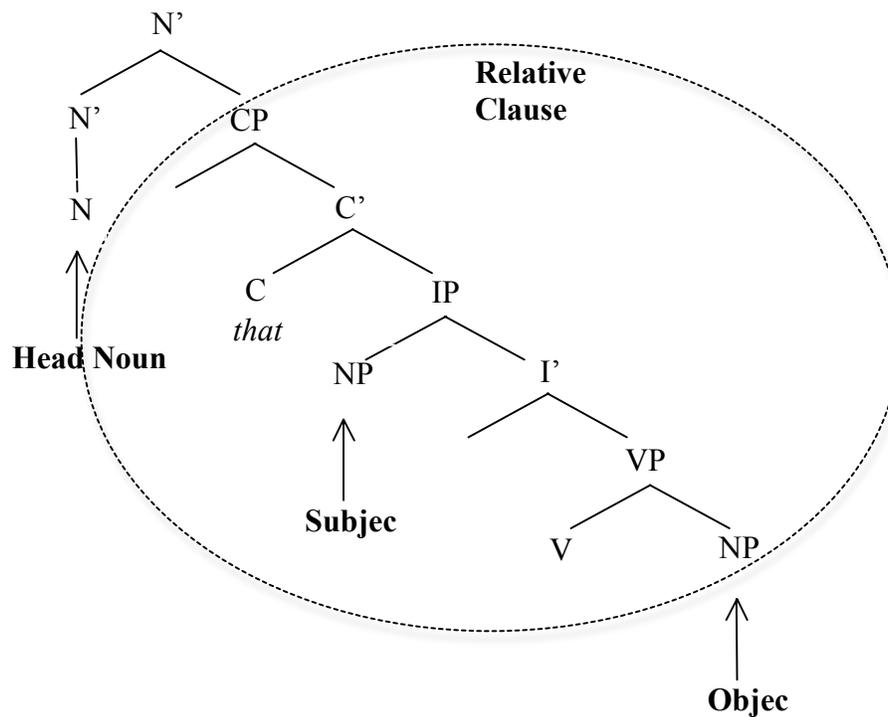
Subject > Direct Object > Indirect Object > Oblique Argument >

Genitive Argument > Object of Comparison

(Keenan & Comrie, 1977: 66)

As illustrated below, when the relativized position is a subject, it is less embedded, i.e., hierarchically closer to the head noun, than when the relativized position is an object.

(7) Structure of Relative Clause (*with irrelevant details omitted*)



Second, the order in the AH can be used to make a clear set of predictions about the order of acquisition of different types of relatives, since according to Keenan and Comrie, it “reflects the psychological ease of comprehension” (Keenan & Comrie 1977: 88). Hence, the higher the argument is, the easier that type of relative is to comprehend. I call this proposal the **Structural Distance Hypothesis** (a term borrowed from e.g., Hsu et al. 2009): More embedded positions will show worse performance/slower acquisition than less embedded positions.

For the frequency-based proposal, I examine the **Frequency of Exposure Hypothesis**: More frequent constructions/lexical items will show better performance than less frequent constructions/lexical items. In other words, frequency of relevant linguistic patterns should correlate with acquisition performance. In many cases, the

Structural Distance Hypothesis and the Frequency of Exposure Hypothesis may make the same predictions. For example, if subject relative clauses are more frequent than object relative clauses, then both the Structural Distance and Frequency of Exposure hypotheses predict earlier acquisition for the subject relative clauses. As we will see, bringing more unusual types of relative clauses into the discussion will be important in considering the predictions and effects of these and related hypotheses.

4 Conclusion

Based on the sentence comprehension data of Chinese-speaking children that I collected and the statistical analyses of corpora that I completed, I suggest that a satisfactory theory of language acquisition needs to still incorporate an innate grammar in order to accommodate the various and sometimes inconsistent input properties I document. Nonetheless, the acquisition pattern cannot be explained well without a good understanding of the language-specific properties and their interactions with the processing system. In particular, Chinese allows grammatical resumptive pronouns in the formation of certain types of relative clauses, including the oblique object relative clause (e.g., 青蛙對牠跳舞的鴨子 ‘the duck that the frog danced for [it]’). Interestingly, Chinese-speaking children appeared to comprehend this type of relative clause better than relative clauses that contain gaps, such as the direct object relative clause (e.g., 小狗親了__的貓咪 ‘the cat that the dog kissed __’). In order to explain children’s better performance on resumptive relative clauses that are lower on the Accessibility Hierarchy and also lower in frequency, I draw evidence from both the syntactic complexity of gap strategy in Chinese and the

facilitating effect of resumptive pronouns in processing. Specifically, the facilitating effect is not an invariant property of resumptive pronouns, but needs to rely on certain language-specific properties, including the grammatical status of resumptive pronouns in a given construction, different ways of embedding and so forth (e.g., Ning 2008; Keshev & Meltzer-Asscher 2016; Chacón 2015).

This dissertation is organized into six main content chapters. In Chapter 2 (*Relative Clauses in the UG-Based Framework*), I provide a literature review of analyses of relative clauses in the framework of generative grammar, and also some cross-linguistic data including relatives using gaps and relatives using resumptive pronouns. Then, I discuss the data from Chinese, and introduce some relevant properties of Chinese relatives.

In Chapter 3 (*Chinese-Speaking Children's Comprehension of Gapped and Resumptive Relative Clauses*), I conducted an experimental study to obtain the acquisition pattern of three types of Chinese relative clauses, including Subject Gapped Relatives (e.g., *The dog that kissed the cat ran away.*), (Direct) Object Gapped Relatives (e.g., *The cat that the dog kissed ran away.*) and Oblique Object Resumptive Relatives (e.g., *The duck that the frog danced for ran away.*). The results indicate that children's performance on Subject Gapped Relatives was as good as their performance on Oblique Object Resumptive Relatives, whereas their performance on Object Gapped Relatives was the worst.

In Chapter 4 (*Non-UG Approaches to Language Acquisition*), I provide a literature review of non-UG approaches to language acquisition. In particular, I

review frequency-based approaches and Bayesian approaches to the acquisition of syntax.

In Chapter 5 (*'A Corpus Analysis of Chinese Relative Clauses and Their Implications for Non-UG Statistical Approaches'*), I start with a review of some previous corpus analyses of Chinese relative clauses, including adults' and children's corpora, and then I lay out my own corpus analyses using the Penn Chinese Treebank (Xue et al. 2005). Based on the results obtained, I discuss the implications for the frequency-based and Bayesian approaches to the acquisition of Chinese relative clauses.

In Chapter 6 (*'The Analysis of Resumptive Pronoun and Gap Structures in Chinese Relative Clauses'*), I provide a comprehensive examination of argument relative clauses using gaps and resumption in Chinese in order to prepare the way for discussion of influence on acquisition patterns from language-specific properties in Chapter 7. I investigate gaps in relative clauses in terms of their sensitivity to different island constraints, and to the Weak Crossover Condition, and point out that they exhibit both movement and nonmovement properties. Then, I examine resumptive pronouns in the island conditions, the Weak Crossover Condition and their reconstruction properties in island and non-island contexts. In addition, based on coordinate structure examples that conjoin relatives using gaps and relatives using resumptive pronouns, I suggest that the syntactic rules governing gaps and rules governing resumption in Chinese relative clauses may not always be distinctive. More precisely, some gaps behave like resumptive pronouns, and are derived via base-generation.

In Chapter 7 (*Toward a UG-Based Processing Account of the Acquisition Pattern*), I review some previous important experimental studies on adults' processing of resumption dependencies where resumptive pronouns are treated as ungrammatical such as English and in languages where resumptive pronouns are considered grammatical such as Hebrew and Chinese. I explore the possible interactions between language-specific properties and the parser in the acquisition of Chinese relative clauses. Specifically, I suggest that resumption itself does not necessarily facilitate comprehension or learning, but needs to coordinate with UG-determined properties such as hierarchical structure, language-specific characteristics such as the grammaticality of structures, distributional cues and so forth in order to yield the observed results in learning.

CHAPTER 2: RELATIVE CLAUSES IN THE UG-BASED FRAMEWORK

1 Introduction

The goal of this chapter is to provide an overview of cross-linguistic data and of the UG-based theoretical analysis of relative clauses. It will equip the reader with the sufficient knowledge to engage with the theoretical and methodological decisions I make throughout the dissertation. I discuss two major issues in the UG-based analysis of relative clauses. One issue concerns the relationship between gapped and resumptive relative clauses in terms of the syntactic distribution and constraints. The other concerns the Accessibility Hierarchy (Keenan & Comrie 1977), which specifies the relativizability of argument positions. I illustrate throughout with data from various languages reported in the literature, and with data from Chinese, the language which is the topic of this dissertation.¹

In the traditional analysis, the linguistic expression of a gap in the relative clause such as t_i in (1) is treated as a variable that is bound by the *wh*-operator in the specifier of CP. The variable is created by movement of the *wh*-word from the original argument position to the A'-position.

(1) the girl [_{CP} *who*_i [_{IP} the boy likes t_i]]

In contrast to the movement analysis of gaps in relative clauses, a resumptive pronoun in a relative clause is usually taken to be a base-generated pronoun (e.g.,

¹ For expository purposes, I will postpone a detailed syntactic analysis of Chinese relative clauses to Chapter 6.

Chomsky 1977). An example of a resumptive pronoun in a relative clause is shown in the underlined element in the Hebrew example (2) (data from Chomsky 1977: 80):

- (2) ze ha-ʔiš še ʔoto raʔiti etmol
this.is the-man that him I.saw yesterday
‘this the man that I saw yesterday’

Despite the movement analysis of the gap or the non-movement analysis of the resumptive pronoun in relative clauses, relative clauses exhibiting the gap and the resumptive pronoun have similar behaviors. For example, Zaenen et al. (1981) show that whereas Swedish gapped relatives obey the Coordinate Structure Constraint proposed by Ross (1967) in that they cannot be conjoined with a clause that has no gaps, as in (3a), Swedish allows coordination of a gapped relative and a resumptive relative, as in (3b).

(3) a. *Där borta går en man_i som jag ofta tärffar
 there goes a man that I often meet
 men inte minns om Marie känner Kalle.
 but don't remember if Marie knows Kalle
 *‘There goes a man that I often meet but don't remember if Marie
 knows Kalle.’

b. Där borta går en man_i som jag ofta tärffar
 there goes a man that I often meet
 men inte minns vad han_i heter
 but don't remember what he is called
 ‘There goes a man that I often meet but don't remember what he is
 called.’

(Zaenen et al. 1981: 681)

Zaenen et al. (1981) then suggest that both the gap and the resumptive pronoun must belong to the same syntactic category; otherwise, the relative clause with a gap cannot be conjoined with the one that has a resumptive pronoun, since they would have discrete syntactic structures.

Moreover, Shlonsky (1992) indicates that resumptive relatives in Hebrew exhibit both *strong crossover* and *weak crossover* effects, which are usually considered properties associated with movement.² In addition, McCloskey (2011) observes that Irish resumptive relatives exhibit weak crossover effects. Therefore,

² See, e.g., Section 2.2 in Chapter 6 for a brief explanation of the Crossover Condition.

each of these studies independently argues that the resumptive pronoun, similar to a gap, is also a variable that is bound by an operator in the A'-position.

Since both the gap and the resumptive pronoun involve similar syntactic behaviors in relative clauses, an immediate question arises as to when a gap and when a resumptive pronoun is used. Shlonsky (1992) argues that the resumptive pronoun, compared to the gap, is a last resort in relative clause formation. Only when movement is prohibited from certain structures, for example, island conditions (e.g., Ross 1967), can the base-generated pronoun be used. However, it is not quite clear whether the resumptive pronoun should still be considered a last resort in languages where it occurs obligatorily in some non-island conditions such as Chinese, Palestinian Arabic (e.g., Shlonsky 1992) or in languages where it occurs interchangeably with a gap in non-island conditions (e.g., Shlonsky 1992 for Hebrew; McCloskey 2011 for Irish, among others). In particular, for children learning a language that allows both the gap and resumption strategies in relative clause formation, that raises several interesting and important questions. For example, can the child “recognize” which strategy is the primary one and then make the right decision when hearing or producing a relative clause? Does the child know both strategies equally well or does one strategy, e.g., the primary one, outperform the other(s) without effort? Moreover, since resumptive pronouns do not form a homogeneous group across, and even within, languages (e.g., Rouveret 2011), it would be hasty to conclude that one strategy is easier (or more difficult) to learn than

the other based on any experimental results without a careful investigation of these strategies within and across languages.³

In addition to the debate on which relative clause forming strategy precedes the other, Keenan and Comrie's (1977) Accessibility Hierarchy (AH) adds more complexity to the original debate. Based on a cross-linguistic survey of about fifty typologically varied languages, the AH specifies the relativizability of different argument positions, as illustrated in (4) (Keenan & Comrie, 1977: 66):

(4) Accessibility Hierarchy (AH)

*Subject > Direct Object > Indirect Object > Oblique Argument >
Genitive Argument > Object of Comparison*

They consider a given relative clause formation strategy to be the “primary strategy” in a language if it can be used to relativize subjects. They also introduce the Primary Relativization Constraint to accommodate the AH, as in (5) (Keenan & Comrie 1977: 68):

(5) The Primary Relativization Constraint

1. A language must have a primary RC[relative clause]-forming strategy.
2. If a primary strategy in a given language can apply to a low position on the AH, then it can apply to all higher positions.
3. A primary strategy may cut off at any point on the AH.

³ Also see Chapter 6 (in particular, Section 3) for more references about the resumption strategy.

In other words, for example, to relativize the subject position in a relative clause as in (6) would be considered “easier” and is more typologically common than to relativize the object position as in (7). As a corollary, languages that have a particular relative clause formation strategy in sentences such as (7) also necessarily use the same formation strategy in (5), but not vice versa.

(6) the dog [that __ chased the cat] (*Subject Relative*)

(7) the cat [that the dog chased __] (*Object Relative*)

One more important feature in the AH that bears directly on the discussion of relative clauses in this study is that *Oblique Arguments* refer to arguments of the main predicate, as *the boy* in *John talked to the boy*, rather than the ones that have an adverbial function like *the garage* in *John's dog slept in the garage*. This allows me to limit my discussion to relative clauses that relativize NPs expressing arguments of predicates such as oblique arguments of intransitive verbs but not adverbial functions including *location, time, reason*, and so forth.

The following sections begin with an introduction of gapped relative clauses in Section 2. Then, I introduce resumptive relative clauses with respect to their distribution and syntactic constraints in Section 3. I discuss Chinese gapped and resumptive relative clauses in Section 4. Section 5 is the conclusion.

2 Gapped Relative Clauses

In relative clauses, the relativized position can be empty, as in the English examples above (e.g., (1)) and in (8a-b) below. The empty position in the relative clause is interpreted as if it contained the head noun, and is sometimes called a gap. As (8a) shows, in English the gap can be in the subject position of the relative clause; it can also be in the object position, as shown in (8b).

- (8) a. the boy [who __ likes the girl]
b. the girl [whom the boy likes __]

In the following discussion, I focus on the argument positions that are relativized, including subject, object, oblique (or indirect) object and possessor positions.

2.1 *The Distribution of Gapped Relative Clauses*

This section is divided into two parts. In Section 2.1.1, I describe the distributions where gapped relative clauses can occur in terms of different argument positions, but I will set aside the question of relativization of adjuncts. In Section 2.1.2, I describe various island conditions where the literature has shown that gapped relative clauses are not allowed to occur.

2.1.1 Argument Positions Where Gapped Relative Clauses Can Occur

Based on Keenan and Comrie (1977), it is normally seen that the subject relativization of the highest clause uses a gap, as shown by the underlined position in (9).

(9) the boy [who likes Mary]

It is also very common to see that the direct object relativization of the highest clause uses a gap, as in (10)

(10) the boy [whom Mary likes]

The gap relative is also seen in the indirect object relativization of the highest clause, as (11).⁴

(11) the boy [that the girl gave the book to]

In possessor relativization, however, the gap is less commonly seen. Here is an example of French possessor relativization using a gap.

⁴ Despite the various grammatical positions mentioned in the AH, Keenan and Comrie (1977) indicate that not all languages distinguish all these grammatical roles, and that some languages may treat two grammatical roles similarly in terms of RC formation or in terms of other syntactic processes. As far as the current study is concerned, I do not distinguish the indirect object role from oblique object role, since Chinese requires a preposition to introduce both an indirect object (i.e., Goal argument, but also see Ning 2008 for the discussion on two types of indirect objects) and an oblique object.

- (12) la femme dont le manteau a été vole
 the woman whose the coat has been stolen
 ‘the woman whose coat was stolen’

(Keenan & Comrie 1977: 91)

In summary, relativization using a gap can occur in the subject, direct object, indirect/oblique object and possessor positions. In the next section, I describe where gapped relative clauses cannot occur.

2.1.2 Where Gapped Relative Clauses Cannot Occur

Ross (1967) observes several island-type conditions where gapped relative clauses are prohibited from being used. These include the Complex NP Condition, as in (13), Subject Island Condition, as in (14), Adjunct Island Condition, as in (15), Coordinate Structure Condition, as in (16), and the Wh-island Condition, as in (17). In each of these island-type configurations, island effects arise with gaps, as illustrated by the ungrammaticality of each example.

- (13) *The hat [_{CP} which I believe [_{NP} the claim [that Otto was wearing __]]] is red. (Ross 1967: 126)

- (14) *The teacher [_{CP} who [that the principle would fire __] was expected by the reporters] is a crusty old battleax. (Ross 1967: 241)

(15) *The deed [_{CP} which [I want to peruse that contract] [before filing away __]]
is probably a forgery. (Ross 1967: 193)

(16) *The lute [_{CP} which Henry [plays __ and sings madrigals]] is warped.
(Ross 1967: 160)

(17) ??He told me about a book [_{CP} which I can't figure out [when I should read
__]]. (Ross 1967: 27)

Moreover, the Crossover Condition indicates that no NP such as *he_i* in (18) can be “crossed” by another corefential NP *man_i*; otherwise, it would result in (strong) crossover effects.⁵

(18) *The man_i [_{CP} who he_i said ___i was tall]
(Ross 1967: 132)

To recap, the distributions where gapped relative clauses are not allowed to occur are mostly associated with island conditions or other configurations where wh-movement is prohibited, confirming that relativization, like wh-questions, is a species of A-bar movement.

⁵ The verb ‘cross’ refers to a movement that occurs during the derivation in pre-Minimalist Program frameworks. Regardless of the actual mechanism that has been revised to account for movement in syntax, the example in (18) is simply used to demonstrate a condition that bans the occurrence of a (moved) gap.

2.2 *Previous Analyses of Gapped Relative Clauses*

In this section, I discuss two major analyses of gapped relative clauses in the literature. One is the head external analysis (e.g., Chomsky 1977), which I discuss in Section 2.2.1. The other approach is the head internal analysis (e.g., Schachter 1973, Vergnaud 1974, Kayne 1994), which I discuss in Section 2.2.2. In Section 2.2.3, I review Sauerland (1998, 2000), where he suggests that both the previous analyses seem to have merit and that there can be a hybrid approach which makes use of both in different environments.

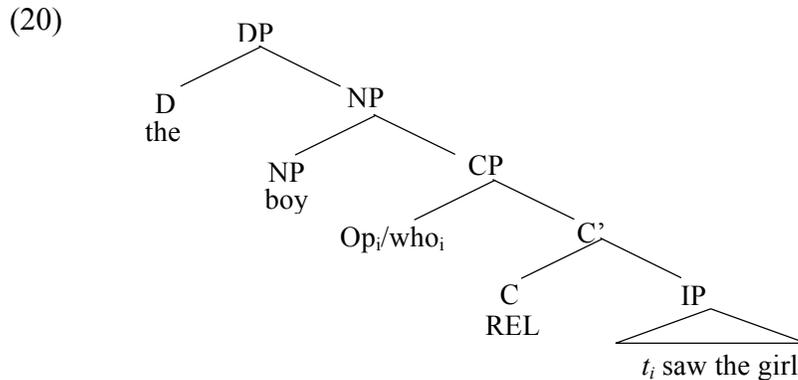
2.2.1 The Head External Analysis: Matching

Since Chomsky (1977), it has been widely accepted that gapped relative clauses are derived via A-bar movement. A relative clause that uses a wh-relative pronoun involves movement of this wh-pronoun from the base-generated argument position to the specifier of CP, as in (19a). A relative clause that lacks an overt wh-pronoun is associated with a null operator in the base position that then moves to the specifier of CP, as in (19b).

- (19) a. the boy [_{CP} who_i [_{IP} *t_i* saw the girl]]
b. the boy [_{CP} Op_i [_{IP} *t_i* saw the girl]]

Under this analysis, the trace position, i.e., *t_i*, introduces a variable that is then bound by the operator in Spec, CP. After the CP is adjoined to the head NP, i.e., *boy*, the entire complex noun phrase obtains its interpretation. In particular, this kind of

movement creates an open clause that needs to be satisfied by an entity: [λx : x saw the girl]. This open clause is then satisfied by the entity referred to by the head NP, i.e., *boy*. The syntactic structure in (20) illustrates this head external analysis of relative clauses.

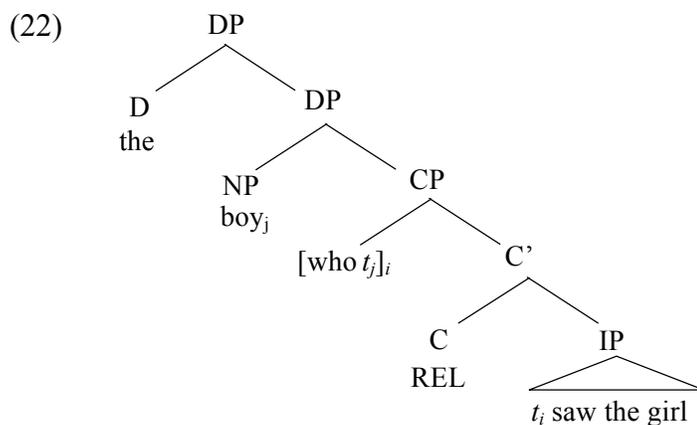


As noted above, one piece of crucial evidence for this analysis is that relative clauses exhibit similar island effects that Ross (1967) takes as diagnostic of *wh*-movement. The island effects that are seen in relative clauses include the examples listed in Section 2.1.2. Additional examples provided here reflect violation of the Complex NP Constraint (i.e., crossing the NP headed by *teacher*), as in (21a), violation of the Adjunct Island Constraint (i.e., crossing the adjunct CP headed by *if*), as in (21b), and violation of the Wh-island Constraint (i.e., crossing the *wh*-island headed by *why*), as in (21c) (data from Aoun & Li 2003: 99):

- (21) a. *the boy_i [who_i I like [_{NP} the teacher [who has taught *t_i]]]]
 b. *the boy_i [who_i I will be happy [_{CP} if you like *t_i]]]
 c. *the boy_i [who_i I wonder [_{CP} why John has taught *t_i]]]***

2.2.2 The Head Internal Analysis: Raising

Apart from the head external analysis, another line of thought is the head internal analysis (e.g., Schachter 1973; Vergnaud 1974; Kayne 1994, among others). Schachter (1973) argues that the analysis of relativization does not involve matching of the head noun with the corresponding nominal in the relative clause. Instead, he argues that the corresponding nominal in the relative clause is itself the one that is promoted, or raised, to the head noun position, which is originally an empty spot in his account. In a more modern formulation, based on Kayne's (1994) analysis, the syntactic structure for a head internal-raising analysis would be like (22):



Under such an analysis, the relative clause CP is a complement of the entire complex noun phrase DP. The DP *who boy* originates from within the IP and then moves to Spec, CP, and *boy* is subextracted and moved to the inner Spec, DP.

There are three main pieces of evidence from reconstruction effects that support the head internal analysis. The first piece of key evidence comes from idiom examples such as *make headway* (e.g., Brame 1968, cited from Sauerland 1998). There is a selectional restriction on the noun *headway* which requires it to occur as the object of the verb *make* in order to receive the non-literal reading, as illustrated by the contrast between (23a) and (23b). However, a relative clause that involves the same idiom with the displaced object is still grammatical, as in (23c). This suggests that the head noun is derived from within the relative clause.

- (23) a. We made headway.
b. *(The) headway was satisfactory.
c. The headway that we made was satisfactory.

(Schachter 1973: 31)

Another piece of evidence derives from coreferent pronouns in relative clauses. Compare non-relative examples in (24a-b) with relative examples in (25a-b):

- (24) a. John₁ thinks that Mary has an unfavorable opinion of him₁.
b. *He₁ thinks that Mary has an unfavorable opinion of John₁.

- (25) a. The opinion of him₁ that John₁ thinks that Mary has is unfavorable.
- b. *The opinion of John₁ that he₁ thinks that Mary has is unfavorable.

(Schachter 1973: 32)

In non-relative examples, the ungrammaticality of (24b) is due to violation of Condition C since a co-referent pronoun *he* c-commands the R-expression *John*. However, in (25b) the co-referent pronoun *he* does not c-command the R-expression *John₁* in the surface structure, and yet it is still ungrammatical. Notice, however, that (25b) would constitute a Condition C violation if the noun phrase *opinion of John* originated from the gap position in the relative clause, i.e., between the verbs *has* and *is*. Hence, the evidence of Condition C reconstruction indicates that the head noun originates from within the relative clause.

One more piece of evidence for the head internal analysis is related to the fact that material internal to the relative clause can take scope over a part of the external head. According to Sauerland (1998, 2000), the preferred interpretation of (26a) is the one where *need* takes scope over *many*. The paraphrase that he provides is: *Gina needs a number of books for vet school that is so large that no linguist would read that many books*. Similarly, the preferred interpretation in (26b) is also one in which the relative clause internal modal *can* takes scope over *few* in the external head. This is paraphrased as: *There is a number such that Mary can take n-many drinks, but she shouldn't even have n-many drinks* (data from Sauerland 1998: 64):

- (26) a. No linguist would read the many books Gina will need for vet school. (need >> many)
- b. Mary shouldn't even have the few drinks that she can take. (can >> few)

2.2.3 Both Analyses Seem Correct

Sauerland (1998, 2000) argues that both the head external and head internal analyses are required in the analysis of English relative clauses because of several discrepancies among relative clauses in terms of reconstruction. For example, (27a) does not show Condition C reconstruction, since it does not exhibit Condition C effects. By contrast, (27b) allows reconstruction for variable binding, since the pronoun *his* can be bound by the quantifier *everybody*.

- (27) a. The relative of John_i that he_i likes lives far away.
- b. The relative of his_i that everybody_i likes lives far away.

(Sauerland 2000: 352)

On the one hand, the absence of Condition C reconstruction in (27a) seems to indicate that no part of the external head noun comes from the internal position of the relative clause. On the other hand, the availability of reconstruction for variable binding in (27b) suggests that the external head should originate from the internal position of the relative clause.

Another divergence is between the material that is pied-piped with the relative clause and the head of relative clause concerning Condition C reconstruction, as illustrated by the contrast between (28a) and (28b).

(28) a. *I respect any writer whose depiction of John_i he_i'll object to.

(Safir 1998:34a, *cited from Sauerland 1998: 65*)

b. I respect any depiction of John_i he_i'll object to.

(Sauerland 1998: 65)

In (28a), the pronoun *he* cannot be coreferent with the R-expression *John* that is within the pied-piped material of the relative operator, since, presumably, the pronoun c-commands the relative clause internal trace position. In (28b), the R-expression *John* is part of the external head, rather than of the relative operator. In such case, coreference between the pronoun and the R-expression is allowed.

In view of these seeming discrepancies, Sauerland follows Carlson (1977), arguing that relative clauses are structurally ambiguous at LF between the matching and raising analyses. The difference lies in the Condition C effects: when Condition C emerges, the raising analysis applies.⁶

⁶ For example, the interpretation for sentence (i) below will not violate Condition C when it has a matching structure, as in (ii), but will violate Condition C when it has a raising structure, as in (iii) (adapted from Sauerland 1998: 67):

- (i) the picture of John_i he_i likes.
- (ii) the picture of John_i λx he_i likes [x] (*matching*)
(head)
- (iii) *the [picture of John_i] λx he_i likes [x, picture of John_i] (*raising*)
(head)

Sauerland uses three cases to support the idea of structural ambiguity in English relative clauses. Specifically, this hybrid approach predicts that, when the head internal analysis is enforced by an independent interpretive constraint, one should observe Condition C effects. The first case involves variable binding. In (29a), the relative clause head includes both an R-expression *John* and a bound variable pronoun *her*. In addition, the pronoun *he* c-commands the trace position in the relative clause. Under such circumstance, variable binding forces the head internal analysis, and therefore Condition C effects occur. *John* cannot be coreferent with the pronoun *he* that c-commands it in its base-position inside the relative clause. This can be contrasted with the grammatical sentence in (29b), where the R-expression is replaced with a pronoun and Condition C does not apply.

- (29) a. *The letters by John_j to her_i that he_j told every girl_i to burn were published.
- b. The letters by him_j to her_i that he_j told every girl_i to burn were published.

(Sauerland 1998: 71)

On the other hand, the interpretation for sentence (iv) will rule out a matching analysis, as in (v), since it contains a variable that needs to be bound by a quantifier inside the relative clause, as in the raising structure (iv) (adapted from Sauerland 1998: 67):

- (iv) the picture of himself_i everybody_i likes
- (v) *the picture of himself_i λx everybody_i likes [x] (*matching*)
(head)
- (vi) the λx everybody_i likes [x_i picture of himself_i] (*raising*)
(head)

The second case is the use of idioms in the relative clause. Sauerland points out that Munn (1994) already observes the reemergence of Condition C effects in examples like (30a). The idiom reading of *take picture of* forces the head internal analysis, which causes the coreferent pronoun *he* to c-command the R-expression *Bill* in the internal relative clause trace position. This leads to a Condition C violation. When a reflexive is substituted for this R-expression, as in (30b), Condition C does not apply, so no ungrammaticality occurs (Sauerland 1998: 71):

- (30) a. *the picture of Bill_i that he_i took (Munn 1994: (15c))
b. the picture of himself_i that Bill_i took

One last case that forces the head internal analysis is when the material in the relative clause scopes over the part in the external head. As shown in (31), *need* scopes over *many*, and Condition C effects emerge (Sauerland 1998: 71):

- (31) *The many books for Gina_i's vet school that she_i needs will be expensive
(need >> many)

To sum up, in Section 2, I reviewed the typical distribution of gapped relatives, the common conditions that prevent gaps from occurring, and three major theoretical analyses within the generative framework. In the following section, I review resumptive relatives.

3 Resumptive Relative Clauses

I adopt Sells' (1984) distinction between *true resumption* and the so-called *intrusive resumption*. In particular, true resumption arises as a grammar option, while intrusive resumption is not licensed by grammar but serves as a last resort device to salvage sentences with processing complexity. Resumption in English is an example of this latter type, as it is exclusively used to save structures that would otherwise incur grammatical violations. As observed by Ross (1967), relative clauses that use a gap exhibit island effects, but relative clauses that use a resumptive pronoun do not. He shows that a dialect of English has relatives using a resumptive pronoun without violation of island constraints, such as in the Complex NP Condition, as in (32a), the Sentential Subject Island, as in (32b), and the Adjunct Island Condition, as in (32c).

- (32) a. All the students who the papers which they submitted were lousy
I'm not going to allow to register next term.
- b. Palmer is a guy who for him to stay in school would be stupid.
- c. King Kong is a movie which you'll laugh yourself sick if you see it.
- (Ross 1967: 433)

In this chapter (and also Chapter 6), I focus on the former type, true resumption.⁷

⁷ But note that this term should not be misunderstood as the same one used by Aoun et al. (2001), as it always refers to resumption that is derived via base-generation in their study, in contrast to the other term 'apparent resumption,' which always involves movement. To clarify, the resumption that is examined throughout this chapter refers to the one that arises as a grammar option, and whether it involves base-generation or A'-movement depends on its

3.1 *The Distribution of Resumptive Relative Clauses*

This section consists of three parts. In Section 3.1.1, I illustrate where resumptive relative clauses can occur in terms of the argument position, in non-island contexts. In Section 3.1.2, I discuss where resumptive relative clauses can occur with respect to island conditions. In Section 3.1.3, I discuss where resumptive relative clauses typically do not occur.

3.1.1 Argument Positions Where Resumptive Relative Clauses Can Occur

Languages that use grammatical resumptive pronouns to form relative clauses normally allow the resumptive pronoun to occur in various argument positions. Here is a list of argument positions with permitted resumptive relative clauses in different languages.

First, the direct object position (of the highest clause) can be relativized using the resumptive pronoun in Hebrew (33), in Irish (34) and in Standard Arabic (35).

(33) ha-ʔiš še- raʔiti (ʔoto)
the-man that-(I) saw (him)

‘the man that I saw’

(Shlonsky 1992: 444)

syntactic behavior and interpretive properties in each different language that are to be discussed

(34) an ghirseach ar ghoid na síogaí í
 the girl COMP_{pro} stole the fairies her
 ‘the girl whom the fairies stole’ (McCloskey 2011: 114)

(35) ʔaʕrifu l-mumaθilata allati sayuqabilu-ha saami
 know.1S the-actress that FUT.meet-her Sami
 ‘I know the actress that Sami will meet.’ (Aoun et al. 2010: 172)

Second, the resumptive pronoun is used in oblique/indirect object relatives in Hebrew (36), in Irish (37), in Standard Arabic (38), and in Welsh (39).

(36) ha-ʔiš še- xašavti ʕal-*(av)
 the-man that-(I) thought about-(him)
 ‘the man that I thought about’ (Shlonsky 1992: 445)

(37) an fear a raibh mé ag caint leis pro
 the man COMP_{pro} was I talk-PROG with-3S-MASC
 ‘the man that I was talking to’ (McCloskey 2011: 77)

(38) qaraʔna (l-)kutuba (allati) ʔaxbarana ʕan-ha kariimun
 read.1P (the-)books (that) told.3S.MASC about-them Karim
 ‘We read the books that Karim told us about.’ (Aoun et al. 2010: 172)

- (39) y merched y siaradodd y dynion am-danynt
 the girls PTC spoke the men about-them
 ‘the girls that the men spoke about’ (Sells 1984: 132)

Third, the relativized position of possessors can be associated with a resumptive pronoun in Hebrew (40), in Irish (41), in Standard Arabic (42), and in Welsh (43).

- (40) ha-ʔiš še- raʔiti ʔet ʔišt-*(o)
 the-man that-(I) saw ACC wife-(his)
 ‘the man whose wife I saw’ (Shlonsky 1992: 445)

- (41) an bhean a raibh a mac pro breoite
 the womanCOMP_{pro} was 3S.FEM son ill
 ‘the woman that her son was ill’ (McCloskey 2011: 74)

- (42) ḥāDarna masraḥiyyatan yaʕrifu kariim mumaθθilii-ha
 saw.1P play know.3S.MASC Karim actors-it
 kulla-hum
 all-them
 ‘We saw a play that Karim knows all its actors.’ (Aoun et al. 2010: 173)

- (43) y dyn yr oedd ei fam gartref
 the man PTC was his mother at-home
 ‘the man whose mother was at home’ (Sells 1984: 132)

Fourth, the resumptive pronoun can occur in the subject position of an embedded clause when it is relativized in Hebrew (44), in Irish (45), and in Welsh (46).

- (44) ha-ʔiš še- xašavt še-(hu) melamed
 the-man that-(you.F) thought that-(he) teaches
 ʔanglit
 English
 ‘the man that you thought teaches English’ (Shlonsky 1992: 444)

- (45) na caiple sin a n-abrann sealgairí
 those horses COMP_{pro} say hunters
 go mbíonn siad ag-léimnigh agus ag-damhsa
 COMP be-HAB they PROG-jump and PROG-dance
 ‘those horses that hunters say are always jumping and dancing’
 (McCloskey 2011: 114)

(46) y dynion y dywedodd Siôn y darll-enent y
the men PTC said John PTC would-read(3P) the
llyfrau
books

‘the men that John said would read the books’

(Sells 1984: 143)

Fifth, like the subject of an embedded clause, the object of an embedded clause can be relativized using the resumptive pronoun in Hebrew (47), in Irish (48) and in Welsh (49) as well.

(47) ha-ʔiš še- xašavt še-Dani pagaš (ʔoto)
the-man that-(you.F) thought that-Dani met (him)

‘the man that you thought that Dani met’

(Shlonsky 1992: 445)

(48) an rud céanna a dtug sé orm mionnú
the thing same COMP_{pro} took he on-me swear-INF
go gcoinneoinn ceilte é
COMP keep-COND-1S concealed it

‘the same thing that he made me swear that I would keep (it) hidden’

(McCloskey 2011: 114)

(49) y llyfrau y dywedodd Siôn y darllenai'r dynion
 the books PTC said John PTC would-read the.men
 hwy
 them
 'the books that John said the men would read'

(Sells 1984: 144)

Below is the table that summarizes the observed argument positions that can be relativized using resumptive pronouns cross-linguistically.⁸ What is particularly interesting is that the distribution of resumption in these languages is the inverse of the distribution of the gap in the Accessibility Hierarchy. Specifically, resumption occurs in argument positions that are LOWER on the hierarchy, while gaps occur in positions that are HIGHER on the hierarchy.⁹

⁸ Until more data is provided, the *Embedded Clause* in Table 1 simply refers to a finite clause and does not include nonfinite clauses.

⁹ I will return to this issue again when discussing the interaction between processing resumption dependencies and language-specific properties in Chapter 7.

Table 1: Resumptive Pronouns in RCs in Different Languages

Language	Argument Positions					
	Matrix Clause				Embedded Clause	
	Subject	Object	Oblique Object	Possessor	Subject	Object
Hebrew	–	(+)	+	+	(+)	(+)
Irish	–	(+)	+	(+)	(+)	(+)
Standard Arabic	–	(+)	(+)	(+)	(+)	(+)
Palestinian Arabic	–	+	+	+	+	+
Welsh	–	–	+	+	+	+

Key: The symbol ‘–’ indicates that a resumptive pronoun is prohibited from occurring in that position; ‘(+)’ indicates that the occurrence of a resumptive pronoun is optional; ‘+’ indicates that the occurrence of a resumptive pronoun is obligatory.

3.1.2 Conditions Where Resumptive Relative Clauses Can Occur

Resumptive relative clauses across languages have been shown to exhibit no island violations in all of the island conditions that we have discussed in the gapped relative clauses. Below is a list of examples that show where resumptive relatives can occur in different kinds of island configurations cross-linguistically.

First, the resumptive pronoun in Hebrew (50), Irish (51), and Standard Arabic (52) can occur in the Complex NP configuration; no island effects occur.

- (50) raʔiti ʔet ha-yeled she-/asher dalya makira ʔet ha-ʔisha
 saw-I ACC the-boy that Dalya knows ACC the-woman
 she- ʔohevet ʔoto
 that loves him

‘I saw the boy that Dalya knows the woman who loves him.’

(Borer 1984: 221)

- (51) Sin madadh nachN bhfaca mé ariamh
 that a.dog NEG-COMP saw I ever
 asal aL bheadh _ chomh mór leis
 a.donkey would.be as big with.it

‘??That’s a dog that I have never seen a donkey that was as big as it.’

(McCloskey 1979: 34)

- (52) ḥaDarna (l-)masraḥiyyata (allati) taʕrifuuna S-Saḥaafiyya
 saw.1P (the-)play (that) know.2P the-journalist
 llaḏi ʔaḫbara-na ʕan-ha
 that told.3S.MASC-us about-it

‘We saw a/the play that you know the journalist that told us about it.’

(Aoun et al. 2010: 174)

Second, no island effects arise with resumptive pronouns in the Coordinate Structure configuration, as in Hebrew (53).

- (53) raʔiti ʔet ha-yeled she-/ʔasher rina ʔohevet ʔoto
 saw-I ACC the-boy that Rina loves him
 ve- ʔet ha-xavera shel-o
 and ACC the-friend of-his

‘I saw the boy that Rina loves him and his girlfriend.’

(Borer 1984: 221)

Third, the resumptive pronoun in Standard Arabic is allowed to occur in the Adjunct Island configuration, as in (54).

- (54) ḥaDarna (l-)masraḥiyyata (allati) ḏahabtum lamma
 saw.1P (the-)play (that) left.2P when
 ʔaʕbara-na kariim ʕan-ha
 told.3S.MASC-us Karim about-them

‘We saw a/the play that you left when Karim told us about (it).’

(Aoun et al. 2010: 174)

Fourth, both Irish and Standard Arabic allow the resumptive pronoun to occur in the Wh-island configuration without obtaining any island effects, as in (55) and (56), respectively.

- (55) na dánta sin nach bhfuil fhios
the poems DEM COMP_{pro}+NEG is knowledge
again cén áit ar cumadh iad
at-us what place COMP were-composed them
‘those poems that we do not know where they were composed’
(McCloskey 2011: 76)

- (56) naʁrifu (r-)raʒula (allaði)tatasaaʔaluna ʃam-man
know.1P (the-)man (that) wonder.2P about-who
sayuqaabilu-hu
meet.3S.MASC-him
‘We know a/the man that you are wondering who will meet him.’
(Aoun et al. 2010: 174)

Finally, resumptive relative clauses in Irish do not exhibit weak crossover effects, as in (57). Here the resumptive pronoun *é* ‘him’ is coreferential with the epithet *bhastaird* ‘bastard,’ but no crossover effects are observed.

(57) Sin an fear ar dhúirt mé le tuismitheoirí
 that the man COMP_{pro} said I with parents
 an bhastaird gur cheart é a-chaitheamh isteach
 the bastard COMP+COP right him throw-INF into
 i bpríosún.
 in prison

‘That is the man that I said to the bastard’s parents that he should be thrown into prison.’

(McCloskey 2011: 80)

Thus far, the examples from languages that allow grammatical resumptive pronouns in relative clause formation demonstrate that the resumptive pronoun does not always obey island constraints. The insensitivity to islands make resumptive pronouns a different type of A’-dependencies from gaps. Besides, resumptive pronouns may not be sensitive to each kind of island configuration, which is a phenomenon that is worth investigating but will not be pursued in the current study.

In the next section, I discuss the configurations that have been reported in the literature to prevent resumptive pronouns from occurring.

3.1.3 Where Resumptive Relative Clauses Cannot Occur

Even though resumptive relatives have been shown to occur in different island conditions without the presence of island effects, there are some contexts where they are prohibited from being used in certain languages.

First, they exhibit strong crossover effects in Hebrew, as in (58), and in Irish, as in (59).

- (58) *Ze ha-baxur še- yidaʕti ʔet ha- ʔidiot
 this the-guy that-(I) informed ACC the-idiot
 še-ha-more yaxšil ʔoto
 that-the-teacher will.flunk him

‘This is the guy that I informed the idiot that the teacher will flunk.’

(Shlonsky 1992: 461)

- (59) *Sin an fear ar dhúirt mé leis an
 that the man COMP_{pro} said I with the
 bhastard gur cheart é a-chaitheamh isteach
 bastard COMP+COP right him throw-INF into
 i bpríosún.
 in prison

‘That is the man that I said to the bastard that he should be thrown into prison.’

(McCloskey 2011: 80)

In addition, Hebrew resumptive relative clauses exhibit weak crossover effects, as in (60).

(60) *ʔZe ha-baxur še- yidaʕti ʔet ha-horim šel
 this the-guy that-(I) informed ACC the-parents of
 ha-ʔidiot še-ha-more yaxšil ʔoto
 the-idiot that-the-teacher will.flunk him

‘This is the guy that I informed the idiot’s parents that the teacher will flunk.’

(Shlonsky 1992: 461)

Moreover, there is a commonly known constraint in resumptive relative clauses, which is sometimes called the Highest Subject Constraint/Restriction (e.g., McCloskey 2011; Borer 1984; Shlonsky 1992; Rouveret 2011). This prohibits resumptive pronouns from occurring in the subject position of the highest clause even when that language allows resumptive subject relativization in an embedded clause, as shown in Hebrew example (61).

(61) ha'iš še ___/*hu pagaš et rina
 the.man that ___/*he met Rina

Intended: ‘the man that met Rina’

(Sells 1984: 64)

In the next section, I review previous analyses of resumptive relative clauses.

3.2 *Previous Analyses of Resumptive Relative Clauses*

According to Chomsky (1977, pp. 80), whereas relative clauses associated with a gap involve wh-movement, relatives that use the resumptive pronoun involve no

movement at all because of their insensitivity to island constraints. He suggests that the latter contains a base-generated pronoun in the relative clause. Shlonsky (1992) examines the distribution of resumptive pronouns in relative clauses in Hebrew and Northern Palestinian Arabic (henceforth Palestinian), and argues that the use of a resumptive pronoun (aka resumptive strategy) can only serve as a last resort in the grammar, and occurs only when wh-movement is blocked. He suggests that resumptive pronouns are variables bound by A'-operators because of the strong crossover (Condition C) and weak crossover effects that they exhibit in Hebrew.

McCloskey examines the patterns of the gap and of the resumptive pronouns in Irish (e.g., McCloskey 1979, 2002, 2005, 2011, among others). According to his observations, gaps are associated with the properties that are listed in (62), and resumptive pronouns are connected with the properties listed in (63) (adapted from McCloskey 2002: 191):

(62) gap

- i. a gap (trace) in the variable position;
- ii. all the standard properties of movement to an A'-position – island effects, weak crossover, and so on (McCloskey 1979, 1985, 1990; Chung and McCloskey 1987);
- iii. morphosyntactic evidence for successive-cyclic movement

(63) resumptive pronoun

- i. a pronoun in the variable position;
- ii. evidence that movement plays no role in the derivation – immunity from islands, absence of weak crossover effects (McCloskey 1990);
- iii. morphosyntactic evidence (in the form of the default complementizer *go*) that intermediate C-positions play no role in establishing the binding relation.

Due to the different syntactic behaviours between gaps and resumptive pronouns in Irish, McCloskey (2002) proposes that the gap is derived by movement to an A'-position and that the resumptive pronoun is derived via base generation. Furthermore, he proposes that it is the different features associated with C that make the distinction in derivation. In the movement cases, the head C of the clause containing the gap is realized as *aL* and bears both an Operator Op-feature and an EPP-feature. This C will enter into an Agree relation with a *pro* that also bears the Op-feature. Since the C also has an EPP-feature, the *pro* will be moved to Spec-CP. In the resumption cases, the C is realized as *aN* and bears only the EPP-feature, so no Agree relation is established with *pro* and thus no movement ensues. Instead, a null operator is externally merged in Spec-CP to satisfy the EPP feature on C, and binds a variable pronoun base-generated in the relevant position in its c-command domain.

4 Chinese Relative Clauses

Chinese relative clauses exhibit the usage of both a gap and a resumptive pronoun, depending on which argument position is relativized. In Section 4.1, I introduce

gapped relative clauses and their distributions. In Section 4.2, I discuss resumptive relative clauses. In Section 4.3, I review Aoun and Li's (2003) analysis of Chinese relative clauses.

4.1 Chinese Gapped Relative Clauses

4.1.1 Chinese Gapped Relative Clauses in Different Argument Positions

Gapped relative clauses in Chinese occur in the subject and object positions of the highest clause, as in (64) and (65), respectively. Resumption is not an option in these cases.¹⁰

(64) zhe-ge [(*ta) xihuan Xiaoying] de nanhai
 this-CL he like Xiaoying REL boy
 'the boy who likes Xiaoying'

(65) zhe-ge [Xiaoying kan-guo (*ta)] de nanhai
 this-CL Xiaoying see-ASP he REL boy
 'the boy who Xiaoying saw before'

When the subject of an embedded clause is relativized, however, a gap can be used, but is interchangeable with a resumptive pronoun, as in (66):

¹⁰ However, there may be some regional differences in the acceptability/grammaticality of resumption in the direct object position of the highest clause, as some previous studies judge resumptive pronouns to be grammatical in objects of the highest clause (e.g., Gu 2001; Ning 2008).

- (66) a. na-ge Zhangsan juede (ta) hui du zhexie
 that-CL Zhangsan think he will read these
 shu de nanren
 book REL man
 ‘the man that Zhangsan thinks (he) would read these books’
- b. na-ge Zhangsan suo (ta) hui du zhexie
 that-CL Zhangsan say he will read these
 shu de nanren
 book REL man
 ‘the man that Zhangsan said (he) would read these books’
- c. na-ge Zhangsan xiangxin (ta) hui du
 that-CL Zhangsan believe he will read
 zhexie shu de nanren
 these book REL man
 ‘the man that Zhangsan (he) believed would read these books’

Similarly, when an object of embedded clauses is relativized, a gap is also optional, as in (67):

- (67) a. na-ge Zhangsan juede Lisi xihuan (ta) de
 that-CL Zhangsan think Lisi like she REL
 nüsheng
 girl
 ‘the girl that Zhangsan thinks Lisi likes (her)’
- b. na-ge Zhangsan suo Lisi hui bangmang
 that-CL Zhangsan say Lisi will help
 (ta) de xuesheng
 he REL student
 ‘the student that Zhangsan said Lisi will help (him)’

When a possessor is relativized, a gap is also optional, as in (68):

- (68) na-ge (ta) haizi shoushang de mama lai le.
 that-CL he child get.hurt REL mother come SFP
 ‘The mother whose child got hurt came.’

4.1.2 Chinese Gapped Relative Clauses in Island Conditions

In Chinese, gapped relative clauses also exhibit island effects in certain island conditions. First, both subject relativization and object relativization exhibit island effects in the Complex NP Condition, as in (69) and (70), respectively.¹¹

¹¹ Xu (2009) has a Chinese example that does not seem to violate the Complex NP constraint, as in (i).

(69) *zhe-ge [wo jian-guo [t_i xihuan t_j] de
 this-CL I see-ASP like REL
 laoshi_j]] de nanhai_i
 teacher REL boy

‘*the boy who I met the teacher who (he) likes’

(70) *zhe-ge [wo jian-guo [t_j xihuan t_i] de na-wei
 this-CL I see-ASP like REL that-CL
 laoshi_j]] de nanhai_i
 teacher REL boy

‘*the boy who I saw the teacher who likes’

(i) t_i chuan t_j de yifu_j hen zang de
 wear REL clothes very dirty REL
 na-ge ren_i
 that-CL person
 Intended: ‘the person such that the clothes that he wears are very dirty’
 *‘that person who the clothes (he) wears are very dirty’

(Xu 2009: p30, ex. 25b)

Compared to the examples in (69-70), one might think that the relativized position, that is, subject position, in the less embedded relative clause, i.e., *who (his) clothes ... are very dirty*, plays a role in making (i) grammatical in Chinese, since that seems to be the only difference between (i) and (69-70). In both (69) and (70), the relativized position in the less embedded relative clause is the object position, e.g., *the teacher*.

However, it becomes more complicated if an aspectual marker is added to the verb in the original example, as in (ii). The relativization of the original element now shows CNPC effects.

(ii) *[[t_i chuan-le t_j] de yifu hen zang
 wear-ASP REL clothes very dirty
 de na-ge ren
 REL that-CL person
 Intended: ‘the person such that the clothes that he wore are very dirty’

So far I do not have a good explanation for such a contrast, and perhaps it is because the subject trace t_i in (i) is in fact a covert pronominal *pro*. I will not pursue this issue in this chapter, though.

Gapped subject relativization and object relativization also appear to exhibit island effects in the Adjunct Island Condition, as in (71) and (72), respectively.¹²

(71) ??zhe-ge [[ruguo t_i xihuan Zhangsan] wo hui
 this-CL if like Zhangsan I will
 hen gaoxing] de nanhai_i
 very happy REL boy
 ‘the boy who I will be happy if he likes Zhangsan’

(72) ??zhe-ge [[ruguo Zhangsan xihuan t_i] wo hui
 this-CL if Zhangsan like I will
 hen gaoxing] de nanhai_i
 very happy REL boy
 ‘the boy who I will be happy if Zhangsan likes’

Gapped subject relativization and object relativization also show island effects in the Coordinate Structure Condition, as in (73) and (74), respectively.

¹² But see Chapter 6 (i.e., Section 2.1) for more relevant examples of gapped relatives in the Adjunct Island Condition. In that section, some gapped relatives do not appear to be sensitive to this constraint. For this puzzle, I don’t have a good answer yet.

(73) *zhe-ge [t_i xihuan Zhangsan] de nanhai_i han
 this-CL like Zhangsan REL boy and
 Lisi xihuan Wanwu
 Lisi like Wanwu
 ‘*the boy who likes Zhangsan and Lisi likes Wanwu’

(74) *zhe-ge [Zhangsan xihuan t_i] de nanhai_i han
 this-CL Zhangsan like REL boy and
 Lisi xihuan Wanwu
 Lisi like Wanwu
 ‘*the boy who Zhangsan likes and Lisi likes Wanwu’

In terms of the Wh-island condition, however, it appears that only gapped object relativization exhibits island effects, as in (75a).¹³ Gapped subject relativization is grammatical in a wh-island, as in (75b).

¹³ But notice that there seems to be a contrast between stage-level predicates and individual-level predicates in terms of island effects for object relativization. As shown in (i), when the predicate in the relative clause is replaced with an individual-level predicate such as *xihuan* ‘like’, there are no island effects in Wh-island anymore.

(i) zhe-ge [wo xiang zhidao [Zhangsan weisheme xihuan
 this-CL I want know Zhangsan why like
 t_i] de nanhai_i
 REL boy
 ‘the boy who I wonder why Zhangsan likes’

So far I have not found a satisfactory explanation for this contrast. It has not been discussed in the literature of relative clauses, but similar issues associated with wh-questions have been discussed in Wei et al. (2014). Nonetheless, I will provide more relevant data in Chapter 6 (i.e., Section 2.1) and show that this is probably just a superficial contrast in gapped relative clauses.

- (75) a. ??zhe-ge [Lisi xiang zhidao [Zhangsan weisheme
 this-CL Lisi want know Zhangsan why
 qin-le t_i] de nanhai_i
 kiss-ASP REL boy
 ‘the boy who Lisi wonders why Zhangsan kissed him’
- b. zhe-ge [Lisi xiang zhidao [weisheme t_i qin-le
 this-CL Lisi want know why kiss-ASP
 Zhangsan de nanhai_i
 Zhangsan REL boy
 ‘the boy who Lisi wonders why he kissed Zhangsan’

Chinese gapped relative clauses exhibit Strong Crossover effects, as in (76).

- (76) *na-ge ren_i jiushi na-ge hundan_i shuo [t_i
 that-CL man is that-CL bastard say
 hui sha-si women] de ren_i
 will kill-dead we REL person

*‘That’s the man_i that the bastard_i said that he_i would kill us.’

Chinese gapped relative clauses also seem to exhibit weak crossover effects,
 as in (77).

- (77) *zhe-ge [ta_i laopo bao-zhe t_i] de nanren_i
 this-CL he wife hug-ASP REL man
 ‘*the man_i that his_i wife is hugging’

4.2 Chinese Resumptive Relative Clauses

4.2.1 Resumptive Pronouns in Different Argument Positions

In Chinese, relativization of oblique object position must use a resumptive pronoun; it is ungrammatical with a gap, as in (78).

- (78) zhe-ge [Lisi zhengzai gen *(ta) shuohua] de
 this-CL Lisi PROG with him talk REL
 nanren
 man
 ‘the man that Lisi is talking with’ (Lit. ‘The man that Lisi is talking with him.’)

Resumption is also obligatory with an indirect object that is not marked by a preposition, as with the Goal argument of an object control verb, as in (79a-b):¹⁴

¹⁴ When the object of a causative verb is relativized, the resumptive pronoun is also obligatory; these may also be object control verbs, like English *force*.

- (i) a. na-ge Zhangsan shi *(ta) diedao de nanren
 that-CL Zhangsan cause he fall REL man
 ‘the man that Zhangsan caused to fall’
 b. na-ge Zhangsan rang *(ta) shengqi de nanren
 that-CL Zhangsan make he angry REL man
 ‘the man that Zhangsan made angry’

- (79) a. na-ge Zhangsan jiao *(ta) du zhexie
 that-CL Zhangsan ask he read these
 shu de nanren¹⁵
 book REL man
 ‘the man that Zhangsan asked to read these books’
- b. na-ge Zhangsan xiangyiao *(ta) qu
 that-CL Zhangsan want he go
 bangmang Lisi de xuesheng.
 help Lisi REL student
 ‘the student that Zhangsan wants to help Lisi’

¹⁵ Lu (2012) argues that *jiao* ‘ask/request’ takes an infinitival complement clause because *jiao* is incompatible with a complement clause that has a temporal phrase such as *zuotian* ‘yesterday’ in it, as in (i):

- (i) *Zhangsan jiao Lisi zuotian xian zou.
 Zhangsan ask Lisi yesterday first leave
 Intended: ‘Zhangsan asked Lisi to leave yesterday.’

(Lu 2012: 45)

In addition, whereas the verb *ask* in English is an ECM (Exceptional Case Marking) verb in that it can assign (accusative) case to the following argument, Li (1990) has argued that there is no ECM structure in Chinese. Even the possible candidates such as *yao* ‘want’, as in (ii-a) and *zhidao* ‘know’, as in (ii-b), are analyzed as object control verbs rather than ECM verbs.

- (ii) a. ta yao ta qu.
 he want he go
 ‘He wants him to go.’
- b. ta zhedao ta shangjin.
 he know he work.hard
 ‘He knows him to work hard.’ (Lit. ‘He knows that he works hard.’)
- (Li 1990: 132)

Here I adopt both Lu (2012) and Li (1990) and simply take *jiao* ‘ask,’ *xiangyiao* ‘want’ and similar verbs as object control verbs.

As noted in example (68) above, Chinese resumptive relatives can also optionally occur in the possessor position, as in (80):

- (80) zhe-ge Lisi kanjian-le (ta) taitai de nanren
this-CL Lisi see-ASP he wife REL man
‘the man whose wife Lisi saw’ (Lit. ‘The man who Lisi saw his wife.’)

In addition, even though resumptive relatives do not occur in the subject position of the highest clause, they can occur in the subject position of embedded clauses, as in (81a-b).

- (81) a. na-ge Zhangsan juede (ta) hui du zhexie
 that-CL Zhangsan think he will read these
 shu de nanren
 book REL man
 ‘the man that Zhangsan thinks (he) would read these books’
- b. na-ge Zhangsan shuo (ta) hui du zhexie
 that-CL Zhangsan say he will read these
 shu de nanren
 book REL man
 ‘the man that Zhangsan said (he) would read these books’

The resumptive pronoun can also occur in the relativized object position of an embedded clause, as in (82a-b):

- (82) a. na-ge Zhangsan juede Lisi xihuan (ta) de
 that-CL Zhangsan think Lisi like she REL
 nūsheng
 girl
 ‘the girl that Zhangsan thinks Lisi likes (her)’
- b. na-ge Zhangsan shuo Lisi hui bangmang
 that-CL Zhangsan say Lisi will help
 (ta) de xuesheng
 he REL student
 ‘the student that Zhangsan said Lisi will help (him)’

Below is a table that summarizes the usage of resumptive relatives in Chinese.

Table 2: Resumptive Pronouns in RCs in Chinese

Language	Argument Positions					
	Matrix Clause				Embedded Clause	
	Subject	Object	Oblique Object	Possessor	Subject	Object
Chinese	–	–	+	(+)	(+)	(+)

Key: The symbol ‘–’ indicates that a resumptive pronoun is prohibited from occurring in that position; ‘(+)’ indicates that the occurrence of a resumptive pronoun is optional; ‘+’ indicates that the occurrence of a resumptive pronoun is obligatory.

4.2.2 Conditions Where Chinese Resumptive Relative Clauses Can Occur

There are some island conditions where resumptive relatives can occur without island effects. These include the Adjunct Island Condition, as in (83), and Wh-island Condition, as in (84).

(83) a. na-ge [[ruguo Zhangsan gen ta shuohua]
 that-CL if Zhangsan with he talk
 Lisi hui hen shengqi]de nanren
 Lisi will very angry REL man
 ‘the man who if Zhangsan talks with him Lisi will be very angry’

b. na-ge [[ruguo ta taitai bu lai] Lisi
 that-CL if he wife not come Lisi
 hui hen shengqi]de nanren
 will very angry REL man
 ‘the man who if his wife does not come Lisi will be very angry’

(84) na-ge [wo xiang zhidao [Zhangsan weisheme dui
 that-CL I want know Zhangsan why to
 ta hui-shou]] de nanhaii
 he wave-hand REL boy
 ‘the boy that I want to know why Zhangsan waved at him’

Moreover, under the Weak Crossover Condition, which governs the potential interpretation of pronouns, there are no weak crossover effects in resumptive relatives, as in (85).

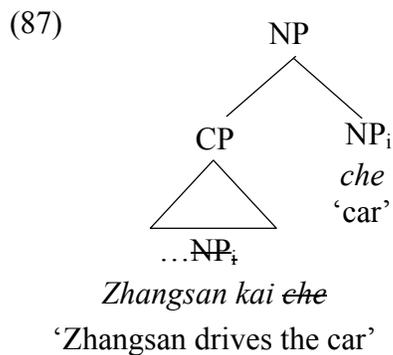
- (85) zhe-ge ta_j laopuo bao-zhe ta_j de nanren_j
 this-CL he wife hug-ASP he REL man
 ‘the man_j that his_j wife hugs him_j’

As shown above, the pronoun *ta* ‘he’ in the object position of the relative clause can refer to the head noun *nanren* ‘man’. In the next section, I review Aoun and Li’s (2003) analysis of Chinese relative clauses.

4.3 *Aoun & Li (2003)*

In terms of relativization on argument positions in Chinese, Aoun and Li (2003) propose two different mechanisms to generate relative clauses using a gap and relative clauses using a resumptive pronoun separately. When relative clauses involve a gap, as in (86), they adopt Kayne’s (1994) promotion, or, raising analysis, where the relativized head originates from the relative clause and then moves to the head of the NP. The corresponding syntactic structure is illustrated in (87).

- (86) [CP Zhangsan kai t_i]de che_i
 Zhangsan drive REL car
 ‘the car the Zhangsan drives’



Notice that Aoun and Li treat the relative clause, that is, the CP, as an adjunct of the entire complex NP. This differs from Kayne’s (1994) raising analysis, which takes CP as a complement of D in the relativized NP, as in (22). According to Aoun and Li, the NP, i.e., *che* ‘car’, can be generated in the relative clause in Chinese, and then raise to the head position of the entire complex NP. In other words, it is an NP raising in Chinese, unlike the DP raising in English. One of the reasons they use to argue for NP raising instead of DP raising is due to the lack of reconstruction effects when a quantifier phrase is the head of relative clauses in Chinese¹⁶. For example, in (88), the noun phrase *mei-ge-ren* ‘everyone’ in the relative clause must take wide scope, and the quantifier phrase *san-ben shu* ‘three books’ always takes narrow scope. No reconstruction effects appear.

¹⁶ In their analysis of noun phrases in Chinese, D and Q(uantifier) occur in the same position.

(88) wo hui zhengli mei-ge-ren hui kan de
 I will arrange every-CL-person will read REL
 san-ben shu
 three-CL book

‘I will put the three books that everyone will read in order.’

(everyone > 3 books; *3 books > everyone)

(Aoun & Li 2003: 134)

In contrast, when the head of a relative clause does not involve a quantifier or determiner phrase, reconstruction effects seem to appear. As shown in (89), the quantifier phrase *mei-ge-ren* ‘everyone’ can bind the reflexive *ziji* ‘self’ even though on the surface structure *ziji* is not within the same clause as its antecedent.

(89) [wo jiao Zhangsan guan mei-ge-ren_i kai t_j
 I ask Zhangsan persuade every-CL-person drive
 lai] de [ziji_i de chezi]_j.
 come REL self POSS car

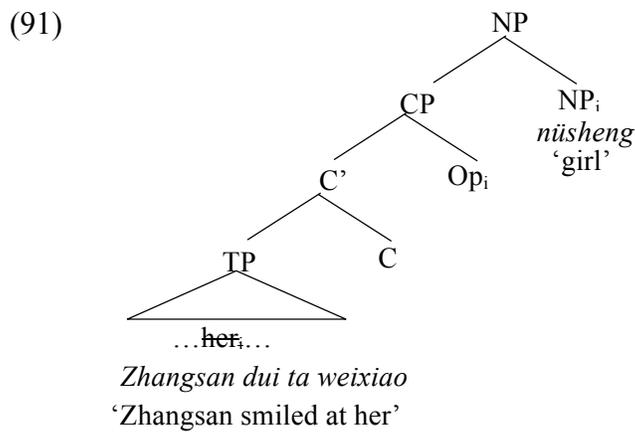
‘self’s car that I asked Zhangsan to persuade everyone to drive over.’

(Aoun & Li 2003: 132)

As for relative clauses that use a resumptive pronoun, as in (90), a base-generation analysis is proposed. Under this analysis, an empty wh-operator (Op) is

required, and the resumptive pronoun can be a different lexical item in the numeration. I illustrate this derivation in (91).

- (90) [CP Zhangsan dui ta_i weixiao]de nüsheng_i
 Zhangsan to her smile REL girl
 ‘the girl that Zhangsan smiled at’



One piece of evidence that Aoun and Li provide to distinguish these two types of relative clauses comes from the sensitivity to island effects. Relatives formed with the resumptive pronoun do not exhibit island effects, as discussed above such as (83-84), and are illustrated here with similar examples. No Adjunct Island effects occur in (92a) and no Complex NP effects are observed in (92b). In contrast, when the resumptive pronoun in these relative clauses is substituted with a gap, the sentences become ungrammatical, thus exhibiting island effects, as in (93a-b):

(92) a. wo xiang kan [na-ge [ni [yinwei ta_i
 I want see that-CL you because he
 bu hui lai] hen shengqi]de xuesheng;_i
 not will come very angry REL student
 ‘I want to see the student that you are angry because he would not
 come.’

b. wo xiang kan [na-ge [ni yaoqing [dai ta_i
 I want see that-CL you invite bring he
 lai]de ren lai zher]de xuesheng;_i
 come REL person come here REL student
 ‘I want to see the student that you invited the person over that brought
 him over.’

(Aoun & Li 2003: 170)

(93) a. *wo xiang kan [na-ge [ni [yinwei t_i
 I want see that-CL you because
 bu hui lai] hen shengqi]de xuesheng;_i
 not will come very angry REL student
 ‘I want to see the student that you are angry because he would not
 come.’

- b. *wo xiang kan [na-ge [ni yaoqing [dai t_i
 I want see that-Cl you invite bring
 lai]de ren lai zher]de xuesheng;
 come REL person come here REL student
 ‘I want to see the student that you invited the person over that brought
 him over.’

Thus, according to Aoun and Li, gapped relatives are derived via movement, while resumptive relatives are derived via base-generation.

5 Conclusion

In this chapter I illustrated gapped and resumptive relatives with data from various languages including English, Hebrew, Irish, Arabic, Welsh and Chinese. So far, the examples discussed indicate three things. First, there is no apparent ban on the occurrence of gaps in different argument positions, although a gap may be prohibited from appearing in island-type configurations due to its association with the movement structure. Second, resumptive pronouns, including obligatory and optional ones, can potentially be present in all argument positions except the subject position of the highest clause. Third, the distribution of resumptive pronouns appears to be the inverse of the distribution of gaps in the AH. Resumptive pronouns tend to occur in argument positions that are lower on the AH, while gaps are the opposite. In the next chapter, I present an experimental study on Chinese-speaking children’s comprehension of gapped and resumptive relatives.

CHAPTER 3: CHINESE-SPEAKING CHILDREN'S COMPREHENSION OF GAPPED AND RESUMPTIVE RELATIVE CLAUSES

1 Introduction

In this chapter, I present data from a truth value judgment task performed with Chinese-speaking children, which demonstrate that a purely UG-based structural proposal, namely the Structural Distance Hypothesis, is not entirely supported.

Recall that for the UG-derived approach to the acquisition of different types of relative clauses, I examine the psychological ease of acquisition implied by the hierarchical relationship between relatives based on Keenan and Comrie's (1977) Accessibility Hierarchy (AH). According to Keenan and Comrie, the higher the grammatical role in the AH, the easier it is to understand the relative clause. Although the AH is an implicational universal derived from data in about fifty languages, this order matches the embeddedness of different grammatical roles in terms of the hierarchical structure proposed in UG-based syntactic theories. For example, the role in the highest position in the AH, namely, subject, is also the least embedded compared to the other roles such as direct and indirect objects or oblique objects in the syntactic structure. Hence, the Structural Distance Hypothesis states that more embedded positions will show worse performance/slower acquisition than less embedded positions.

In the present experimental study, I tested Chinese-speaking children's comprehension of the top three types of relative clauses in the AH: Subject Gapped

relatives, as in (1), Object Gapped relatives, as in (2), and Oblique Object Resumptive relatives, as in (3).

(1) Subject Gapped relatives

[___ qin-le maomi de] xiao-gou pao-zou le.
 kiss-ASP cat REL little-dog run-away ASP

‘The dog that kissed the cat ran away.’

(2) Object Gapped relatives

[xiao-gou qin-le ___ de] maomi pao-zou le
 little-dog kiss-ASP REL cat run-away ASP

‘The cat that the dog kissed ran away.’

(3) Oblique Object Resumptive relatives

[qingwa dui ta tiaowu de] yazi pao-zou le.
 frog to it dance REL duck run-away ASP

‘The duck that the frog danced for ran away.’

Note that, in addition to the different embeddedness among relativized arguments, the third type of relative clause in Chinese, Oblique Object Resumptive RC, requires a resumptive pronoun in the relativized position where the other two types use gaps. As mentioned in Chapter 2 (and also see Keenan & Comrie 1977), while gaps tend to appear in argument positions that are higher on the AH,

resumptive pronouns appear in positions that are lower on the AH. If this indicates that gaps have a preference for less embedded positions over more embedded ones and that resumptive pronouns prefer more embedded positions over less embedded ones, then this tendency may reduce the predictive power of the Structural Distance Hypothesis when both the gap and resumption strategies are at work. Because of this additional contrast, it allows us to examine, on the one hand, whether the hierarchical structure is the only factor that affects Chinese-speaking children's acquisition of relative clauses, and on the other hand, how well the resumption strategy fares against the gap strategy in acquisition.

The remaining sections of this chapter are organized as follows. In Section 2, I review past studies on the acquisition of RCs. In Section 3, I describe the experimental study in detail. In Section 4, I show the results and analyses. Section 5 is the general discussion, and Section 6 is the conclusion.

2 Past Research on the Acquisition of RCs

Many previous studies on the acquisition of relative clauses have found that children perform better on subject than object relatives, i.e., the subject-object asymmetry (e.g., Brazilian Portuguese: Corrêa 1995; European Portuguese: Costa et al. 2014; Chinese: Lee 1992; Hsu et al. 2009; Hu, Gavarro, & Guasti 2016; Hu, Gavarro Vernice, & Guasti 2016; Cantonese: Lau 2016; English: McKee et al. 1998; German: Diessel & Tomasello 2005; Hebrew: Arnon 2005; Friedman et al. 2009; Palestinian Arabic: Botwinik et al. 2015; among many others). Such a consistent finding appears to support the Structural Distance Hypothesis, where the subject is considered less

embedded than object in the hierarchical structure, thus making subject relatives easier to learn than object relatives.

However, in languages where both the gap and the resumption strategies are grammatical options in relative clause formation, results are mixed concerning which strategy would outperform the other. For example, Lau (2016) tested Cantonese-speaking children's comprehension and production of subject gapped RCs, object resumptive RCs and object gapped RCs. The results from the comprehension task showed that children performed best on subject RCs, and equally bad on both types of object RCs. In the production task, children never used the resumption strategy to form any of these three types of relative clauses. In Hebrew, Arnon (2005) found that children performed slightly better on object resumptive RCs than object gapped RCs, whereas Friedman et al. (2009) did not find any significant difference in children's comprehension of these two types of object RCs. More compellingly, Rahmany et al. (2014) tested Persian-speaking children's comprehension of subject gapped RCs, object gapped RCs and object RCs containing a resumptive pronoun or an object clitic. Their results showed that children performed better on object RCs containing the resumptive elements than subject gapped and object gapped RCs.

Due to the potential of the resumption strategy to outperform the gap strategy, it is worth investigating how Chinese-speaking children perform on gapped and resumptive RCs. But to the best of my knowledge, few studies have been dedicated to the influence of a resumptive pronoun on the acquisition of Chinese relative clauses. In the following sections, I review two relevant experimental studies on the

acquisition of gapped and resumptive RCs in Chinese: Lee's (1992) comprehension study and Su's (2004) production study.

2.1 Lee (1992)

Lee (1992) investigated three main questions. The first is whether the functional position that is relativized in the RC makes a difference in children's performance (i.e., where the RC's gap is). The second is whether the attachment site of the complex NP has any effect on performance. The third is whether there is a difference between gapped RCs and resumptive RCs.

He conducted two experiments using an act-out task. The procedures of an act-out task are as follows. The audio recording of the test sentences were made earlier by a native speaker of Beijing Mandarin Chinese. Two experimenters interviewed the subject. While one experimenter controlled the tape recorder and arranged the toys, the other experimenter recorded the subject's responses. For each test sentence, the subject was presented with relevant toys and was instructed to arrange them according to the sentences on the tape. One more thing that is worth mentioning is that, originally Lee tried to satisfy the felicity conditions that Hamburger and Crain (1982) discussed by including at least two toys of the type that corresponds to the head noun of a RC, but his pilot study showed that 4-year-olds did not work well with too many toys. Thus, he took out the toys that fulfil the felicity conditions in the experiments. In view of these difficulties in Lee's act-out task with young children, the current study uses another task that avoids these problems, and I will come back to the discussion of this task used in my study in Section 3.2.

The subjects in Lee's first experiment included five age groups of children (mean ages: 4;7, 5;6, 6;6, 7;5 and 8;5), and each group had 12 subjects, except for the 7-year-old group that had 13 subjects. There was no adult control group.

This experiment tested six types of RCs. Examples in (4-9) below illustrate these types. They can be divided into two major groups. In one group, the attachment site of the complex NP functions as the subject of the main sentence, as in (4-6). In this group, the functional position that is relativized in the RC includes a subject, as in (4), an object, as in (5), and an indirect object, as in (6). Also in (6), a resumptive pronoun such as *ta* 'him/her' is required since it is an indirect object in the RC.

(4) [__ bao zhe xiaoxiong de] neige baitu ti houzi
 hug ASP teddy-bear REL that rabbit kick monkey
 'The rabbit that is hugging the teddy-bear kicks the monkey.'

(5) [xiaoxiong bao zhe __ de] neige baitu ti houzi
 teddy-bear hug ASP REL that rabbit kick monkey
 'The rabbit that the teddy-bear is hugging kicks the monkey.'

(6) xiaogou ti ta shuaya de] neige baitu gen
 puppy for him brush-tooth REL that rabbit follow
 zhe xiaoxiong
 ASP teddy-bear
 'The rabbit for whom the puppy is brushing his teeth follows the teddy-bear.'

In the other group, the attachment site of the complex NP functions as the (direct or indirect) object of the main clause, as in (7-9). The functional position that is relativized in the RC includes a subject, as in (7), an object, as in (8), and an indirect object, as in (9).

- (7) xiaogou cai [___ bao zhe houzi de] neige
 puppy step-on hug ASP monkey REL that
 baitu
 rabbit

‘The puppy steps on the rabbit that is hugging the monkey.’

- (8) xiaogou cai [houzi bao zhe ___ de]
 puppy step-on monkey hug ASP REL
 neige baitu
 that rabbit

‘The puppy steps on the rabbit that the monkey is hugging.’

- (9) xiaoxiong gen zhe [xiaogou ti ta shuaya
 teddy-bear follow ASP puppy for him brush-tooth
 de] neige baitu
 REL that rabbit

‘The teddy-bear follows the rabbit for whom the puppy is brushing his teeth.’

Lee collected the number of correct responses on each type of RC from five age groups. A response was considered correct only when both the main clause and the RC were acted out correctly; otherwise, the score was zero.

Lee's results from the first experiment are summarized as follows. First, when all ages were combined, children performed better on subject gapped RCs than object gapped and indirect object resumptive RCs. Second, there was no significant difference between object gapped RCs and indirect object resumptive RCs in the performance of any age group. Third, all ages grouped together performed better on RCs associated with the subject attachment site than RCs with the object attachment site.

The first result aligns with what cross-linguistic studies have found with respect to the subject-object asymmetry; that is, people perform better on subject gapped RCs than object gapped RCs. The second result that object gapped RCs and indirect object resumptive RCs did not differ seems to suggest that children were not affected by the extra structural distance between the head and the relativized position in indirect object resumptive RCs. While the former supports the Structural Distance Hypothesis, the latter is at odds with it.

Furthermore, because Lee was concerned that the length and the complexity of sentences in the first experiment might be the cause of children's general poor performance, he conducted a second experiment using sentences that replace the transitive verb in the matrix clause with an intransitive verb. This reduces the RC types into three: subject gapped RCs, as in (10), object gapped RCs, as in (11), and indirect object resumptive RCs, as in (12).

(10) [__ bao zhe xiaoxiong de] neige baitu shui le.

hug ASP teddy-bear REL that rabbit sleep ASP

‘The rabbit that is hugging the teddy-bear has fallen asleep.’

(11) [xiaoxiong bao zhe __ de] neige baitu shui le

teddy-bear hug ASP REL that rabbit sleep ASP

‘The rabbit that the teddy-bear is hugging has fallen asleep.’

(12) [houzi gei ta xilian de] neige xiaogou

moneky for him wash-face REL that dog

shui le.

sleep ASP

‘The dog for whom the monkey washing face has fallen asleep.’

In this second experiment, he tested 8 4-year-olds and 8 5-year-olds, and found no significant difference among these three types of RCs in both groups.

Finally, there is one big drawback in Lee’s study. None of his experiments included an adult control group. This may be a problem when one wants to compare the grammar among different age groups: at least a group that has a mature grammar needs to serve as the reference level. Even though he had 8-year-olds in the first experiment, and their best performance in a type of RC was 93.8%, there were still two types of RCs where their performances were below 50%, which were object gapped RCs and resumptive RCs with the attachment site in the object position. In

order to see whether there is a developmental change in children's knowledge about RCs, an adult control group should be included.

2.2 *Su (2004)*

Su (2004) investigated two questions that overlap with those in Lee's study. First, she asked whether the gap position in the RC makes a difference for children. Second, she asked whether children perform differently on gapped RCs and resumptive RCs.

She used an elicited production task based on the adjustments made by Hamburger and Crain (1982). In particular, she included two identical toys of the type that corresponds to the head noun of a RC, which differed only in the actions they are doing (e.g., one ate cookies and the other ate chicken). In each experiment, there were two experimenters. During each trial, Experimenter 1 acted out a story with toys and props in front of the participant and the puppet (Big Bird) played by Experimenter 2. After the story, the puppet turned around, and then Experimenter 1 pointed to one of the toys that the subject can see but the puppet cannot. Experimenter 1 then told the subject to guide the puppet to pick up the toy that s/he just pointed to. Su was also cautious about avoiding using any relative clauses or any pronouns that may correspond to a resumptive pronoun in an RC in the storytelling part, since these are the crucial linguistic information that the experiment was designed to investigate.

The subjects included a younger group of 20 children (with mean age 5;3), an older group of 20 children (with mean age 6;1) and 31 adults in the control group.

Her materials included gapped RCs and resumptive RCs, as shown in (13-14) and (15-17), respectively. In addition, (13) is a subject gapped RC, and (14) is a gapped object RC, where the target utterance is an NP with a RC; the RC's

attachment site is not probed because it is not embedded in a full sentence. The resumptive RCs include the one where the relativized position is an object of a preposition, as in (15), and a subject of an embedded clause, as in (16), or a subject of a wh-island clause, as in (17) (Su 2004, ex. (4)):

(13) [___ chi binggan de] nühai
 eat cookie REL girl
 ‘the girl that ate cookies’

(14) [gongzhu qi ___ de] ma
 princess ride REL horse
 ‘the horse that the princess rode’

(15) [baixue-gongzhu dui ta dazhaohu de] Kitty mao
 snow-princess to her greet REL Kitty cat
 ‘the Kitty cat to whom Snow White greeted’

(16) [wupo jiao [ta qu xizao] de] laoshu
 witch ask him go bathe REL mouse
 ‘the mouse that the witch asked him to go bathing’

- (17) [tianshi zhidao [ta xihuan chi sheme]de] nanhai
angel know he like eat what REL boy
‘the boy that the angel knew what he liked to eat’

She collected all RCs that each of the three age groups produced. She presented the percentage of target RCs for each group, and the percentage of RCs that co-occurred with resumptive pronouns, which included grammatical and ungrammatical ones. For example, subject RCs, as in (14) and object RCs, as in (15) are considered to be ungrammatical with a resumptive pronoun in the relativized position, but the children in her study produced resumptive pronouns in some of these RCs.

Her results are summarized here. First, there was no significant difference between subject gapped RCs and object gapped RCs. Second, the younger age group of children generally produced a lower percentage of RCs using a resumptive pronoun than the older age group of children, and this difference between child groups was significant in the resumptive RC that involves a wh-island, as in (17). She then suggests that there is a developmental change toward the usage of complex structures. Third, based on the percentage of produced resumptive pronouns in different types of RCs, she indicates that all age groups, including children and adults, produced more resumptive pronouns in more complicated RCs, as in (15-17) than in simpler RCs, as in (13-14).

Concerning her data and results, what is most relevant here is that her study seems to suggest that there is a correlation between a complex structure and the usage of a resumptive pronoun, since younger children produced a lower percentage of

resumptive RCs that involve a wh-island than older children, and all age groups produced more resumptive pronouns in more complicated RCs, which are the ones where a resumptive pronoun is required in Chinese. However, even though it might be true that there is a correlation between a complex structure and the usage of a resumptive pronoun in the RC, this can only be validated when the experimental materials involving a complex structure, e.g., longer, are not always linked with a resumptive RC and that materials involving a simpler structure, e.g., shorter, are not with a gapped RC. Thus, there is a confounding factor involved in Su's study. That is, her significant results that suggest there is a correlation between complex structures and the usage of a resumptive pronoun only come from experimental materials that involve *both* a complex structure such as a wh-island and a resumptive RC, as in (17). Therefore, one cannot tell, from her study, whether the resumptive strategy is generally more difficult or easier than the gap strategy in RCs.

2.3 *Interim Summary*

Lee (1992), but not Su (2004), found that Chinese learners' performance was better on subject gapped RCs than object gapped RCs. Thus, while Lee's study supports the Structural Distance Hypothesis, Su's study does not. Despite the apparent conflict, there is one possibility that may have led to the absence of subject-object asymmetry in Su's results. That is, her examples of subject gapped RCs and object gapped RCs were all free-standing NPs and not included in a sentence. It could be that these RCs are too simple so that the effects did not show up in her production task. If this concern is valid, then this suggests that either a different subtle on-line task needs to

be utilized, or an even younger age group of children should be tested, especially those that are still learning words and phrases.

Moreover, both Lee (1992) and Su (2004) examined the effect of different RC formation strategies, and only Su found a difference between gapped RCs and resumptive RCs. But as mentioned before, Su had a confounding factor in those RCs using a resumptive pronoun. That is, her examples of RCs using a resumptive pronoun are longer than gapped RCs and some involve more than one layer of embedding. If the length or complexity of sentences is indeed the primary reason for poorer performance of her participants on resumptive RCs than gapped RCs, then it is problematic to conclude that the resumptive strategy is more difficult than the gap strategy in RCs. As for why Lee did not find an effect, there are at least two possible explanations. One is that all these participants are old enough to master both strategies equally well when they do not consider subject RCs. The second possibility is that children find both object gapped RCs and indirect object resumptive RCs equally easy or hard to act out. As the object gapped RC in (5) and the indirect object resumptive RC in (6) show, the head noun that is associated with the object/indirect object, whether it is a gap or a resumptive pronoun, always come immediately after the RC. As a result, it is possible that, when children were acting out the relevant events, they found the same clue for the direct object and the indirect object. They simply chose the animal that followed the relativizer *de*, so that their performance on the two types of RCs did not differ.

Finally, I would like to draw the reader's attention to the participants' ages in these studies. The youngest group in Lee (1992) was comprised of 4-year-olds, and

those in the youngest group in Su (2004) were 5-year-olds. None of these studies on the acquisition of Chinese RCs has tested children younger than 4, but Lee's study showed that 4-year-olds already have some, if not all, knowledge about RCs. Thus, it is important to have a study which examines knowledge about RCs in children younger than 4 years old. That is why the present study has included subjects younger than 4 years old, (more detailed information about the subjects in this study is discussed in Section 3.1).

3 Experimental Research

In this study, I used a truth value judgment task to examine the acquisition of RCs in Chinese-speaking children between 2;9 and 6;10. In particular, I compared three types of RCs: Subject Gapped RCs, Object Gapped RCs, and Oblique Object Resumptive RCs, as in (1-3), respectively. To avoid effects from the length or multiple layers of embedding, I simply focus on RCs that modify the subject position of sentences whose matrix verbs are intransitive.

A pilot study was conducted in English before this experiment was done. (See Appendix A for a brief summary of the pilot study.)

Next, I provide detailed information on the participants, and then describe the truth value judgment task and the experimental materials.

3.1 Subjects

Child group. I tested 73 children (38 boys and 35 girls) from three different cities in Taiwan (16 from Taipei, 25 from Hsinchu and 32 from Tainan). They ranged in age from 2;9 to 6;10. They were native speakers of Chinese as spoken in Taiwan (with

exposure to some other second languages such as Taiwan Southern Min, Hakka, Vietnamese, English, or some Cantonese). The experiment was conducted in Taipei in May, in Hsinchu in June and in Tainan in July and August 2013. The analyzed data comes from 68 children. The reduction of participants is due to failure to pass the pre-test (3 out of 4 sentences; n=3), and response bias (over 70% yes in fillers where the correct percentage of yes is 50%; n=2). The mean age of the analyzed subjects is 5;2. Table 3 summarizes these numbers of participants, with the gray rows referring to the 68 subjects whose data I analyze here.

Table 3: Number of Participants by Category

<i>Participants</i>	<i>Number</i>
Failure to pass the pre-test	3
Yes bias in fillers (over 70%)	2
Data analyzed	68
Female/Male	33/35
From each city	Taipei (16); Hsinchu (24); Tainan (28)

Adult group. 22 native speakers of Chinese spoken in Taiwan participated in this study. They were above 20 years old, and were students from universities in Taiwan and from the University of Arizona.

3.2 *Tasks and Materials*

3.2.1 The Truth Value Judgment Task

This study uses a Truth Value Judgment (TVJ) task. It is a comprehension task that probes children's linguistic competence (see e.g., Gordon 1996; Crain & Thornton 1998). There are two experimenters involved in this task: Experimenter 1 tells a story using toy props, and Experimenter 2 utters grammatical sentences that are true or

false in the context of the other experimenter's stories and takes notes of the participant's responses.

Before the experiment starts, Experimenter 1 asks the participant if he or she wants to play a game that involves stories, toys and a puppet. In this game, Experimenter 1 introduces a puppet (played by Experimenter 2) to the participants and tells them that the puppet will play the game with them. Specifically, Experimenter 1 indicates that the puppet will listen to stories together with them, and will say what happens at the end of each story. Experimenter 1 then emphasizes that the puppet is goofy so that it would sometimes say things right, but sometimes say things wrong, so that the puppet needs the participant's help. Furthermore, Experimenter 1 introduces a cookie and a rag, and indicates that when the puppet says the right thing, the participant can give it a cookie as a reward, but when the puppet says the wrong thing, the participant can give it a rag as a punishment. The cookie/rag detail motivates children to give a positive response or a negative response. This kind of reward/punishment task, rather than just a yes/no task, is crucial in this study because it makes both responses as enjoyable as possible for children, especially those that are younger than 4 years old. Potentially there might be a problem about distracting children using this cookie/rag, but in the real experiments, this did not cause any problem. More importantly, it was observed that this cookie/rag procedure helped the child to get engaged in the game at the beginning. Once the child got engaged and learned to play the game well, some of them just started saying yes/no to the puppet without remembering to give the puppet the cookie or the rag. It is important to understand that the primary purpose of introducing the cookie/rag was

to help the child learn to play the game. If it becomes a distractor, then the experimenters can judge whether to remove it or not during the experiment.

After these instructions, Experimenter 1 tells the participant a story using toys and props. In the meantime, Experimenter 2 plays the puppet, who watches the story with the participant. At the end of each story, the puppet (that is, Experimenter 2) makes a series of statements about the story, and the participant judges the truth of the puppet's statements based on what happens in the story. When the puppet describes the event correctly, the participant gives the puppet a cookie. When the puppet describes the event incorrectly, the participant gives the puppet a rag.

One reason for using this task is that it engages children in fun activities. To make it fun, a puppet is chosen to utter the trial sentences. This is to make children feel comfortable when they are about to judge whether each sentence matches the story or not. Children usually do not feel comfortable correcting an adult, but they may feel more comfortable correcting a puppet, especially when the puppet is said to be goofy sometimes and thus would make mistakes.

In addition to being easy and fun for children, the TVJ task has some merits for researchers as well. In this task, both true and false versions of the same test sentence are presented to the participant. Consider (18). This test sentence is presented in two different stories. In one story, what happens matches this sentence, so that this sentence is true. In the counterpart story, what happens does not match this sentence, so that this same sentence is false.

- (18) [qingwa dui ta tiaowu de] yazi pao-zou le.
frog to it dance REL duck run-away ASP
‘The duck that the frog danced for ran away.’

A picture in Figure 1 illustrates the layout of one of the paired stories, where a target sentence like (18) will be uttered at the end of the story.



Figure 1: Sample Experimental Layout

The availability of both the true and false versions of the test sentence allows the researcher to deal with the potential problem of random responses. On the one hand, when the participants do not pay attention to the story, they are likely to give random responses. As a result, they may respond correctly by chance only on half of the true and false versions of the same sentences. On the other hand, if the participants do pay attention to the story, then their responses of the true and false versions of test sentences should consistently reflect their knowledge about the sentences. This offers a solid way to test whether the participant understands the test sentences or not.

3.2.2 Materials

In the following, I first discuss the experimental factors that this experiment manipulated, and then move to other considerations in designing these materials.

3.2.2.1 Experimental Factors

This study tests Chinese-speaking children's comprehension of three types of RCs: Subject Gapped RCs, Object Gapped RCs, and Oblique Object Resumptive RCs. The attachment site of the complex NP that contains the RC is restricted to the subject position of the main clause in all target sentences.

There are four target sentences for the Subject Gapped RC, two target sentences for the Object Gapped RC, and two target sentences for the Oblique Object Resumptive RC. Each target sentence occurs once in two different stories. In one of the two stories, the target sentence is true, and in the counterpart story, the same sentence is false. The total number of target sentences is 16 (i.e., $(4+2+2)*2=16$).

3.2.2.2 Other Considerations

In this experiment, a pre-test precedes 8 test stories. The pre-test serves as the criterion for evaluating participants' understanding of the experiment. There are four sentences in this pre-test. As for each test story, there are five sentences. The total number of sentences presented for the child to judge in this experiment is thus 44, including the pre-test items.¹⁷ (Also see Appendix B for a complete list of the sentences.)

¹⁷ That is, 4 (in the pre-test) + 8 (stories) * 5 (in each story) = 44 .

To illustrate what the layout looks like in each story, consider the examples in (19a-e). In the layout of all sentences in each story, there are two target sentences, as in (19c) and (19e), and three filler sentences, as in (19a-b) and (19d). The organization of the layout is always the same, so the order of both target sentences and filler sentences never changes in each story. Despite this, the order of correct yes/no answers varies randomly in each story, so that there is no fixed order of correct yes/no answers. (See Appendix C for the correct yes/no answers in each story.)

(19) Examples in Story 1

a. you yi-zhi xiaogou zai wazi-shang tiaowu.
 have one-CL dog in sock-on dance

‘A dog was dancing on the sock.’

b. you yi-zhi maomi zai wazi-shang tiaowu.
 have one-CL cat in sock-on dance

‘A cat was dancing on the sock.’

c. qin-le maomi de xiaogou pao-zou le.
 kiss-ASP cat REL dog run-away ASP

‘The dog that kissed the cat ran away.’

d. zhe-zhi maomi zai chi yu.
 this-CL cat PROG eat fish

‘This cat was eating a fish.’

e. xiaogou qin-le de maomi pao-zou le.
 dog kiss-ASP REL cat run-away ASP

‘The cat that the dog kissed ran away.’

The fixed order of the sentences in all stories might be problematic. But there are a few good reasons to use the same story paradigm throughout the experiment. I will explain each in detail.

To begin with, a pair of filler sentences that have the same structure are presented first. Take the first two fillers in Story 1 for example. As illustrated in (19a-b), both are simple sentences, and they differ only in the animal referred to. The purpose of this order is to see if the participants are paying attention to the story. In Story 1, it is the dog, but not the cat, that was dancing on the sock. If the participant is paying attention to the story, there should be no contradictory responses on this pair.

Moreover, as shown above, each story has two target sentences that both contain RCs, as in (19c) and (19e). Therefore, to avoid confusion since both sentences contain the same number of toys and actions, it would be wise to have another filler in-between.

Turning to the role of the context, I will focus on the importance of plausible deniability. Crain et al. (1996) argue that children’s poor performance on sentences with quantification is not always due to their lack of linguistic knowledge. Instead, it is because the experimental designs do not conform to the condition called plausible dissent (or deniability). They use Philip’s (1995) examples to illustrate this point. Consider a picture in which there are three farmers and each is feeding a donkey and

one additional farmer who is not feeding a donkey. In order to reject (20a), the child should at some point during the trial be presented with another event where (20b) is true. The rationale behind this condition is this: the child may think that the extra farmer is relevant, that is, being there for some reason. If there is no actual event such (20b) that takes place, then the child may infer that the extra farmer is feeding a donkey at some point, too. Thus, the child would give a response that contradicts the experimenter's prediction in (20a).

- (20) a. Every farmer feeds a donkey.
b. Some farmer feeds a dinosaur.

This is relevant to my study. In the story where a dog kissed a cat and did not run away, to have a No answer to the target sentence as in (19c), the dog must perform some other action before ending the story; otherwise, this would violate the plausible deniability condition. Were this condition not met, children might be curious about what this dog was doing after kissing the cat. They might thus infer that at some point this dog that kissed the cat probably did run away, which is not an inference that we want.

Other pragmatic factors also include the presuppositions of restrictive RCs. According to Hamburger and Crain (1982), a sentence such as *The lion licked the zebra that jumped over the fence* cannot be used felicitously if there is only one zebra present. Specifically, the RC *that jumped over the fence* entails that one particular zebra is chosen from a set of zebras. Therefore, to meet this felicity condition, I also

use two pairs of identical animals in each story. To illustrate, consider the examples in (19c) and (19e) again. In one story, there are two identical dogs and two identical cats. In each pair, whereas one animal is involved in the target sentence, the other identical animal is doing something else in the context, for example, dancing on the sock.

One final issue to note is whether to point to the animal or the item (location) when the experimenter encounters the pair of fillers, as in (19a-b). In the story that is relevant to this pair of filler sentences, the experimenter can either consistently point to the same animal (e.g., dog throughout or cat throughout), or just point to the sock in both sentences. Here I suggest that the latter is easier than the former for children to judge. The reason is this: If the experimenter points to the dog that was dancing on the sock when uttering the first filler sentence, then there should be no immediate problem. However, when the experimenter moves on to utter the second filler counterpart while still pointing to the dog, in order to say No, pragmatically the child would need to first search for every cat in the context, which are two, and then make sure that both of them were not dancing on the sock, and then the child can finally come to the conclusion. This involves more cognitive resources than when the experimenter just points to that one prominent sock, where a dog was dancing. In the latter choice, in order to say No to the sentence in (19b), the child simply needs to make sure that there was no cat dancing on this sock. The child does not need to search for all cats and to see what they were doing in the context. Thus, pointing to animals requires more memory load for the child to recall what happened in the story

than pointing to the item. Since these are just filler items, it would be wise to make them as simple and easy as possible for children to judge.

3.3 Data Coding

Experimenter 2 recorded each participant's responses on a paper score sheet. The researcher carefully checked the responses and transferred them from paper to a spreadsheet. Participants' responses were classified into "yes" or "no." Cases where participants refused to answer were not analyzed.

In addition, sometimes there were situations where participants' responses were not clear. For example, they sometimes gave both a nodding and a shaking head without saying yes or no to the puppet. When these situations occurred, they were presented with the test sentence again and their additional response to the same sentence was recorded on the score sheet as well. Then, their response was transferred to the spreadsheet.

Furthermore, the experimenters asked random questions about both target and filler sentences to the participants. They asked these additional questions in order to make sure that children's responses truly reflect their grammatical knowledge. Specifically, the experimenters did not restrict their random questions to any particular type of the participant's response; they asked questions after a yes response, a no response, a correct response, an incorrect response, a no response, or a confusing response. In such cases, both the questions and participants' comments were recorded on the paper sheet.

A yes/no response was scored 1 if it matched what happened in the story; otherwise, it was scored 0.

4 Results and Analyses

As noted earlier, I examine only data from participants who passed the pre-test and did not show a response bias (over 70% yes in fillers where the target percentage of yes is 50%). There were 68 children and 22 adults whose data were analysed.

Correct responses to the filler sentences show that subjects overwhelmingly understood the game: 97% of children's responses to the fillers were accurate (with the younger age group having 95% accuracy). The true fillers were accepted 99% of the time and the false fillers were rejected 94% of the time. The children's responses to the fillers motivate confidence that when they said "yes" to a meaning-utterance pair, it was acceptable to them, and that when they said "no" to a meaning-utterance pair, it was unacceptable to them. Such high accuracy in the filler items not only indicates that the experimental items are solid, but also shows that the remaining children do not have a strong bias towards yes or no. (Also see Appendix D for the percentage accuracy of yes/no responses in fillers for the child group.)

The percentages of correct responses for both the adult and child groups in each type of RC are reported in Table 4.

Table 4: Percentage of Correct Responses by RC Type and Age Group

RC type	Age Group			
	Adult	Child <i>younger</i>	<i>middle</i>	<i>older</i>
Subject Gapped RC	100%	67%	83%	81%
Object Gapped RC	99%	58%	63%	79%
Oblique Object Resumptive RC	89%	66%	78%	86%

In the child group, I also ranked the 68 subjects from youngest to oldest and divided

them into three groups: The 20 children in the younger group ranged from 2;9 to 4;11 and had a mean age of 4;0.¹⁸ The 31 children in the middle group ranged from 5;0 to 5;11 and had a mean age of 5;4. The 17 children in the older group ranged from 6;0 to 6;10 and had a mean age of 6;3. This piece of information is summarized in Table 5. In addition, Figure 2 illustrates the means (and error bars) for three types of RCs for each age group.

Table 5: Information of Each Child Group

<i>Age group/Information</i>	<i>number</i>	<i>age range</i>	<i>mean age</i>
younger age	20	2;9 to 4;11	4;0
middle age	31	5;0 to 5;11	5;4
older age	17	6;0 to 6;10	6;3

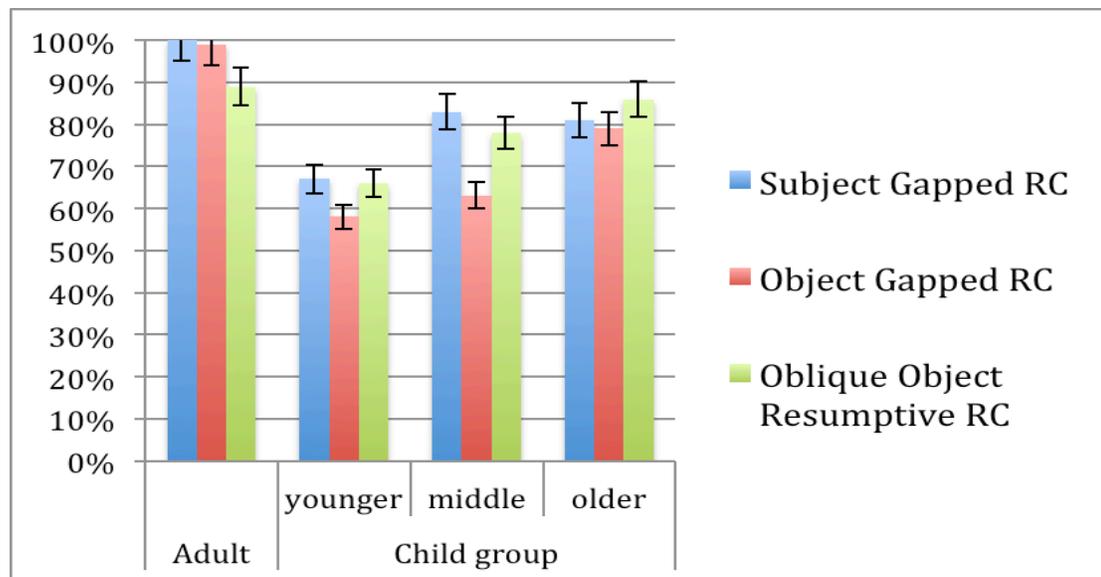


Figure 2: Means for Group by RC Type for Correct Comprehension Percentage

¹⁸ Due to the small number of 2-year-olds (*one*), 3-year-olds (*eight*) and 4-year-olds (*nine*), I collapsed these age groups into one.

The asymmetry in correct responses among three types of RCs was striking among children (e.g., 63% for the middle age in the Object Gapped RC vs. 83% in the Subject Gapped RC, compared to the older age figures of 79% vs. 81%). These effects were tested for significance with a two-way mixed model ANOVA with age group (adult vs. child: younger/middle/older) as the between-subject factor and RC types (Subject Gapped/Object Gapped/Oblique Object Resumptive) as the within-subject factor. The analysis revealed a significant main effect of age group ($F(3, 86) = 21.85, p < 0.001$) and a main effect of types of RCs ($F(2, 86) = 6.24, p = 0.003$). Furthermore, there was a significant interaction between age group and RC types ($F(6, 86) = 4.20, p = 0.01$).

Simple main effects analysis showed that, for the younger group, Subject Gapped RCs led to a better performance than Object Gapped RCs ($F(3, 86) = 2.79, p = 0.021$). For the middle group, Subject Gapped RCs led to a better performance than Object Gapped RCs ($F(3, 86) = 15.98, p < 0.001$), and Oblique Object Resumptive RCs led to a better performance than Object Gapped RCs as well ($F(3, 86) = 15.98, p = 0.001$). For adults, Subject Gapped RCs led to a better performance than Oblique Object Resumptive RCs ($F(3, 86) = 3.55, p = 0.01$), and Object Gapped RCs led to a better performance than Oblique Object Resumptive RCs as well ($F(3, 86) = 3.55, p = 0.048$). In addition, there was a simple main effect of age group on types of RCs. For Object Gapped RCs, older group and adults performed better than younger and middle groups (older group: $F(3, 86) = 23.11, p < 0.001$; adults: $F(3, 86) = 23.11, p < 0.001$).

In terms of each age group on each type of RC, the younger group tended to perform worse on Object Gapped RCs. The middle group seemed to catch up with the older group in both Subject Gapped RCs and Oblique Object Resumptive RCs but not in Object Gapped RCs. These findings suggest a developmental change among children in different types of RCs: The older they grow, the more they improve their performance in RCs. However, it is not quite clear why adults' performance on Oblique Object Resumptive RCs was worse than the other two types of RCs.

A paired *t*-test was also conducted in each pair of RCs across the child groups in order to see whether there is any significant difference among types of RCs. The results showed that the performance on Subject Gapped RCs was significantly better than on Object Gapped RCs; $t(67)= 4.2375$, $p < 0.001$; performance on Oblique Object Resumptive RCs was significantly better than on Object Gapped RCs; $t(67)= 3.0686$, $p < 0.05$; however, the difference of performance between Subject Gapped RCs and Oblique Object Resumptive RCs did not reach significance; $t(67)= 0.66634$, $p = 0.5$.

Overall, the results of Chinese-speaking children's performance on the three types of RCs indicate that they performed equally well on both Subject Gapped RCs and Oblique Object Resumptive RCs, and performed worst on Object Gapped RCs. Their performance is also summarized as in (21) (where ">" means "better than"):

(21) Subject Gapped RCs = Oblique Object Resumptive RCs > Object Gapped RCs

5 General Discussion

Two of my results align with previous studies on the acquisition of RCs in different languages. The first result pertains to the hierarchical relation about the subject-object asymmetry in RCs. That is, Chinese-speaking children also performed better on Subject Gapped RCs than on Object Gapped RCs. This result also supports the Structural Distance Hypothesis.

The second result suggests that different RC formation strategies have an influence on the acquisition of RCs along with the hierarchical structure. Chinese-speaking children's better performance on Oblique Object Resumptive RCs than on Object Gapped RCs suggest that relatives that are lower on the AH can still surpass their rivals which are higher on the AH. Unlike the first result, the second result does not support the Structural Distance Hypothesis.

Moreover, there are two new findings. First, Chinese-speaking children's performance on Subject Gapped RCs was as good as their performance on Oblique Object Resumptive RCs. Second, children performed better on Oblique Object Resumptive RCs than Object Gapped RCs. Here I explore one possible explanation for these new findings. To begin with, consider the sentence with Oblique Object Resumptive RCs, as in (3) (repeated here as (22)):

(22) Oblique Object Resumptive RC

[qingwa dui ta tiaowu de] yazi pao-zou le.
frog to it dance REL duck run-away ASP

‘The duck that the frog danced for ran away.’

Notice that the prepositional phrase (PP) *dui ta* ‘for it’ that takes the resumptive pronoun precedes the verb *tiaowu* ‘dance’ in the relative clause. Instead of focusing on the hierarchical structure of different arguments in syntax, Chinese-speaking children may just find it easier to recognize any argument that comes *before* the verb in the RC. This linear relation would treat direct objects as one group, and subjects and adjuncts as the other. But to see whether it is the hierarchical structure or the linear relation that plays a major role in determining the RC acquisition, more studies on different types of RCs are required. In particular, one can compare RCs that relativize an adjunct that follows the verb with RCs that relativize an adjunct that precedes the verb. Chinese does not seem to offer this option of comparison, since adjuncts in Mandarin usually come before the verb (see Huang 1982).

Alternatively, I could compare the indirect object RC and the direct object RC that are involved in the double object construction, since both arguments precede the verb in the RC and they use different strategies to form relative clauses. The indirect object position is filled with a resumptive pronoun when relativized, as in (23), whereas the direct object position is a gap when it is relativized, as in (24).

(23) Indirect Object RC in the Double Object Construction

[qingwa	song	ta _i	hua	de]	yazi _i	pao-zou	le.
frog	give	it	flowers	REL	duck	run-away	ASP

‘The duck that the frog gave flowers ran away.’

(24) Direct Object RC in the Double Object Construction

[qingwa song yazi ____i de] hua_i huai le.
frog give duck REL flowers wither ASP
'The flowers that the frog gave the duck withered.'

Thus, my data is not conclusive, and the Structural Distance Hypothesis cannot fully account for Chinese-speaking children's performance in the current experimental study.

6 Conclusion

This chapter reports a study of Chinese-speaking children's comprehension of RCs. The study tests three types of RCs: Subject Gapped RCs, Object Gapped RCs, and Oblique Object Resumptive RCs. Data comes from 68 children aged 2;9 to 6;10 and 22 adults who completed a truth value judgment task. There are two new findings. One is that Chinese-speaking children's performance on Subject Gapped RCs was as good as their performance on Oblique Object Resumptive RCs. The other is that they performed better on Oblique Object Resumptive RCs than Object Gapped RCs.

The UG-derived proposal, Structural Distance Hypothesis, is able to account for why children performed better on Subject Gapped RCs than on Object Gapped RCs. However, it does not provide a good explanation of why children performed better on Oblique Object Resumptive RCs than on Object Gapped RCs and also why they performed equally well on both Subject Gapped and Oblique Object Resumptive RCs. Thus, the Structural Distance Hypothesis is unable to gain full support.

In the following two chapters, I start with a review of non-UG approaches to

language acquisition in Chapter 4, and then I conduct corpus analyses to examine whether the non-UG proposal, the Frequency of Exposure Hypothesis, can account for the same set of acquisition data in Chapter 5.

CHAPTER 4: NON-UG APPROACHES TO LANGUAGE ACQUISITION

1 Introduction

In this chapter, I discuss two non-UG approaches to language acquisition, which dispute whether a theory that postulates the existence of innate grammar, namely, the initial state of the language faculty, is necessary (e.g., Chomsky 1965, 1986), and raise the question of to what extent the learner can draw statistical information from the linguistic input. One approach depends on frequency information extracted from the input, which has been considered an important factor in the acquisition of syntactic structures within various frameworks (e.g., Goldberg et al. 2004; Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007; Huttenlocher et al. 2002; Theakston et al. 2004; Diessel 2007). The other is the Bayesian inference approach to language acquisition; it focuses on a learner's capacity to use Bayesian inference to learn the complete knowledge of grammar even with sparse data (e.g., Regier & Gahl 2004; Xu & Tenenbaum 2007; Kemp et al. 2007; Perfors et al. 2010; Perfors et al. 2006, 2011, among others).

Before laying out these non-UG approaches, I provide a brief overview of the background of the debate over the existence of poor vs. sufficient data. Dating back to the sixties, Chomsky (1965), for example, has argued that there is not enough data in the corpus of material available to the acquiring child to learn a language—the “poverty of the stimulus argument”:

A consideration of the character of the grammar that is acquired, the degenerate quality and narrowly limited extent of the available data, the

striking uniformity of the resulting grammars, and their independence of intelligence, motivation, and emotional state, over wide ranges of variation, leave little hope that much of the structure of the language can be learned by an organism initially uninformed as to its general character.

(Chomsky 1965: 58)

The poverty of stimulus argument is fortified by examples of grammar learning that appear to manifest “structure-dependent” rules, since hierarchical structure is not obviously a feature of the overt language input, which consists of linearly ordered strings of words. One of these examples is subject-auxiliary inversion in yes-no questions (e.g., Chomsky 1965; Berwick et al. 2011). In a yes-no question as in (1a), English reverses the order of the subject noun phrase *this woman* and the auxiliary *can* compared to its declarative counterpart as in (1b).

- (1) a. Can this woman play the guitar?
b. This woman can play the guitar.

This kind of displacement, however, is not affected by the length or internal complexity of the noun phrase but only by the way that the structure situates these elements. Consider (2a-b):

- (2) a. Can that woman who will wear a blue hat play the guitar?
b. That woman who will wear a blue hat can play the guitar.

In (2a), the main clause auxiliary inverts with the entire noun phrase that occurs in the subject position—even when that noun phrase, as in (2b), itself contains an auxiliary

which precedes the main clause auxiliary in the linear string. The fact that it is the subject-auxiliary relationship in the main clause that matters for auxiliary inversion, rather than the presence of an auxiliary in an arbitrary linear string of words, is an argument for structural dependence. Thus, according to Chomsky, the grammar utilizes rules that must be structure-dependent. In other words, if separated from the structure, operations in the grammar find no way to work. The poverty of stimulus argument says that such a requirement of structure-dependency in grammatical operations cannot be detected in the input, whether such operations are formulated in the early framework of transformational grammar (e.g., Chomsky 1965) or in the current Minimalist Program (e.g., Chomsky 2011).

Crucially, neither the frequency-based approach nor the Bayesian approach agrees that input is deficient. Rather, both have examined the potential inherent in the statistical information that a learner can potentially draw from the input, and investigated its role in language acquisition. Several frequency-based studies have shown that children's acquisition of syntactic constructions is highly influenced by frequency information in the input (e.g., Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007). Based on this line of research, lexical items with higher frequencies and regular distributions play a vital role in the acquisition of syntactic structures.

In another line of thinking, Bayesian approaches to language acquisition propose that a learner mainly utilizes Bayesian statistical inference, which is treated as a general cognitive capacity, to learn a language, and that with this technique it is possible to learn the grammar even with few examples (e.g., Perfors et al. 2010; Perfors et al. 2006, 2011). More interestingly, some Bayesian studies have claimed to

prove that a learner with Bayesian inference can learn from the input the kind of grammar that is specified by hierarchical phrase structures, hence suggesting that there is no need for the existence of innate grammar.

These two non-UG approaches taken together share some similarities. Both argue that there is more information in the input, especially the statistical information that a learner can draw from, than generative linguists have traditionally assumed. They argue that the input data is sufficient to form a grammar. Both challenge the idea that language acquisition must begin with a set of innately endowed grammatical principles that interact with presented experience. Most importantly, they propose that language acquisition starts with the learning of specific items from the input and extends to abstract principles and constructions.

They contrast with each other in their predictions of how much can be learned or inferred from the observed data. For instance, the frequency approach indicates that, even though frequency is the key in the acquisition of syntax, it is not the only factor that contributes to the acquisition of a grammar; other factors such as social skills and discourse functions also contribute to syntactic knowledge. By contrast, some Bayesian proposals have claimed that they can obtain whatever linguistic features from the input a learner needs, including the features for a learner to derive structured representations of sentences. Since the Bayesian approach appears to be able to extrapolate more statistical information from the input than the frequency approach, I spend relatively more time discussing this approach in the following sections.

This chapter is organized as follows. In Section 2, I illustrate the main ideas of the frequency approach by reviewing some major works on syntax acquisition. In Section 3, I lay out the Bayesian approaches to the acquisition of syntactic knowledge. Section 4 is the conclusion.

2 Frequency in the Acquisition of Syntax

The frequency information that a learner can draw from the input has been shown to be key in the acquisition of syntactic knowledge (Huttenlocher et al. 2002; Goldberg et al. 2004; Theakston et al. 2004; Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007; among others). Under the usage-based framework, the learner uses lexically specific items that are frequent and distributionally regular to learn more abstract structures or syntactic knowledge. For instance, Matthews et al. (2005) did a production study to examine whether the input frequencies of English verbs play a role in children's use of word order. They tested two groups of children whose mean ages were 2;9 and 3;9, respectively. The experimental materials consist of four high frequency, four medium frequency and four low frequency verbs that were chosen based on verb frequency counts in the child directed speech of the 12 mothers in the Manchester corpus (Theakston et al. 2001), which is available on the CHILDES database (MacWhinney 2000). In particular, it was made sure that verbs chosen across frequency conditions belong to the same semantic class and must be transitive verbs, such as *'push'*, *'shove'*, and *'ram'*, so that the verbs could presumably be used in the same scenario. Each child was shown videos of hand puppets acting out the verbs. The experimenter described the videos using the verbs in SOV order. The results showed that all 2 year olds reverted to SVO word order in the high frequency

verb condition, but they were more likely to use SOV word order in the lower frequency verbs than in the higher frequency ones. As for the 3 year olds, they generally reverted to SVO order regardless of which frequency condition they were in. Based on these results, they suggest that it is frequency information that influences children's acquisition of word orders. The younger children's preference of the canonical English word order SVO or the deviant (for English) word order SOV is mainly determined by the frequency of verbs that were used.

Kidd et al. (2010) examined whether frequencies of individual lexical items are associated with children's acquisition of argument structure. Specifically, they looked at whether the frequency of a verb taking a finite complement clause would influence the acquisition of the finite complement clause construction. They tested sixty children aged 4- and 6-years using a sentence recall/lexical priming experiment. The materials consist of four pairs of complement-taking verbs with both high and low frequencies. The frequency counts were calculated on the CHILDES database (MacWhinney 2000). In the experiment, the experimenter presented a picture to the child and described it using either a high-frequency or a low-frequency complement-taking verb. The child was asked to remember how to describe that picture, which is the target sentence. Then, the child was presented with another picture, which was described by a prime sentence using a verb that shares semantic similarity but is contra-matched for frequency with the target sentence. The child was then shown the target picture again and required to recall the target sentence. Overall, the results showed that the children recalled better the test sentences that had high-frequency complement-taking verbs than in the ones that had low-frequency complement-taking

verbs. The authors took these results as evidence to support the claim that the child's knowledge of complementation is linked to their acquisition of high frequency verbs that take finite complements.

Thus, based on these frequency studies of children's syntactic knowledge, it is claimed that the frequency information leads the child from learning individual lexical words with higher frequencies to more abstract syntactic representations such as word order and verb complement constructions.

In the next section, I review Kidd et al.'s (2007) frequency-based study that looks at the effects of frequency on the acquisition of relative clauses in English- and German-speaking children. Compared with another relevant frequency work on relatives (Diessel & Tomasello 2005), this work bears more directly on the current study's interest in the acquisition of relative clauses, as frequency information was argued to play the main role in children's performance on relatives. Thus, I take it to be the main work for comparison.

2.1 Kidd et al. (2007)

Kidd et al. (2007) conducted two corpora and two experimental studies to examine English- and German-speaking children's processing of subject and object relative clauses. In particular, they showed that young children are sensitive to the distributional frequencies of relatives and use this information in acquisition.

To begin with, they argue that the subject-object asymmetry that is commonly observed in psycholinguistic research on relatives is due to the lack of consideration of distributional frequency information that the learner attends to in the input. Among these they argue that two factors play a crucial role in the processing of object

relatives. One factor is the animacy of the head noun. Since the head noun that children most often use in object relatives is inanimate, it is argued that using inanimate head nouns should make object relatives easier to process. The other factor is the type of subject in object relative clauses. It is argued that, since the subject within the relative clause is often a pronoun rather than a lexical NP in naturalistic speech, using a pronoun subject in the object relative would also make it easier to process.

In their first corpora study, they conducted an analysis of speech from four English-speaking children from the CHILDES corpus (MacWhinney 2000). They extracted all object relative clauses and analyzed them for both the animacy of the head noun and the type of NP in subject position of the relative clause. Of all the head nouns, the results showed that 75% were inanimate, 10.5% animate, and the rest were ambiguous. For the type of subject of the relative clause, first person pronoun and second person pronoun make up the largest portion, 86.6%, while only 1.5% of the subjects of relative clauses are lexical NPs.

Then, they conducted a sentence repetition experiment with English-speaking 3- and 4-year-old children with subject and object relative clauses. The object relatives were manipulated for both the animacy of the head noun and the type of subject NP. The subject relatives were also manipulated for the animacy of the head noun. Overall, the results showed that, in the object relatives, children performed better on sentences with pronouns in the subject position of the relative than those with lexical NPs. They also performed better on sentences with an inanimate head noun than those with an animate head noun. In the subject relatives, children

performed better on sentences containing an animate head noun than those with an inanimate head noun. More importantly, their results showed that there was no significant difference in children's performance between subject and object relatives, contrary to the subject-object asymmetry that has been commonly observed in psycholinguistic research on the acquisition or processing of relatives. Thus, the results suggest that, when the object relatives match the type typically encountered in the input, children are just as good as they are at the subject relatives. They conclude that children are better at the types of sentences they hear most frequently.

Next, they did a corpus study of one German-speaking child's use of object relatives, who was between the ages of 2;0 and 3;0 during the recording of the corpus. Again, they looked at the animacy of the head noun and the type of NP in the subject position. The results showed that 2% of the object relatives had an animate head noun, and that 77% had an inanimate head noun. For the distribution of NP types in subject position, the results showed that first and second person pronouns make up 62% of the NPs, and that only 19% of the NPs are lexical NPs. In spite of the variability, the distributional frequency information about the animacy of head nouns and the NP types in the subject position displayed the same trend as in the English corpus.

They then conducted a sentence repetition study comparable to the English one with German-speaking 3- and 4-year-old children. In the object relatives, the results showed that children performed better on object relatives with a pronominal subject than those with a lexical NP. The children also performed better on object relatives with an inanimate head noun, but only when the subject was a lexical NP.

As for the subject relatives, children performed best on sentences with an inanimate head. There was no significant difference between these two types of relatives when the children's performance on the subject relatives was compared with their performance on the object relatives, suggesting that neither is more difficult than the other.

In summary, this study demonstrates that, the patterns of children's performance on relatives correlate with the distributional frequency information of the input.

In the next section, I describe the Bayesian approach to language learning.

3 Bayesian Models of Language Learning

This section provides background information on Bayesian models of cognition and describes how they have been applied to language learning. The basic assumption of Bayesian models of cognition is that humans use the principles of Bayesian probabilistic inference to generalize from examples. One of the most important and central questions it addresses is how human minds can learn beyond the superficially sparse data of experience (e.g., Griffiths et al. 2008; Perfors et al. 2011).

Recently, Bayesian models of cognition have been adapted to investigate how language learners make inferences about correct hypotheses given the limited linguistic data that are available to them. This approach to language learning has gained more and more popularity for its interesting results and implications. For example, Xu and Tenenbaum (2007) demonstrated that a Bayesian model is able to make inferences about different word meanings and derive the most probable meaning of a word given the limited data. Moreover, a growing body of work has

implemented a certain type of Bayesian model, namely, the hierarchical Bayesian model, to investigate the potential of the model to make inferences about overhypotheses, or, more abstract knowledge, such as the shape bias in word learning (Kemp et al. 2007). Of importance in this chapter are the studies that adopt the hierarchical Bayesian model to examine syntactic phenomena. Perfors et al. (2010) showed that a hierarchical Bayesian model could make inferences about the correct verb argument construction even in the absence of negative evidence. Further, Perfors et al. (2006, 2011) demonstrated that, by adopting a domain-general ability and Bayesian inference, the hierarchical Bayesian model is capable of learning hierarchical structures of sentences. These latter two studies bear direct relevance to the current study, as they specifically argue that innate language-specific constraints which have long been assumed by nativists such as Chomsky (1965) are not required, since the hierarchical Bayesian model has demonstrated a potential to derive abstract syntactic principles. I review these two studies in detail in Section 3.3.

The organization of this section is as follows. To explain what Bayesian probabilistic inference is, I start with a brief description of Bayes' theorem in Section 3.1. In Section 3.2, I introduce the basics of Bayesian models of cognitive development. In Section 3.3, I review previous studies that have adapted the Bayesian models to investigate language learning. Specifically, I review two studies on hierarchical Bayesian models about different syntactic phenomena, including Perfors et al. (2010) on the acquisition of verb argument constructions, and Perfors et al. (2006, 2011) on abstract syntactic principles.

3.1 *Bayes' Theorem*

Bayesian probabilistic inference uses Bayes' Theorem to update beliefs when observing new data. Bayes' theorem can be used to derive the probability of an event that occurs in light of new evidence. The theorem goes as follows:

(3)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Simply put, it states that the probability of some event A, given that B is true, $P(A/B)$, is the probability of B given that A is true, $P(B|A)$, multiplied by the probability of A, $P(A)$, and then divided by the probability of B, $P(B)$. I will give more relevant details in Section 3.2.1.

In the next section, I review some basics of Bayesian models of cognition.

3.2 *The Background of Bayesian Models of Cognition*

The essential idea of Bayesian models of cognition is that they are formulated based on understanding the logic of how one makes inferences when generalizing from evidence, rather than focusing on any physical property of the system that is under study. Thus, intrinsically the models do not require assumptions to be made about any specific cognitive processes or algorithms in natural tasks. The basic question that Bayesian approaches to cognition have tried to address is how to update beliefs and to

make inferences given some observed data. Therefore, to explain how this kind of model works, I explain its main features in the following sections. These features include degrees of belief as probabilities, as discussed in Section 3.2.1, the size principle, as discussed in Section 3.2.2, and a particular type of the Bayesian model that has claimed success in grammar acquisition, as discussed in Section 3.2.3.

3.2.1 Degrees of Belief as Probabilities

Bayesian models take as a central assumption the hypothesis that degrees of belief are represented by probabilities (e.g., Griffiths et al. 2008; Perfors et al. 2011). For example, if a person attempts to infer a process that generates some data, d , and that she has different hypotheses h_i, h_{i+1}, \dots, h_n , about this process, then she uses probabilities to represent degrees of belief in each h . Each hypothesis can be represented by the probability $P(h)$. The number of $P(h)$ ranges from 0 to 1, where 0 would mean that she believes that the hypothesis h is completely false, and that 1 would mean that she believes h is completely true. More importantly, the gist of the Bayesian models of cognition is that degrees of beliefs, now corresponding to probabilities, may be updated rationally based on the given data.

Two components are required in order to compute the degrees of belief in some hypothesis h_i given some data d . One component is called the *prior probability*, which is represented by $P(h_i)$. It is the probability that she attributes to the hypothesis h_i (for example, a hypothesis about the sentence-generating process) before observing the data d . The other component is called the *likelihood*, which is represented by $P(d|h_i)$. It is the probability of the data d given the hypothesis h_i , that is, how likely it

is that she would expect to see the data d if she already knew that the hypothesis h_i is true. When we replace B with d and A with h_i in (3), it gives (4).

$$(4) \quad P(h_i|d) = \frac{P(d|h_i)P(h_i)}{P(d)}$$

Given Bayes' Theorem and that H is a set of all mutually exclusive hypotheses, these two components can combine to compute the degrees of belief in some hypothesis h_i given some data d , $P(h_i|d)$, which is called the *posterior probability*, as in (5):

$$(5) \quad P(h_i|d) = \frac{P(d|h_i)P(h_i)}{\sum_{h' \in H} P(d|h')P(h')}$$

Thus, the posterior probability is the product of the likelihood and the prior probability.

To help clarify these ideas, I use an example from Griffiths et al. (2008). Suppose that we have two identical-looking coins in a bag, thus two hypotheses h_0 and h_1 . One coin produces heads with probability 0.5, but the other produces heads

with probability 0.9. The probabilities to of heads and tails for each hypothesis are summarized in (6).

- (6) a. h_0 : producing heads with probability 0.5 and tails with 0.5
b. h_1 : producing heads with probability 0.9 and tails with 0.1

Assume also that each of the coins will be chosen with the same chance, so both hypotheses will have equal prior probabilities: $P(h_0) = P(h_1) = 0.5$. We obtain data by flipping the coin. The probability of a particular sequence of coinflips that has a heads and b tails given h_0 is $(0.5)^a(0.5)^b$, whereas the probability of the same sequence given h_1 is $(0.9)^a(0.1)^b$. If we flip a coin ten times and obtain the sequence HHHHHHHHHH, then the probability of this sequence given h_0 is approximately 0.348678. By contrast, the probability of this sequence given h_1 is approximately 0.000976. In order to determine which hypothesis is more likely than the other to generate this sequence of data, we can take the prior probabilities and likelihoods now to compute the posterior probabilities for both hypotheses. In particular, since there are only two hypotheses in this case, we can simply compute the ratio of these two posterior probabilities. The posterior odds in favor of h_1 given the sequence HHHHHHHHHH is:

(7)

$$\frac{P(h_1|d)}{P(h_0|d)} = \frac{P(d|h_1)P(h_1)}{P(d|h_0)P(h_0)} = \frac{0.348678 * 0.5}{0.000976 * 0.5} \approx 357.25$$

Note that since $P(d)$ is the same for both hypotheses, it is not included in this computation.

In the next section, I discuss the size principle.

3.2.2 Size Principle

Another crucial concept that Tenenbaum and Griffiths (2001) incorporate into the Bayesian inference model of cognition is the *size principle*. It states that the less inclusive a hypothesis h_i is, the greater likelihood $P(d|h_i)$ that hypothesis will receive for an observed piece of data d . Consider the following formula under strong sampling (adapted from Tenenbaum & Griffiths 2001: 633):

$$(8) \quad P(d|h) = \begin{cases} \frac{1}{|h|} & \text{if } d \in h \\ 0 & \text{otherwise} \end{cases}$$

Here $|h|$ refers to the size of the area of hypothesis h . This formula shows that the probability of a piece of data d given that h is true is inversely proportional to the size of h . When the area of h consists of discrete spaces, $|h|$ is the cardinality, or, the number of elements of the subset that is associated with h . When the area of h is

composed of continuous spaces, $|h|$ would correspond to the *measure* of the hypothesis, for instance, the length of the interval in one dimension. In other words, if there are two hypotheses that both generalize to a given piece of data d , and one is smaller h_s , the other is larger h_l , then the smaller hypothesis h_s will receive a higher likelihood $P(d|h_s)$ than the larger hypothesis $P(d|h_l)$.

A more concrete example of the size principle is illustrated as follows. Imagine that there are two bags A and B. Bag A has two marbles in it: one red and one green. Bag B has three marbles in it: one red, one green and one yellow. Thus, the probability of choosing the red marble out of bag A is $\frac{1}{2} = 0.5$, whereas the probability of choosing the red marble out of bag B is $\frac{1}{3} \approx 0.33$. If it is already known that a red marble has been picked up out of a bag, without knowing which bag, then it is rational for a person to show a bias toward bag A. The size principle encodes the reasoning behind this preference.

3.2.3 The Hierarchical Bayesian Model

As the hierarchical Bayesian model of language learning has demonstrated its potential to generalize abstract knowledge about language, especially abstract syntactic knowledge (e.g., Perfors et al. 2010, Perfors et al. 2011), this section serves to explain this particular feature of hierarchical Bayesian models.

A hierarchical Bayesian model can make inferences about data on multiple levels, including item-based hypotheses and higher-level hypotheses about the hypothesis space of those item-based or lower-level hypotheses. These higher-level hypotheses are sometimes called overhypotheses, or abstract knowledge. The

hypothesis space corresponds to the set of hypotheses that are currently considered in the model. One of the significant features of the hierarchical model is that it differs from non-hierarchical models in that it does not need to specify parameters that are used to define lower-level hypotheses. Rather, it only needs to specify the parameters of highest overhypotheses at most and can learn the range of parameters that define lower-level hypotheses based on few observed examples (see e.g., Kemp et al. 2007; Perfors et al. 2011).

In the next section, I turn to the discussion of Bayesian models of language learning, and examine some important studies on syntax learning.

3.3 Bayesian Models of Language Learning

From the perspective of a Bayesian framework on language learning, the corresponding components that can be used to apply Bayes' Theorem are determined by the type of the examined linguistic expressions. Take word learning as an example. Xu and Tenenbaum (2007) took each word meaning to be a hypothesis, and these meanings constitute the hypothesis space of a novel word. The learner will update the probabilities of these hypotheses given what word meanings of the word have been supplied in the input. At the level of grammar learning, Chater and Manning (2006) reframe Chomsky (1965), where language acquisition is considered choosing among a hypothesis-space of (innate) candidate grammars. They indicate that, in the Bayesian framework, each candidate grammar that a learner has can be represented by a prior probability, and the probabilities of all these candidate grammars will be adjusted based on Bayes' Theorem given some data. The learner will then choose a grammar, or, a language that has the highest posterior probability.

Such a Bayesian approach to grammar learning has been applied in Perfors et al. (2010) on learning verb argument construction and in Perfors et al. (2006, 2011) on learning abstract syntactic principles.

I review these latter two studies on Bayesian approaches to syntax in Section 3.3.1 and Section 3.3.2, respectively.

3.3.1 Perfors et al. (2010)

Perfors et al. (2010) conducted three studies to examine the acquisition of verb argument constructions using a hierarchical Bayesian framework. They showed that by using a domain-general approach, this model can make inferences about the distributional statistics of verbs on both the item-specific level and the verb-general knowledge level. They also take this model to examine the negative evidence problem. For example, in English the verb *load* can appear in two kinds of locative constructions, as in (9a-b) (Perfors et al. 2010: 608):

- (9) a. He loaded apples into the cart.
b. He loaded the cart with apples.

But there are similar verbs such as *pour* that cannot alternate between these two constructions, as in (10b):

- (10) a. He poured apples into the cart.
b. *He poured the cart with apples.

This poses a problem as to how speakers learn which verbs can alternate and which cannot between these constructions given no explicit negative evidence.

To explore these questions, they speculate about three levels of knowledge that will be considered in the Bayesian models in their studies. Level 1 knowledge concerns how often each verb co-occurs with each construction. In their model, this knowledge is represented by θ . Level 2 knowledge concerns the distribution of verb constructions, which is represented by two parameters, α and β . The parameter α is related to how consistently each individual verb occurs in one construction, and the parameter β corresponds to the overall frequency of each construction. Level 3 knowledge deals with prior knowledge about Level 2 knowledge. Specifically, Level 3 knowledge includes the range of values that is expected about the uniformity of constructions within the verb, and the range of values that is expected about the distribution of verb constructions across the language. These are represented by two parameters, λ and μ , respectively. In such a hierarchical Bayesian model, learning takes place when the model makes inferences about higher-level parameters based on data that are observed in the input. The model will then predict the parameters of novel verbs. For a new verb x , the model can make inferences about the construction distribution with the highest probability.

In study 1, they consider two models, Model L2 and Model L3. Both models can make inferences on multiple levels. While Model L2 assumes that the Level 3

knowledge is already known, Model L3 assumes that knowledge at higher levels than Level 3 knowledge is known. Thus, Model L2 learns the α and β values that correspond to the highest posterior probability for the given data. Model L3 learns not only α and β , but also λ and μ . Both models learn Level 1 knowledge.

The constructions that they examine in study 1 are VAP (verb agent patient) and VPA-*ka* (verb patient agent particle (*ka*)) of an artificial language, which are adapted from Wonnacott et al. (2008). Five novel nouns and eight novel verbs occur in one of these constructions. Two conditions are considered. In the Generalist condition, each of the eight verbs occurs in both constructions, and each verb occurs seven times as often in the VPA-*ka* construction as in the VAP. In the Lexicalist condition, seven verbs occur in the VPA-*ka* construction only, and one verb occurs in the VAP only. A completely novel verb is also presented in either VAP or VPA-*ka* construction in order to assess whether the model can generalize beyond learned lexical items.

The results show that both models can make inferences about verb-general knowledge, and can use that to make generalizations about novel input. In both Model L2 and Model L3, novel verbs occur in both constructions when they are in the Generalist condition, and the same novel verbs occur in only one construction when they are in the Lexicalist condition.

In study 2, they continue to examine whether Model L2 and Model L3 can learn verb alternations given that they already possess verb-general knowledge as shown in study 1. The rationale behind this study is that, even though both models can acquire verb-general knowledge, it is not clear whether they can predict that

sometimes only a subset of verbs but not all the verbs exhibit the same verb-general properties. Similar to the locative examples mentioned above, for example, some English verbs can occur in both the prepositional dative and direct object dative constructions, as in (11a) and (11b), respectively (adapted from Perfors et al. 2010: 610):

- (11) a. Debbie gave a pretzel to Dean.
b. Debbie gave Dean a pretzel.

But some verbs cannot alternate between two constructions, as in (12a) and (12b), respectively (data from Perfors et al. 2010: 610):

- (12) a. Jonathan confessed the truth to Doug.
b. *Jonathan confessed Doug the truth.

To examine this question, they extend both models to allow them to discover verb classes based on the data, which are represented by Model K-L2 and Model K-L3 thereafter. They use real-world data from sentences spoken by adults in the Adam corpus (Brown, 1973) of the CHILDES database (MacWhinney, 2000), and focus on the dative alternation.

The results show that Model K-L3 makes appropriate predictions about verb classes. The verbs that are associated with one construction are not treated the same class as those associated with both constructions.

In study 3, since it has been shown that there are strong correlations between verb semantics and verb syntax, they extend the previous questions and further examine how semantic features can contribute to the generalization of verb constructions. They incorporate semantic features into the current model. They introduce a semantic feature to the corpus which corresponds with different classes of verbs. Since the semantics of the dative corpus can be divided into three classes, there are three possible values of this semantic feature. One value is associated with the class of alternating verbs, termed “semantic class A”. One value corresponds to verbs that occur only in the prepositional dative construction, termed “semantic class P”. The last value corresponds to verbs that occur only in the direct object dative construction, termed “semantic class D”.

As predicted, the results indicate that the model exhibits less overgeneralization than equivalent models that do not utilize semantic features.

In brief, they demonstrated that a domain-general hierarchical Bayesian model with only syntactic input can generalize the usage of constructions even without explicit negative evidence. It is also argued that, since a domain-general model can accomplish the task of learning language, there is no need to assume any language-specific or innate linguistic constraints.

3.3.2 Perfors et al. (2006, 2011)

Perfors et al. (2006, 2011) use the Bayesian framework to investigate the learnability of abstract syntactic principles, and specifically, they look at the phenomenon of

auxiliary fronting in English interrogatives. Some interrogatives correspond to simple declaratives, as in (13):

- (13) a. Eagles can fly.
b. Can eagles fly?

Others have a more complex form, as the subject noun phrase is modified by a relative clause:

- (14) a. Eagles that are alive can fly.
b. Can eagles that are alive fly?

To begin with, they assume that an ideal human language learner has two domain-general abilities that are powerful enough to enable her to *learn* hierarchical phrase-structure grammars. First, the learner has the potential to represent structured grammars of different forms. The hierarchical phrase-structure grammars are merely a subset of those structured grammars that she can represent. Second, the learner can use the principles of Bayesian probabilistic inference to update the probabilities of each different structured grammar given observed data.

The hypothesis space of grammars in their study consists of seven grammars. Two of them are context-free grammars (CFGs), which are treated as grammars with hierarchical phrase structure. These include CFG-S, which is smaller but fits the data less precisely, and CFG-L, which is larger and fits the full corpus more precisely. The

other five grammars are considered without hierarchical phrase structure. Three of them are based on regular-grammars, or, finite-state grammars. They are REG-B, the smallest with the least precise fit, REG-N, the largest with most precise fit, and REG-M, which takes a middle ground. Another grammar is what they term as FLAT grammar, and it is just a memorized list of all sentence types. The last grammar is a one-state (1-ST) grammar. This grammar has the lowest fit, but is the simplest; it allows any syntactic category to come after any other.

The corpus that they use to examine these grammars is composed of sentences spoken by adults in the Adam corpus (Brown, 1973) from the CHILDES database (MacWhinney, 2000). In their corpus, they replace each word with its syntactic category and remove complicated sentence types. The final corpus has 21,671 individual sentence tokens, and 2,336 distinctive sentence types.

Their results show that both context-free grammars, which correspond to hierarchical phrase structures, have higher posterior probabilities than the other grammars in the full corpus. They conclude that their learning analysis thus can infer that a grammar with hierarchical phrase structures accounts for the syntax of English better than a grammar with linear structures.

More importantly, they show that both context-free grammars can parse both simple and complex interrogative sentences such as (13b) and (14b), even though they have not encountered any example of these forms in the input. In contrast, the three regular grammars and FLAT grammar can only parse simple interrogatives.

Overall, they use these results to argue that a language learner, who is equipped with both the capacity to represent structured grammars and the capacity to

access Bayesian probabilistic inference, can infer that a language needs a grammar with hierarchical phrase structure based on evidence in the input. Since these two capabilities are domain-general, they suggest that there may be no need for domain-specific predispositions in language acquisition. In other words, they suggest that the input contains enough information for a learner to generalize a grammar, contra what generativists have assumed (e.g., Chomsky 1965).

4 Conclusion

In this chapter, I reviewed two non-UG approaches to syntax learning. These include the frequency approach, which mostly lies within the usage-based frameworks, and the Bayesian approach. Both approaches bear on the issue of whether a theory that postulates the existence of the Universal Grammar is required (i.e., Chomsky 1965, 1986), and on the question of the extent to which the learner can draw from the linguistic input the statistical information.

The frequency approach in different frameworks has demonstrated that in certain contexts such as the learning of word order, verb complement constructions and subject and object relative clauses, children's performance correlates with the frequencies of particular lexical items of the input. The frequency approach suggests that there is sufficient statistical information in the input. If true, such a claim would serve to weaken the poverty of stimulus argument.

The Bayesian approach considers Bayesian inference to be the main general cognitive capacity which enables children to learn a grammar. More importantly, a particular type of Bayesian model, namely, the hierarchical Bayesian model, has been shown to be able to learn grammars with hierarchical structured representations.

Thus, according to some researchers working on the Bayesian framework, there is no need for a theory of language to postulate the existence of an innate grammar, since the input provides enough information for the learner to learn a language. What the learner needs to learn a language are just few examples in the input and general cognitive capacities that include Bayesian inference and the ability to represent hierarchical structures.

CHAPTER 5: A CORPUS ANALYSIS OF CHINESE RELATIVE CLAUSES AND THEIR IMPLICATIONS FOR NON-UG STATISTICAL APPROACHES

1 Introduction

The purpose of the corpus study discussed within this chapter is two-fold. First, it serves to contribute to a relatively small body of work on corpus analyses of Chinese relative clauses. Many of the previous studies that involve corpus analyses of Chinese relatives have focused on adults' processing of subject and object relatives, and therefore these analyses are mostly restricted to these two types (e.g., Hsiao & Gibson 2003; Wu 2009; Chen et al. 2012; Jäger et al. 2015, among others). The results in these studies consistently showed that subject relatives are more frequent than object relatives.

Second, it compares the results of the corpus analyses with my experimental study in Chapter 3 on Chinese-speaking children's comprehension of Subject Gapped relatives, Object Gapped relatives, and Oblique Object Resumptive relatives. In particular, that study showed that children before age 6 performed significantly better on Subject Gapped relatives than on Object Gapped relatives. Furthermore, whereas there was no significant difference between Subject Gapped relatives and Oblique Object Resumptive relatives among them, children between 5;0 and 5;11 performed significantly better on Oblique Object Resumptive relatives than Object Gapped relatives. The overall results from the study showed that children performed equally well on Subject Gapped and Oblique Object Resumptive relatives, and worst on Object Gapped relatives. Thus, this corpus study examines the frequency of these

three types of relatives more carefully, and to examine how this may or may not fit in to frequency-based or Bayesian approaches to language acquisition (e.g., Goldberg et al. 2004; Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007; Huttenlocher et al. 2002; Theakston et al. 2004), and Bayesian approaches (e.g., Xu & Tenenbaum 2007; Kemp et al. 2007; Perfors et al. 2010; Perfors et al. 2006, 2011).

This chapter is organized as follows. In Section 2, I review the corpus analyses of Chinese relatives from previous studies. In Section 3, I describe the methods and tools that are used in the current study. In Section 4, I present and discuss the results of the corpus analysis. In Section 5, I compare the results from the corpus analyses with my experimental results and discuss how they fit in the frequency-based approaches and Bayesian approaches to language acquisition. Section 6 is the conclusion.

2 Previous Corpus Analyses of Chinese Relative Clauses

In this section, I review five corpus studies on Chinese relatives. Four of them analyze different versions of the Chinese Treebank (Hsiao & Gibson 2003; Wu 2009; Hsiao & MacDonald 2013; Jäger et al. 2015), each of which is discussed in Section 2.1. The last one analyzes relatives in children's spontaneous speech (Chen & Shirai 2015), and is discussed in Section 2.2.

2.1 Corpus Studies of the Chinese Treebank

To pave the way for the discussion that follows, I provide some basic properties of the Chinese Treebank and an overview of the four studies that analyze the Treebank. The data in the Penn Chinese Treebank consists of news and magazine articles from

Xinhua newswire, Hong Kong News and Sinorama magazine (now called Taiwan Panorama magazine). The corpus was parsed by the Penn Chinese Treebank Project (e.g., Xue et al. 2005). Among the four studies that analyze the Chinese Treebank, three focus on adults' processing of relatives. Hsiao and Gibson (2003) examine the frequency of subject and object relatives. Wu (2009) looks at additional factors that may affect adults' processing, including the syntactic role of the matrix clause that the relative modifies and the animacy of the modified noun head. Jäger et al. (2015) investigate whether the conditional probability of subject and object relatives play a role in adults' processing of relatives. The last study is Hsiao and MacDonald (2013). They analyze subject and object relatives in the Treebank, and, based upon it, assess a computational model of relative clause processing.

Hsiao and Gibson (2003) analyzed the Chinese Treebank (version 3.0) and examined only active relatives that modify arguments, i.e., thus no passives nor adjunct relatives (e.g., *the reason why he left*), for the purpose of designing their experiment. They extracted 882 instances of relatives, and 375 (42.5%) of them are object relatives, whereas the remaining 507(57.5%) are subject relatives.

Wu (2009) analyzed the Chinese Treebank (version 5.0) and extracted several pieces of distributional information on relatives. The first piece of information is the frequency of subject and object relatives in both subject-modifying and object-modifying positions. The results showed that subject relatives are more frequent than object relatives in both modifying positions. The second piece of information is the distribution of different grammatical positions in the matrix clause that are modified by relatives. The results indicate that out of 1,209 relatives that were found in the

corpus, 530 (43.83%) of them modify the subject of a matrix clause, 292 (24%) modify the direct object, and 229 (18.94%) modify the object of a preposition (e.g., ‘*John and Mary talked in the room where Bill was killed*’). The other positions found to be modified by relatives include predicate nominal (e.g., ‘*This is a great achievement that makes him proud.*’), existential head noun (e.g., ‘*a barn that has a lot of lambs*’), and so forth, but they only comprised less than 13% of the total. Another piece of distributional information that was extracted is related to the animacy of subject and object relatives. This category consists of 331 relatives that contain transitive action verbs. Among them, 86.7% of 128 object relatives are found to be associated with inanimate heads, and thus only 13.3% with animate heads. For the 203 subject relatives, 64.5% of them occur with animate heads, and 35.5% with inanimate heads. The overall results in this study showed that object relatives tend to occur with inanimate heads more often than animate heads in general, whereas subject relatives are only predominantly attached to animate heads when occurring in the subject position of a matrix clause.

Jäger et al. (2015) carried out corpus searches on relatives in the Chinese Treebank (version 7.0). In particular, they eliminated all syntactic configurations that might cause local ambiguities, i.e., different potential ways to parse the clauses, including analyzing the relativizer *de* as a possessive marker, analyzing a subject gap as a covert pronominal, and so forth.¹⁹ They extracted all tokens that match

¹⁹ In Chinese, the relativizer *de* and the possessive marker *de* are homonyms, with the same character and same pronunciation. In addition, Chinese allows subjects and objects to drop from finite sentences (e.g., Huang 1989), and the references of these covert pronominals can be recovered from the context.

Det(erminer) + Cl(assifier) + Adv(erb) sequences with the determiner being restricted to just *zhe* ‘this’ or *na* ‘that,’ as in (1) (adapted from Jäger et al. 2015: 103):

- (1) na-ge [RC t_i shanggeyue yaoqing-le nanhai ji-ci
 that-CL last.month invite-ASP boy several-CL
 de] nūhai renshi Wang.laoshi...
 REL girl know Wang.teacher
 ‘The girl who invited the boy several times last month knows teacher
 Wang...’

Then, they calculated counts for any different structure following this sequence. The results showed that 98.6% of all 73 tokens that match the Det + Cl + Adv sequence come with a relative clause. More interestingly, for relatives that modify the subject of a matrix clause, 77.5% of the tokens matching Det + Cl + Adv sequence are followed by subject relatives and 20% by object relatives. For relatives that modify the object of a matrix clause, 85% of the tokens matching the same sequence are followed by subject relatives and 15% by object relatives. The overall analysis shows that the structural probabilities of subject relatives are more frequent than object relatives.

The last study that involves the analysis of relatives in the Chinese Treebank (version 7.0) examines a computational model of relative clause processing (Hsiao & MacDonald 2013). More precisely, they extracted all subject and object relatives that modify either the subject or the object position of the matrix clause. Three pieces of

relevant frequency information are summarized here. First, there are more subject relatives than object relatives in both modifying positions. Second, in terms of the distributional frequency of the syntactic position that relatives modify, 60% of all relatives modify the subject position, and 40% modify the object position. Third, in terms of the animacy of the head noun, object relatives mostly occur with inanimate head nouns regardless of which grammatical position they modify. On the other hand, subject relatives that contain transitive verbs occur more often with animate head nouns in subject-modifying position (but not in object-modifying position), while subject relatives that contain intransitive verbs occur more often with inanimate heads than animate heads. Table 6 is the summary of the corpus results of these four studies:

Table 6: Frequency Patterns of Chinese RCs in the Adults' Corpora

	Categories of RCs	Frequency Patterns (“>” means “higher than”)
Hsiao & Gibson (2003)	Relativized head	Subject RCs > Object RCs
Wu (2009)	1. Relativized head	Subject RCs > Object RCs
	2. Grammatical role of Head Noun	Subject > Object > Object-of-Preposition > Predicate Nominal
	3. Animacy of the head noun (Subject vs. Object RCs)	Object RCs with inanimate heads; Subject RCs (+transitive verbs) with animate heads (only in subject-modifying positions)
Jäger et al. (2015)	1. Tokens that match Det + Cl + Adv order	98.6% are all relative clauses.
	2. Relativized head	Subject RCs > Object RCs
Hsiao & MacDonald (2013)	1. Relativized head	Subject RCs > Object RCs
	2. Grammatical role of Head Noun	Subject > Object
	3. Animacy of the head noun (Subject vs. Object RCs)	Object RCs with inanimate heads; Subject RCs with animate heads (+transitive verbs) and with inanimate heads (+intransitive verbs)

Thus, quite consistently, subject relatives are found to occur more frequently than object relatives in all these studies. However, it is not clear whether there are other

types of relative clauses that have a higher frequency of occurrence than any of these two types in the corpus, such as indirect/oblique object relatives as in (2a-b), since these other types have not received much attention in sentence processing studies.

(2) a. na-ge Zhangsan song ta shu de xuesheng
 that-CL Zhangsan give him book REL student
 ‘that student who Zhangsan gave books to’ (Lit. ‘that student who
 Zhangsan gave him books’)

b. na-ge Zhangsan dui ta xiao de xuesheng
 that-CL Zhangsan to him smile REL student
 ‘that student who Zhangsan smiled at’ (Lit. ‘that student who
 Zhangsan smiled at him’)

Perhaps only one of the above studies may imply that there is a lower frequency of the indirect/oblique object relatives compared to subject and object relatives, that is, Jäger et al. (2015). According to them, 98.6% of all tokens matching Det + Cl + Adv sequence are all relatives, among which 77.5% are subject relatives and 20% are object relatives. This implies that there are still 1.1% of these relatives, whichever types they are, that are neither subject nor object relatives. Note that an indirect/oblique object relative is fine after Det + Cl +Adv sequence, as shown in (3).

- (3) na-ge shanggeyue nanhai dui ta nuhuo-le
 that-CL last.month boy tu her roar.at-ASP
 ji-ci de nühai renshi Wang.laoshi
 several-CL REL girl know Wang.teacher
 ‘The girl who the boy roared at several times last month knows teacher Wang.’ (Lit. ‘The girl who the boy roared at her several times last month knows teacher Wang.’)

Due to lack of detailed frequency information of relatives other than subject and object relatives, I conducted a more comprehensive corpus analysis of relatives and the details are discussed in Section 3. In the next section, I present Cheng and Shirai’s (2015) corpus analysis on Chinese relative clauses in children’s spontaneous speech.

2.2 *Child Speech Corpus Analysis*

Chen and Shirai (2015) investigated the naturalistic data of four children from the Fang corpus (Min 1994, cited from Cheng & Shirai 2015). It is a very detailed study that examines the developmental trajectory of relative clauses in the longitudinal speech of monolingual Chinese-speaking children and their caregivers. The age of children ranged from 0;11 to 3;5, and the data were collected in Beijing during late 1980s and early 1990s.

The data analyses center around both the relativized position in the relative clause and the syntactic position of the head noun that a relative modifies. First, four categories are identified based on the relativized position, which include subject

relatives, (direct) object relatives, oblique argument relatives, including relatives that modify a location (i.e., ‘*in the cage in which small animals were enclosed*’), and relatives that they refer to as having no grammatical relation, such as gapless relatives (e.g., ‘*They chased to a place in which there is no road.*’).²⁰ The results showed that object relatives are the most frequent type, which account for 53.8%. Subject relatives are the second most frequent type, which make up 23.4%. Oblique argument relatives comprise 11% of the total, and gapless relatives amount to 11.7%. Interestingly, according to the study, this frequency trend of relatives with different relativized positions also matches the trend in caregiver speech, which is summarized in Table 7:

Table 7: Relative Frequency of RC by Relativized Position in Children and Caregivers

	Object	Subject	Gapless	Oblique arguments	Total counts
Children	53.8%	23.4%	11.7%	11%	145
Caregivers	57.4%	18.6%	15.9%	8.1%	333

In terms of the modified syntactic position, they identify five categories. These categories include subject position of a matrix clause, (direct) object position of a matrix clause, oblique object position (e.g., ‘*It is said that a child should not read in places where the light is no good.*’), a predicate NP in a copular clause (e.g., ‘*This is the house that I made.*’), and isolated NPs, i.e., relative clauses that modify an isolated NP (e.g., ‘*the board daddy bought*’). They found that the isolated relative clauses are the predominant type, which constitute 45.5% of all relatives. Subject-

²⁰ For more details of the theoretical approaches to Chinese gapless relative clauses, the interested reader is referred to Tsai (1997), Aoun and Li (2003), Cheng and Sybesma (2006), Zhang (2008) and also Tsai (2008).

modifying relatives account for 18.6%, object-modifying relatives account for approximately 16%, and relatives that modify a predicate NP in a copular clause (e.g., ‘*This is the house that I made.*’) account for 13.8% of the total. The relative that modifies an oblique NP has the lowest frequency of occurrence, which accounts for 6.2% of all relatives. This trend of frequency among relatives in child speech mostly also matched the trend in their caregiver speech, as summarized in Table 8. Also, Table 9 summarizes the frequency patterns found from this study.

Table 8: Relative Frequency of RC by Modified Syntactic Function in Children and Caregivers

	Isolated NP	Subject	Object	Predicate NP	Oblique NP	Total counts
Children	45.5%	18.6%	16%	13.8%	6.2%	145
Caregivers	27.9%	29.4%	17.7%	13.2%	12%	333

Table 9: Frequency Patterns of RCs in Children and Caregivers’ Spontaneous Speech

	Categories of Relative Clauses	Frequency Patterns
Children	1. Relativized head (Subject vs. Object vs. Oblique argument vs. Gapless RCs)	Object RCs > Subject RCs > Oblique RCs/Gapless RCs
	2. Grammatical role of Head Noun	Isolated NPs > Subject > Object > Predicate NP > Oblique object
Caregivers	1. Relativized head	Object RCs > Subject RCs > Oblique RCs/Gapless RCs
	2. Grammatical role of Head Noun	Subject > Isolated NPs > Object > Predicate NP > Oblique object

Thus, in terms of the relativized head, children produced more object relatives than subject relatives. This matches the pattern found in the caregivers' speech. Accordingly, Chen and Shirai suggest that the distributional pattern in the input is a good predictor of the child's acquisition pattern of relatives. As for the frequency pattern of grammatical positions of head nouns, isolated NPs were the most frequent type in the children's speech, while subject positions were the most frequent in the caregivers' speech. To explain this discrepancy, Chen and Shirai suggest that relatives associated with isolated NPs are structurally similar to adjective-modified or noun-modified nouns, as in (4), which are typically treated as involving simpler structures.

- (4) baba de ban
 daddy POSS board
 'daddy's board'

However, the frequency of relatives in terms of the relativized head in the child-directed speech corpus is opposite to the frequency found in the adults' corpora discussed in the previous section. In view of Chen and Shirai's analysis of children's data, it seems that the distributional information extracted from the child-directed speech corpus would serve as a better predictor of the child's acquisition pattern than the distributional information extracted from the adults' corpora. But when we consider previous studies on Chinese-speaking children's acquisition of relatives discussed in Chapter 3, e.g., Lee (1992), Su (2004), or even my experimental results,

it becomes unclear whether the child-directed speech corpus is still the optimal input for acquisition, since none of these studies showed that object relatives outperformed subject relatives. I will return to this issue when comparing my corpus analysis results with my experimental results in Section 5.

In the next section, I lay out the current corpus analysis of Chinese relative clauses.

3 The Current Corpus Study: Chinese Treebank

The set of data used in this study is composed of a comprehensive set of frequency information of different types of Chinese relative clauses. The majority of frequency information was automatically extracted from the Penn Chinese Treebank, which is a large-scale, syntactically annotated corpus (Xue et al. 2005). Apart from the automatic extracted data, this study also contains a small amount of data that was manually searched due to the distributional properties of certain types of relatives that make them difficult to extract using automatic methods.

3.1 The Corpus: The Penn Chinese Treebank and Relative Clauses

In this section, I provide some more details of the Chinese Treebank, and describe how relative clauses are annotated and parsed in the Treebank.

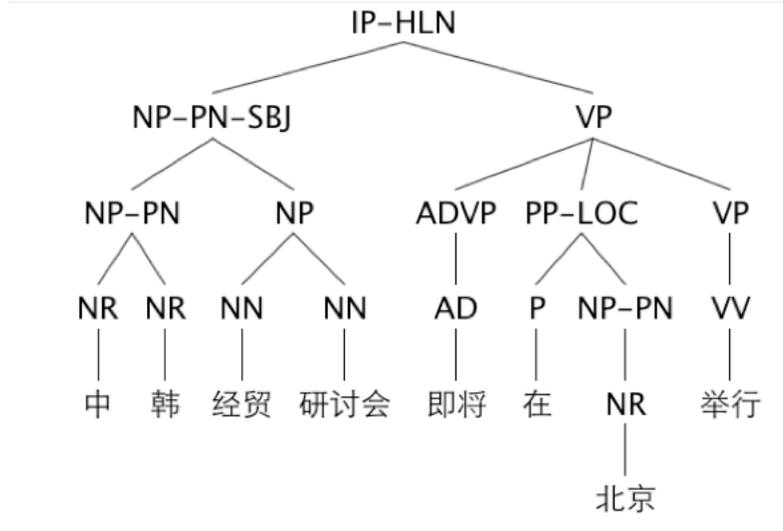
As mentioned earlier, the data in the Penn Chinese Treebank is mainly composed of news and magazine articles from Xinhua newswire, Hong Kong News and Sinorama magazine (now called Taiwan Panorama magazine). The corpus was parsed by the Penn Chinese Treebank Project (Xue et al. 2005). In this study, I mainly

rely on the Chinese Treebank version 8.0 as it was the most up-to-date version when this study was conducted. This version contains 3,007 text files, which include 71,369 sentences and approximately 1.5 million words. The Treebank is annotated with part-of-speech tags and phrase structure trees.

Below is an example that illustrates a sentence with part-of-speech tags on the words and the corresponding phrase structure tree in the Treebank (i.e., a screenshot).²¹

²¹ The part-of-speech tags used in this diagram include: AD for adverb, HLN for headline, NN for common noun, NR for proper noun, VV for other verb. More details can be found in the guidelines of Penn Chinese Treebank.

(5)



zhong han jingmao yantaohui jijiang zai beijing
China Korea economic.trade conference soon in Beijing
juxing
held

‘The China-Korea economic-trade conference will be held in Beijing soon.’

As can be seen in this phrase structure tree, each word has its corresponding part-of-speech tag, i.e., syntactic category, and all heads project their phrases. For example, a noun head such as 研讨会 ‘conference’ will project a noun phrase 经贸研讨会 ‘economic-trade conference’ in the tree. Moreover, noun phrases that occur in the subject or object position are indicated by their corresponding grammatical functions in the sentence. In this case, 中韩经贸研讨会 ‘China-Korea economic and trade conference’ is the subject of the sentence, and therefore it is indicated by SBJ. Certain

phrases are also tagged with their thematic roles, so *在北京* ‘in Beijing’ indicates the location of this entire event; therefore, it has this additional tag LOC.

In addition to these part-of-speech tags that indicate overt elements in sentences, the Treebank contains a particular list of part-of-speech tags that allow us to manipulate covert elements, which include *T*: trace of A’-movement, *NP*: trace of A movement, *PRO*: the null element in control constructions, *pro*: pro-drop, *RNR*: used in right node raising, and *OP*: used for the empty operator in relative constructions.²²

Among these tags on covert elements, both the tag *T*, which indicates a trace of A’-movement, and the empty operator *OP* in relative constructions become very useful in searching for relative clauses in the Treebank, since relative clauses that appear to contain gaps, including argument and adjunct relatives, are analyzed as a kind of A’-movement.²³ Thus, to search for trees where there is a relative clause, we can take advantage of both *OP* and *T*. Moreover, since argument relatives, such as subject and object relatives, are annotated differently from adjunct relatives, including relatives that modify location, manner, and reason to name a few, either just the former or the latter type can be extracted. Specifically, argument relatives are

²² The relative operator in Chinese Treebank is always considered empty, as there are no elements that are like English *wh*-words such as *who*, *what*, *where*, *how* in Chinese relative clauses.

²³ Note that the analysis of relatives in the Treebank is not necessarily the same as in certain theoretical works on Chinese relatives. For example, in Aoun and Li’s (2003) analysis, argument relatives that contain a gap involve raising of the head noun, and no relative operator; argument relatives that contain a resumptive pronoun involve an empty relative operator which binds the pronoun; adjunct relatives that appear to contain a gap are associated with an empty relative operator that originates from the gap position. These discrepancies, however, do not render themselves real hindrances in a search for different types of relatives, as will be made clear in the discussion below.

tagged with NP in conjunction with the trace tag *T*, whereas adjunct relatives are tagged with PP (even though they may not necessarily be PPs) together with the trace tag *T*.

Furthermore, since a tree or a sentence may potentially carry more than one relative clause, to ensure that a relative operator always binds the exact trace, i.e., both c-commanding its trace and being coreferential with each other, we can use the co-index feature that is indicated in the Treebank. Treebank 8.0 uses indices for elements that express relationships such as binding. In the case of relative clauses, therefore, the empty relative operator will be tagged by an index that is the same as the other index in its bound trace. However, such kind of co-indexing is not annotated among (overt) pronouns and their antecedents. This means that, when a relative clause is formed by using a resumptive pronoun instead of a gap, for example, indirect/oblique object relatives in Chinese, one cannot use the same principle that applies to gap relatives. But in Section 4, I will show that this is not entirely true in the current parsed Treebank and I will also indicate how the potential indirect/oblique object relatives, at least the items that I am concerned with, can still be extracted manually.

In the following section, I introduce the tool Tregex and how it can be used to search for desired patterns and constructions.

3.2 The Tool for Querying Tree Data Structures: Tregex

Tregex is a tree query tool that can be applied to syntactically annotated corpora (Levy and Andrew 2006). Specifically, it can be used to search for matching patterns in tree data structures based on tree relationships and matches of regular expressions

on nodes. The tree relationships which Tregex can query include some core structural relations assumed in the generative grammar, such as dominance, precedence, and kinship (e.g., Haegeman 1994; Carnie 2010). I briefly describe each of the relevant structural relations and how Tregex is used to search for them. Then, in Section 3.3, I discuss a very useful feature that is manipulated in particular for extracting relative clauses in this study.

First, Tregex allows for extracting nodes that have dominance relation. Dominance is a relation of containment, and its definition can be found in e.g., Carnie (2010), as in (6).

(6) **Dominance** (Carnie 2010: 29, *informal definition*)

Node A dominates node B if and only if A is higher up in the tree than B and if you can trace a line from A to B going only downwards.

More specifically, if node A dominates node B, A contains B. Take the tree in (5) for example, the node highest node IP-HLN dominates all the other nodes, including the intermediate nodes such as NP-PN-SBJ, NP and ADVP, terminal nodes such as NR, NN, VV, and also the words at the bottom, such as 中 ‘China’, 韩 ‘Korea’ and 研讨会 ‘conference’.

In Tregex, to search for syntactic trees where node A dominates node B, the pattern “A<<B” or “B>>A” is used. For example, to search trees that contain node IP-HLN that dominates 中 ‘China’, we will have the search pattern “IP-HLN<<中”. If more restrictions need be imposed on the dominance relation, such as having no

intervening node between A and B, then the pattern will be modified as “A<B” or “B>A”, which indicates an immediate dominance relation. Thus, if we want to find trees where the node IP-HLN immediately dominates NP-PN-SBJ, then the search pattern is “IP-HLN< NP-PN-SBJ”.

The second kind of tree relation that can be searched for in Tregex is precedence, which concerns the linear order of the constituents. Precedence, according to Carnie (2010), can be defined by appealing to the immediate dominance relation in conjunction with the orderings of elements among sisters. This is called sister-precedence, and is defined in (7).

(7) **Sister Precedence** (Carnie 2010: 40)

Node A sister-precedes node B if and only if both are immediately dominated by the same node M, and A emerges from a branch from M that is to the left of the branch over B.

General precedence can be, therefore, defined on sister precedence, as in (8).

(8) **Precedence** (Carnie 2010: 40): Node A precedes node B if and only if

- (i) Neither A dominates B nor B dominates A *and*
- (ii) Some node E dominating A sister-precedes some node F dominating B (because domination is reflexive, E may equal A and F may equal B, but they need not do so).

A more restricted type of precedence relation is immediate precedence, and is defined in (9).

(9) **Immediate Precedence** (Carnie 2010: 41)

A immediately precedes B if A precedes B and there is no node G that follows A but precedes B.

Again, take the tree in (3) for example. The node NP-PN-SBJ immediately precedes the higher VP, ADVP, AD, and the word 即将 ‘will’. However, it sister-precedes only the higher VP but not the others, as NP-PN-SBJ and the higher VP are both immediately dominated by IP-HLN and that NP-PN-SBJ emerges from a branch from IP-HLN that is to the left of the branch over the higher VP.

Taken together, the search pattern for A to precede B in Tregex is “A..B” or “B,,A”, and the search pattern for A to immediately precede B is “A.B” or “B,A”. Moreover, when two nodes A and B are sisters, then the search pattern for A to be a sister of B (or B to be a sister of A) is “A\$B”. If A is the *i*th child of C, then the search pattern is “C<*i*A”. Thus, if we want to find trees where NP-PN-SBJ is a sister of VP and that it immediately precedes VP, then the search pattern is “NP-PN-SBJ\$.VP”. But if we just want to search trees where NP-PN-SBJ precedes VP, then the search pattern is “NP-PN-SBJ.VP”

In the next section, I describe a particular feature in Tregex that can be used to extract relative clauses in the Treebank for the purpose of this study.

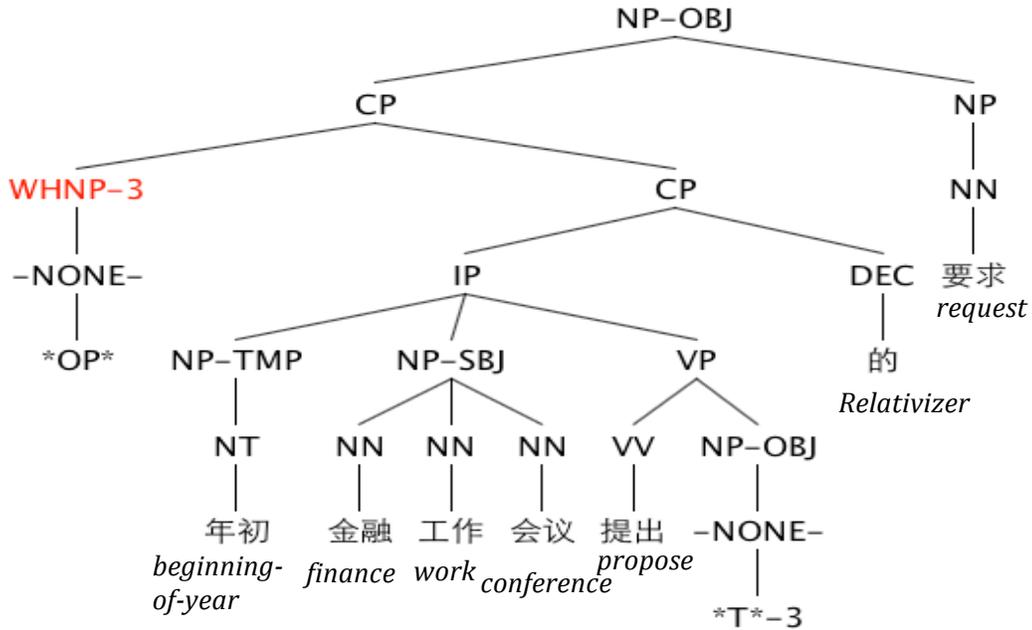
3.3 *Extracting Relative Clauses with Tregex in the Treebank*

The particular feature in Tregex that is very useful for the current study is the use of variable groups. Specifically, when using a regular expression R, i.e., a sequence of characters, to describe a node label, one can assign any group in R to a variable name, e.g., *i*, *j*, and so forth. If there is more than one group matching the same variable name in the search pattern, then matching will only occur when all the groups that are assigned the same variable name capture the same string. This helps to impose coindexation relationships among nodes in a search pattern, such as structures that involve movement and trace. The syntax for the variable group is as shown in (10).

(10) /<regular expression> /#<group-number>%<variable-name>

The following tree that contains a relative clause in the Treebank may be used to illustrate how a relative construction can be extracted using this pattern. Consider (11) (English glosses are added for expository purposes):

(11)



‘the request that the finance work conference proposed at the beginning of the year’

This entire tree corresponds to a noun phrase NP that occurs in the object position of the clause containing it, as indicated by the notation NP-OBJ in the highest node. It means ‘(the) request that the Financial Work Conference proposed at the beginning of this year.’ To simply extract trees that contain such a type of relative clause, where the relative operator *OP* dominated by WHNP-3 binds the trace *T*-3, we can use the search pattern such as (12).

(12) /[^]WH.*-([0-9]+)/#1%i../[^]*T*-([0-9]+)/#1%i

To begin with, let me break down this pattern a little bit more to facilitate understanding. The caret symbol “[^]” in regular expression indicates start of string

(unless it is within brackets). The dot or period “.” matches any character except line break. The asterisk or star “*” means to match the preceding token zero or more times, and “[0-9]+” means any number. Thus, the regular expression “/^WH.*” matches any string that starts with a capitalized WH and is followed by any character(s). Then, when we add “[0-9]+” to this string, the entire regular expression “/^WH.*-[0-9]+” means to match any string starting with a capitalized WH which is followed by or any or no character(s) which is (are) then followed by a hyphen “-” which is then followed by any number.²⁴

Since we want to ensure that the wh-operator binds the trace which is coreferential with it, the index in the wh-operator and the index in the trace need to be the same. That is why we use the same variable name “i” for both the wh-operator and the trace in (10). For the time being, the two dots “..” that separates the wh-operator and the trace just show a precedence relation between them.

But as we want to be more specific about the matches, for instance, structures that contain the overt relativizer DEC *de* that is immediately dominated by a CP that is dominated by any kind of NP, we need to modify the search pattern such as in (13):

(13) All relatives with the relative head binding its variable.

$$\begin{aligned} @NP < (CP < /^WH.*-[0-9]+)/\#1\%i < (CP << (@NP < (-NONE- < /^*T*- \\ ([0-9]+)/\#1\%i)) < DEC)) < NP \end{aligned}$$

²⁴ For more details about regular expressions, the reader is referred to, e.g., Friedl (2006).

This will give us all relative constructions that contain the overt relativizer *de* where the head NP is immediately dominated by an NP in the clause that contains it. Note that this pattern does not yet distinguish argument from adjunct relatives, nor subject from object relatives. As the expression for the lower NP is “@NP”, this means that it matches any basic category NP without adhering to functional tags or other annotations on the label.

In the next section, I discuss how I categorize and extract different types of relative clauses in the Treebank.

3.4 *Types of Relative Clauses*

In this corpus study, I identify two main types of relative clauses. One is the argument relative clause. This type consists of three subtypes: subject gapped relatives, as in (14), (direct) object gapped relatives, as in (15), and indirect/oblique object relatives, as in (2) (repeated here as (16)).

(14) xihuan nanhai de ren
like boy REL person
‘the people that like the boy’

(15) nanhai xihuan de ren
boy like REL person
‘the person that the boy likes’

- (16) a. na-ge Zhangsan song ta shu de xuesheng
 that-CL Zhangsan give him book REL student
 ‘that student who Zhangsan gave books to’ (Lit. ‘that student who
 Zhangsan gave him books’)
- b. na-ge Zhangsan dui ta xiao de xuesheng
 that-CL Zhangsan to him smile REL student
 ‘that student who Zhangsan smiled at’ (Lit. ‘that student who
 Zhangsan smiled at him’)

The other type is the adjunct relative clause, including several sub-types. As mentioned earlier in Section 3.1, the parsing in Treebank has different notations in the trace for argument and adjunct relatives. As a result, to extract an adjunct relative, one can add the additional notation to the search pattern in (13) for each type of adjunct relative. For instance, to search for a relative that relativizes on a location, the search pattern will be as in (17):

- (17) Locative Relatives
- @NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-
 LOC < (-NONE- < /^*T*-[0-9]+)/#1%i)) < DEC)) < NP

Since the focus of this study is on argument relatives, I leave the search patterns for adjunct relatives in Appendix E.

Similar to adjunct relatives, each type of argument relative also carries an additional notation in the Treebank. In particular, to search for subject relatives, I modified the NP that contains the trace by adding SBJ to it, namely, “NP-SBJ”. For object relatives, this NP was modified to “NP-OBJ”. For indirect/oblique object relatives, however, since they use overt pronouns instead of traces, the Treebank does not provide us with the coindexation between the empty relative operator and the overt pronoun; thus it seems challenging to extract them at first sight. Despite this, I still applied the same principle to modify the same NP by adding the IO (i.e., indirect object) to it, and interestingly, there were some results. I will discuss these results in more detail in the following section. Table 10 lists the search pattern for each type of argument relative.

Table 10: Tregex Search Patterns for Subject, Object and Indirect Object RCs

Type of RC	Search Pattern in Tregex
Subject RC	@NP < (CP < / [^] WH.*-([0-9]+)/#1%i < (CP << (/ [^] NP- SBJ / < (-NONE- < / [^] *T*-([0-9]+)/#1%i)) < DEC)) < NP
Object RC	@NP < (CP < / [^] WH.*-([0-9]+)/#1%i < (CP << (/ [^] NP- OBJ / < (-NONE- < / [^] *T*-([0-9]+)/#1%i)) < DEC)) < NP
Indirect/Oblique object RC	@NP < (CP < / [^] WH.*-([0-9]+)/#1%i < (-NONE- < *OP*) < (CP << (NP- IO < (-NONE- < / [^] *T*-([0-9]+)/#1%i)) < DEC)) < NP

4 Results

Overall, the search pattern in (13) allows me to extract all types of relatives, and the results show that there are 24,837 relative clauses in the Chinese Treebank. Among these, subject relatives are the most frequent, accounting for 71.2% of the total.

Object relatives account for 21.5% of the total relatives, which are the second most frequent type. For the indirect object relatives, there are only two examples found in the corpus. Table 11 is the summary of raw counts and frequencies of these relatives. Table 12 includes an example for each type of relative found in the corpus.

Table 11: Raw Counts and Distributional Frequency of 3 Types of RCs

Structure	Counts	Frequency
Subject RC	17,695	71.2%
Object RC	5,330	21.5%
Indirect object RC	2	N/A (or 0)
Total	23027	92.71%

Table 12: Examples of RCs from the Chinese Treebank

Structure	Example
Subject RC	已建立的乡镇企业集团 The rural enterprise groups that have set up.
Object RC	金融工作会议提出的要求 The requirements that the financial work conference proposed.
Indirect object RC	他们称之为‘马拉松谈判’的第一天会谈 The person that the royal family awarded the degree to.

Moreover, adjunct relatives all together account for 6.9% of the total.²⁵ See Appendix F for the summary of raw counts and frequency for different types of adjunct relative clauses.

As indicated earlier, there is no coindexation guide in the Treebank that allows me to search for trees where an empty relative operator binds an overt pronoun in the relative clause. In consequence, as indirect/oblique object relatives in Chinese

²⁵ There are 12 cases of relative clauses that belong to neither of the above categories.

are formed by an overt pronoun, we would expect there be no results even if we modify the search pattern by adding the indirect object notation to it as we did in subject and object relatives. Unexpectedly, not only were two examples found, but these two examples were analyzed as relatives that contain a trace in the Treebank:

- (18) huangjia shou-yu t_i xuwei de ren_i
 royal.family give-give degree REL person
 ‘the person that the royal family awarded the degree to’
(chtb_2710.bn-7)

- (19) tamen cheng-zhi-wei t_i ‘malasong’ tanpan de diyi
 they call-it-as marathon negotiation REL first
 tian huitan_i
 day talk
 ‘the first-day talk that they refer to as a “marathon negotiation”’
(chtb_3119.bn-3)

On closer inspection, though, there are some interesting facts that are worth discussing about the example in (19).²⁶ Although the parse tree in (19) is associated

²⁶ Even though the complex verb *shou-yu* ‘give-give’ in (16) is a ditransitive that appears to be associated with a gap in its indirect object position, it is noteworthy that a resumptive pronoun can actually substitute for the gap, as in (i):

with a trace rather than a resumptive pronoun, there is in fact a word (or character) within it whose meaning is a pronoun, that is, *zhi* ‘it’. This pronoun, perhaps due to its archaic use, is rarely used in Modern Chinese. But since *zhi* ‘it’ in (19) is coreferential with the head noun *diyi tian huitan* ‘first-day talk,’ I would like to reanalyze this relative clause as one associated with a resumptive pronoun instead of a gap, as in (20). Note that this resumptive pronoun cannot be omitted.

- (20) tamen cheng *(zhi_i) wei “malasong” tanpan de diyi
 they call it as marathon negotiation REL first
 tian huitan_i
 day talk
 ‘the first-day talk that they refer to *(it) as a “marathon negotiation”’

Furthermore, in view of the fact that the search pattern for indirect/oblique object relatives cannot possibly identify all desired ones, as there is no coindexation

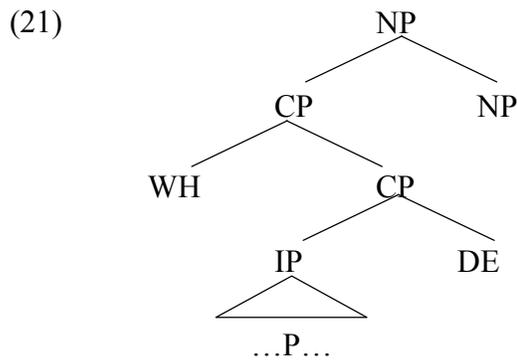
-
- (i) huangjia shou-yu ta_i xuewei de ren_i
 royal.family give-give he degree REL person
 Lit. ‘the person that the royal family award the degree to him’
 Intended: ‘the person that the royal family award the degree to’

In other words, it is a relative construction where a resumptive pronoun is optional. Given our general discussion on the obligatory and optional resumptives in Chapter 2, it is likely that this complex verb *shou-yu* ‘give-give’ in fact takes a small clause as complement, and that the originally assumed indirect object position is the subject of the small clause, as reanalyzed in (ii):

- (ii) huangjia shou-yu[ta_i xuewei] de ren_i
 royal.family give-give he degree REL person

Whether this is true or not will not be investigated here, but this suggests that there may be some misanalyses in the Treebank in some relatives. For the time being, I will simply consider (16) to be a relative that involves a gap.

between the relative operator and the resumptive pronoun, I used two strategies to try to exhaust the matches that may potentially be of particular interest to the current study. First, based on the idea that a less specific search pattern may lead to more results than the original one for indirect/oblique object relative clauses, I sketched a more general, abstract structure for this type, as in (21)



Then, I took this as the base for the search pattern in the Treebank. All things being equal, this tree still indicates that there is a wh-operator in the relative clause, that there is a preposition, which potentially may take a pronoun as its complement, that there is a relativizer DE and that there is a head NP that follows the entire CP. But it is more general, since it removes the indices and specific annotations on the nodes. The search pattern is sketched as in (22).

(22) @NP < (CP < @WH < (CP << (P..DEC)) < NP)

This yields no result in the Treebank. But because this may be due to the fact that there is no empty operator tagged in the tree when a resumptive pronoun is involved,

I removed the “@WH* from (21) and searched again. This time there were 21 matches in total, but none of them was a resumptive relative clause.

The second strategy was used to search for items that were more directly related to my experimental items. I looked for all prepositional phrases whose head preposition is *dui* ‘to’, as the experimental items used in my oblique object relatives are always associated with this preposition followed by a resumptive pronoun. This search was to guarantee that there were no residual matches of *dui* ‘to’ prepositional phrases that might be associated with a relative clause that the first strategy missed or were misanalyzed in the Treebank. Thus, all the sequences of *dui* immediately followed by a pronoun, including *dui ta* ‘to him/her/it’ and *dui tamen* ‘to them,’ were examined. Again, the results showed that none of them involves a relative clause, as summarized in Table 13:

Table 13: Results of the Sequence “*Dui* ‘to’ + Pronoun”

Items Extracted	Number of Occurrences	Correspond to RCs?
<i>dui ta</i> ‘to it’	22	No
<i>dui ta</i> ‘to him’	86	No
<i>dui ta</i> ‘to her’	18	No
<i>dui tamen</i> ‘to them’ (NEU.)	4	No
<i>dui tamen</i> ‘to them’ (MAS.)	74	No
<i>dui tamen</i> ‘to them’ (FEM.)	1	No

In sum, the corpus analysis in Chinese Treebank 8.0 shows that subject gapped relatives are the most frequent, object gapped relatives the second, and indirect/oblique object resumptive relatives are by far the least frequent, which are summarized in (23):

(23) **Relative Frequency of Relative Clause Pattern in the Chinese Treebank**

Subject Gapped relatives > Object Gapped relatives > Indirect/Oblique Object
Resumptive relatives

5 Discussion

In this section, I compare the results in the previous corpus analyses with my earlier experimental results and discuss whether the two non-UG approaches to language acquisition find support based on these data. In Section 5.1, I examine the data within frequency-based approaches to language acquisition. In Section 5.2, I discuss the potential and limits of Bayesian approaches to account for the current data.

5.1 Data Within Frequency-Based Approaches

I begin with an overview of frequency-based approaches to language acquisition. Then, I will focus on comparing Treebank corpus analyses, including those in the previous and the current studies, and my experimental results. Following this, I will compare the child-directed speech, both in children and caregivers, with my results.

According to the frequency-based approaches, the frequency information of the input plays a significant role in language acquisition, so the higher frequency of occurrence a linguistic pattern has, the better it will be learned (e.g., Goldberg et al. 2004: Construction grammar; Matthews et al. 2005; Kidd et al. 2006, 2010; Kidd et al. 2007: usage-based framework; Huttenlocher et al. 2002; Theakston et al. 2004). In terms of the acquisition of syntax, the elements whose frequency information matters can be specific lexical items that are associated with the relevant structure, or they can be merely the structures that are to be acquired. To learn an abstract syntactic

pattern, such as word order or argument structure, the child may attend to the frequent and distributionally regular verbs. To learn complex structures such as relative clauses, the child may look at the type that is more frequent than the others.

First, the frequency pattern extracted from the Chinese Treebank would predict that, among these three types of argument relatives in Chinese, children should have the best performance on subject gapped relatives, and worst on the indirect/oblique object relatives. However, Chinese-speaking children's comprehension of these three types in my experimental study indicates a different story. The results of that study showed that children performed equally well on both subject gapped and oblique object resumptive relatives, and worst on object gapped relatives. On the one hand, the fact that Chinese-speaking children performed better on subject than object gapped relatives seems to support the frequency-based approaches. On the other hand, the fact that children performed equally well on subject and oblique object resumptive relatives and that they performed better on oblique object resumptive relatives than object gapped relatives alike seem to undermine frequency-based explanations. The control group's data (i.e., the adults' data) in my experiment, though, intriguingly, matches the frequency pattern extracted from the adults' corpora. Perhaps this suggests that adults' corpora do not serve as a good baseline for comparison with children's data; namely, adults' corpora are not the right type of input from which children would work on their grammar. Despite this potential drawback, the results from comparison between the current Treebank analysis and my experimental suggest that the frequency-based approaches are not well supported.

Second, if one examines alternative factors such as the frequency of grammatical functions of head nouns, the subject function is the most frequent, the object function the second (Hsiao & Gibson 2003; Wu 2009; Hsiao & MacDonald 2013; Jäger et al. 2015), object-of-preposition the third, and predicate nominal the last (Wu 2009). With regard to the grammatical function, my experimental items of three types of relative clauses all modify the subject of the matrix clause, as in (24-26).

(24) Subject Gapped relative

[___ qin-le maomi de] xiao-gou pao-zou le.
 kiss-ASP cat REL little-dog run-away ASP
 ‘The dog that kissed the cat ran away.’

(25) Object Gapped relative

[xiao-gou qin-le ___ de] maomi pao-zou le
 little-dog kiss-ASP REL cat run-away ASP
 ‘The cat that the dog kissed ran away.’

(26) Oblique Object Resumptive relative

[qingwa dui ta tiaowu de] yazi pao-zou le.
 frog to it dance REL duck run-away ASP
 ‘The duck that the frog danced for ran away.’

This piece of frequency information suggests that the status of the relatives in

examples in (24-26) should be the same, and in fact they all modify the most prominent grammatical function. Therefore, the frequency of grammatical functions may not play a role in the children's various performances on the relatives; however, it may boost the entire performance since subject-modifying relatives are the most frequent type.

Another frequency factor that may affect children's performance on relatives is the animacy of head nouns, since previous studies also showed that object relatives tend to occur with inanimate heads, whereas subject relatives (i.e., with transitive verbs) tend to occur with animate nouns (Wu 2009; Hsiao & MacDonald 2013). In my study, the head nouns of the three types of relatives are all animate. For this reason, this piece of frequency information about animacy may suggest why children performed worse on object gapped relatives than subject gapped relatives, since object gapped relatives with animate noun heads have lower frequency than subject gapped relatives with animate noun heads. However, if one considers the children's performance on oblique object resumptive relatives, then it becomes puzzling why the performance is as good as on subject gapped relatives. Recall that I only found two cases of indirect/oblique object relatives, and only one of them is a real resumptive relative, which is to say that this type always has the frequency that is approximately equal to zero. Therefore, the frequency-based approaches are faced with a dilemma as to which part of frequency information they should use to explain the data.²⁷

After looking at the news writing corpus, now we may turn to the frequency

²⁷ Nonetheless, since I did not manipulate the animacy factor in my experimental study, nor did I extract the frequency of animacy on different types of relatives, I will leave the issue of influence of animacy on acquisition of relatives for future research.

analysis from the child-directed speech corpus (Chen & Shirai 2015), especially the data summarized in Table 9. First, consider the relativized head. From their study, the frequency pattern for both the children and caregivers is: Object relatives > Subject relatives > Oblique relatives/Gapless relatives. This immediately leads to two clear discrepancies between the experimental results and the child-directed speech corpus. Based on the frequency pattern between object relatives and subject relatives in the child-directed speech corpus, children are supposed to perform better on object relatives than subject relatives, which is contrary to the experimental results.

Moreover, since object relatives have the highest frequency among all relatives, this should predict that at least children's performance on object relatives is better than indirect/oblique object relatives, which is, again, contrary to my experimental results. Thus, the frequency-approaches would have some difficulties explaining the mismatch between the input in the corpus and the experimental data. Nonetheless, one possible solution is to consider the big differences in production and judgment data (LouAnn Gerken, p.c.). Specifically, children's frequency data in the child-directed speech corpus is not just frequency data; rather, it is behavioral data about children's abilities in their language, which does not match my experimental data. Thus, it is not surprising to see that the frequency pattern from the child-directed speech corpus contradicts the pattern of my comprehension data, since the former type of data may reflect a different kind of ability than that which is obtained from the comprehension study.

Second, the frequency information about grammatical roles in the child-directed speech corpus will tell a similar story as in the Treebank corpus. Regardless

of the isolated NPs, the subject role in both children and caregivers is still the most frequent, so again, this plays no role in affecting different types of relatives that are controlled in my experiment.

Alternatively, instead of using the entire relative clause as a frequency pattern, one can adopt the frequency-approaches that utilize frequencies of individual lexical items to learn more abstract or higher-level syntactic knowledge. Specifically, I extracted the frequency for all verbs used in my study from the Sinica Corpus (e.g., Chen et al. 2003).²⁸ This is to test whether the frequency distribution of individual verbs can predict the performance on their corresponding relative clauses. Specifically, transitive and intransitive verbs both occur with subject relatives. Transitive verbs co-occur with object relatives, and intransitive verbs co-occur with indirect object relatives. The verbs in Huang’s tested relatives consist of two transitive verbs 親 (*qin*) ‘kiss’, 拍 (*pai*) ‘pat’, and two intransitive verbs 跳舞 (*tiaowu*) ‘dance’ and 揮手 (*hui-shou*) ‘wave at’. Their frequencies are summarized in Table 14:

Table 14: Raw Counts and Frequency of Experimental Verbs in Sinica Corpus

	Verb	Frequency	Raw Counts	Average Frequency
Transitive	親 ‘kiss’	0.000	13	0.0035
	拍 ‘pat’	0.007	333	
Intransitive	跳舞 ‘dance’	0.002	108	0.001
	揮手 ‘wave at’	0.000	19	

²⁸ The online word frequency list in Sinica Corpus can be found here: http://elearning.ling.sinica.edu.tw/eng_teaching.html.

As shown in Table 14, the average frequency of experimental transitive verbs is still higher than intransitive verbs. As direct objects are associated with transitive verbs and indirect/oblique objects with intransitive verbs, the higher frequency of transitive verbs should predict that children's performance on object relatives is better than indirect/oblique object relatives, which is contrary to the result.

Thus, although the frequency pattern from the adults' corpora can explain Chinese-speaking children's better performance on subject gapped relatives than object gapped relatives, it cannot explain children's overall performance on oblique object resumptive relatives. As for the frequency pattern extracted from the child-directed speech corpus, it completely contradicts my experimental data. But if this problem lies in the common discrepancies between production and comprehension in language acquisition (e.g., Hendriks & Koster 2010), then, instead of treating this mismatch as a weakness for frequency-based approaches, I would suggest that future studies compare data of the same type (for example, comprehension versus production, adults versus children, etc.) in order to find frequency results that are more likely to inform the experimental data.

5.2 Data Within Bayesian Approaches

In Section 5.2.1, I discuss how a Bayesian approach to language acquisition can potentially provide an account of the current data and point out the need to delineate the hypothesis space, or, grammar, for relative clauses in Chinese that can fit in the Bayesian approach. In Section 5.2.2, I then investigate two sets of plausible hypothesis space for relative clauses. In 5.2.3, I discuss the remaining issues in this framework with regard to language acquisition.

5.2.1 Bayesian Approaches to Language Acquisition and Results

The essence of Bayesian approaches to language learning is that a learner can make generalizations about the most plausible hypothesis from just one or a few positive examples of a linguistic expression in the input (e.g., Xu & Tenenbaum 2007; Kemp et al. 2007; Perfors et al. 2010; Perfors et al. 2006, 2011; Gerken et al. 2015). For example, Gerken et al. (2015) showed that 9-month-old infants can generalize to new input types from multiple tokens of just a single input example of the syllable string. Xu and Tenenbaum (2007) demonstrated that both adults and 3- and 4-year-olds can make inferences about plausible word meanings based on just one or a few examples to which they were exposed. Furthermore, Kemp et al. (2007) showed that a hierarchical Bayesian model can learn ungrammatical verb argument constructions even if those are not specified as being ungrammatical in the input. Perfors et al. (2006, 2011) also demonstrated that a hierarchical Bayesian model can correctly learn that hierarchical phrase structures, rather than linear structures, are the scaffold of syntax.

One interesting characteristic of the Bayesian approach is that the same set of data cannot be explained by more than one hypothesis in the same hypothesis space, according to Bayes' Theorem. Thus, for example, given a set of syntactic constructions (SC) as the data, say, $\{SC_1, SC_2, SC_3\}$, and a set of syntactic rules, or, potential grammars (G) in the same hypothesis space, say, $\{G_1, G_2, G_3, G_4\}$, then only one of the grammars will exactly account for the data. Assume that G_1 accounts for $\{SC_1\}$, G_2 accounts for $\{SC_2, SC_3\}$, G_3 accounts for $\{SC_1, SC_2\}$, and G_4 accounts for $\{SC_1, SC_2, SC_3\}$, then the winner is G_4 .

With regard to Bayesian models' ability to make plausible inferences from scarce data, my experimental results appear provide a potential piece of evidence to support this framework given the Treebank corpus analysis. That is, with only two examples of indirect/oblique object relatives in the input, children's performance was still as robust as on subject relatives. This suggests that, regardless of the comparative frequency among relatives, children can learn such a type that has so few occurrences. In other words, children seem to be able to make inferences about the most plausible hypothesis, e.g., correct syntactic structure of indirect/oblique object relatives, given only two examples.

However, exactly how a Bayesian model works in this case merits caution, especially in terms of the hypothesis space. Recall that the syntactic structure of indirect/oblique object relatives in Chinese differs from subject or object relatives not simply in the relativized position. An additional, and crucial feature that distinguishes subject and object relatives from indirect/oblique object relatives is that the latter use a resumptive pronoun, instead of a gap, in the relativized position, as shown by the contrast between (24-25) and (26). In other words, the hypothesis space for the grammars of relative clauses in Chinese cannot simply be the one that concerns whether there exists a relativizer or a relative feature in the clause. Rather, it needs to be properly defined. In the next section, I discuss two probable kinds of hypothesis space for the relative clause formation in Chinese.

5.2.2 Hypothesis Space for Relative Clauses Within Bayesian Approaches

The first kind of hypothesis space can be perceived as one that consists of two

strategies in the Chinese relative clause formation: a gap strategy and a resumptive strategy. Then, given the data of relative clauses based on the relativized position, two sets of data can be generated, as in (27):

- (27) a. *Set 1: Relatives that use a gap*
{Subject relatives, object relatives}
- b. *Set 2: Relatives that use a resumptive pronoun*
{Indirect/Oblique object relatives}

If given data in Set 1, the learner will learn the gap strategy. On the other hand, if given data in Set 2, the learner will learn the resumptive strategy. But if both sets of data are provided, then the learner will, supposedly, learn both strategies. The question is, given such kind of hypothesis space, will the learner be confused about when to use a gap and when to use a resumptive pronoun? I will, however, not explore this question in this study but leave it for future research.

The second kind of hypothesis space concerns different grammars that already specify when to use a gap and when to use a resumptive pronoun (p.c. Heidi Harley). Thus, if a learner considers three types of relatives in terms of the relativized positions, such as subject, object and indirect/oblique object relatives, then potentially there are eight grammars competing with each other in the same hypothesis space, as illustrated in Table 15.

Table 15: Eight Potential Grammars in the Same Hypothesis Space for RCs

Grammar	Subject relatives	Object relatives	Indirect object relatives
G ₁	GAP	GAP	GAP
G₂	GAP	GAP	PRONOUN
G ₃	GAP	PRONOUN	GAP
G ₄	GAP	PRONOUN	PRONOUN
G ₅	PRONOUN	GAP	GAP
G ₆	PRONOUN	GAP	PRONOUN
G ₇	PRONOUN	PRONOUN	GAP
G ₈	PRONOUN	PRONOUN	PRONOUN

For example, G₁ will account for exactly the set of data where all three types of relative clauses use the gap, but not others. In this case, given the Chinese data, G₂ will be the winner, as it accounts for exactly the data where subject and object relatives use a gap, and that indirect/oblique object relatives use a resumptive pronoun.

The second kind of hypothesis space seems more plausible than the first kind, as it can explain where to use which strategy. However, one potential problem for this hypothesis comes from the data where a gap and a pronoun are interchangeable in the same relativized position. But again, I leave this issue for future research.

5.2.3 Remaining Issues

Even though the Bayesian approach seems to be able to explain why Chinese-speaking children can learn the indirect/oblique object relatives given scarce data, it has not explained why children's performance varies on different types of relatives. In particular, although the Bayesian view could be good at explaining why the subject and object relatives are parallel, it CANNOT explain why the object relatives are not learned as well; if learners can learn from one or two examples of oblique object

relatives, they should also be able to learn from hundreds of object relative examples.

Moreover, so far Bayesian models have been dealing with the question of whether a learner can learn based on sparse data, and the results have been quite promising. But they have not handled sentences that can potentially share the same set of rules or principles, such as wh-questions, relative clauses, raising and passives, to name a few (e.g., Chomsky 1981) or the ambiguous sentences that children can learn (e.g., Omaki et al. 2014). This makes it difficult to provide a sound explanation as to how subtle distinctions within each type of clause can lead to varied performances in the learners. Thus, it is not surprising to me that such kind of question has been mostly handled either by UG-based generative grammar (e.g., McDaniel & Cowart 1999; Han et al. 2007) or by processing accounts (e.g., Arnon 2005; Hsu et al. 2009; Rahmany et al. 2014, among others).

In brief, the poorer performance on object relatives is a serious problem for the Bayesian approach.

6 Conclusion

In this study, I conducted a corpus analysis of relative clauses in the Chinese Treebank 8.0. Specifically, I extracted the frequency information of subject relatives, object relatives, indirect/oblique object relatives and all other adjunct relative clauses. The results were compared with my experimental results as well as corpus analyses from previous studies, and then discussed within both the frequency-based and Bayesian approaches to language acquisition. It was shown that none of the frequency patterns from the corpus studies match the experimental results exactly, so the frequency-based approaches are unable to gain full support from these overall results.

On the other hand, Bayesian approaches seem to find some support from these results, because Chinese-speaking children can learn indirect/oblique relatives well even with very scarce available data in the input. Despite this, Bayesian approaches are still facing the challenge of explaining why Chinese-speaking children's performance on certain types of relatives, such as on subject relatives and indirect/oblique object relatives was better than on direct object relatives.

CHAPTER 6: THE ANALYSIS OF RESUMPTIVE PRONOUN AND GAP STRUCTURES IN CHINESE RELATIVE CLAUSES

1 Introduction

After seeing that both the purely UG-derived proposal and the frequency-based proposal failed to fully account for Chinese-speaking children's comprehension pattern of gapped and resumptive relatives, I will reconsider some language-specific properties that may have come into play in acquisition but have been neglected at the beginning of the discussion. In particular, since the main puzzle is why children had a better performance on Oblique Object Resumptive RCs that not only are lower on the Accessibility Hierarchy but have a lower frequency than Subject Gapped and Object Gapped RCs in general, I will focus on the comparison of syntactic structures between the gap and the resumption strategies in Chinese relatives.

Thus, this chapter serves to provide a comprehensive examination of Chinese speakers' use of gaps and resumptive pronouns in argument relative clauses in order to prepare the way for discussion of influence from language-specific properties on acquisition in Chapter 7. As discussed in Chapter 2, gaps in Chinese relative clauses are obligatory in matrix subject and matrix object positions, interchangeable with a resumptive pronoun in possessor, embedded subject and embedded object positions, and are excluded from the oblique (or indirect) object positions, as summarized in Table 16.

Table 16: Gaps in Chinese Argument RCs

Matrix Clause				Embedded Clause	
Subject	Object	Oblique Object	Possessor	Subject	Object
+	+	–	(+)	(+)	(+)

Key: The symbol ‘–’ indicates that a gap is prohibited from occurring in that position; ‘(+)’ indicates that the occurrence of a gap is optional; ‘+’ indicates that the occurrence of a gap is obligatory.

On the other hand, the occurrence of resumptive pronouns is the opposite of gaps in argument relative clauses. Resumptive pronouns are excluded from matrix subject and matrix object positions, obligatory in oblique object positions, and optional in possessor, embedded subject and embedded object positions, as shown in Table 17.

Table 17: Resumptive Pronouns in Chinese Argument RCs

Matrix Clause				Embedded Clause	
Subject	Object	Oblique Object	Possessor	Subject	Object
–	–	+	(+)	(+)	(+)

Key: The symbol ‘–’ indicates that a resumptive pronoun is prohibited from occurring in that position; ‘(+)’ indicates that the occurrence of a resumptive pronoun is optional; ‘+’ indicates that the occurrence of a resumptive pronoun is obligatory.

While previous studies have shown that gaps in argument relative clauses generally exhibit island effects (e.g., Ning 1993; Aoun & Li 2003; Huang et al. 2009), as in the Adjunct Island Condition (1), or in the Complex NP Condition (2), resumption does not appear to be sensitive to island constraints, as in (3) (the italic symbol *e* indicates an empty category without presuming which type of NP may be present there).

(1) *ruguo e_i lai wo jiu lai de na-ge ren_i
 if come I then come REL that-CL person
 ‘the man_i that I will come if e_i comes...’

(Ning 1993: 69)

(2) *wo xihuan e_i chuan e_j de yifu_j de na-ge ren_i
 I like wear REL clothes REL that-CL person
 ‘the person_i that I like the clothes_j that e_i wears e_j ’

(Aoun & Li 2003: 178)

(3) a. wo xiang kan na-ge ni yinwei ta_i bu
 I want see that-CL you because he not
 hui lai hen shengqi de xuesheng_i
 will come very angry REL student
 ‘I want to see the student_i that you are angry because [he_i] would not come.’

b. wo xiang kan na-ge ni yaoqing e_i
 I want see that-CL you invite
 dai ta_j lai de ren_i lai zher de xuesheng_j
 bring him come REL person come here REL student
 ‘I want to see the student_j that you invited the person_i over that brought [him_j] over.’

(Aoun & Li 2003: 170)

Because of the contrast between gaps and resumption, Aoun and Li (2003) propose a movement analysis for argument relatives using gaps and a base-generation analysis for argument relatives using resumptive pronouns.²⁹

²⁹ Aoun and Li (2003) distinguish argument from adjunct relativization in Chinese, whether it is associated with a gap or a resumptive noun phrase. An adjunct relative is defined as relativization of a PP or an adverb but not an NP category. In addition, it is the argument relatives, not adjunct relatives in Aoun and Li's term that belong to the types in Keenan and Comrie's (1977) Accessibility Hierarchy. Hence, even though an indirect object, or, an oblique object in Chinese is introduced by a preposition, as long as it denotes a core argument of the predicate, as *the boy* in *John talked to the boy*, rather than serves as an adverbial function like *the garage* in *John's dog slept in the garage*, relativization of the former type is still considered an argument relative clause.

In this study, I also adopt Aoun and Li's distinction between argument and adjunct relative clauses. Two pieces of empirical evidence that distinguish argument relatives from adjunct relatives are discussed as follows. Consider a list of adjunct relatives, including relativization on location (ia), time (ib), manner (ic), and instrument (id) (data from Ning 1993: 95):

- (i) a. ta xiu che de cheku
 he fix car REL garage
 'the garage where he fixed his car'
- b. ta xiu che de na-ge wanshang
 he fix car REL that-CL evening
 'the evening when he fixed his car'
- c. ta xiu che de fangfa
 he fix car REL way
 'the way he fixed the car'
- d. ta xiu che de qianzi
 he fix car REL plier
 'the pliers with which he fixed the car'

Note that each example allows the relativized NP/PP to be empty in the original position of the relative clause, which differs from argument relative clauses that relativize indirect or oblique arguments, since the latter require both a preposition and a resumptive pronoun to be present in the original position.

For a relative clause that relativizes an instrument, like (id), a VP that consists of a light verb *yong* 'use' and an optional resumptive pronoun can be inserted back into the original position of the relative clause, as in (ii).

- (ii) ta **yong** (**ta**) **lai** xiu che de qianzi
 he use it to fix car REL plier
 'the pliers with which he fixed the car'

As illustrated above, the resumptive pronoun that refers to the instrument is not obligatory, as opposed to the oblique argument relative clauses.

However, some data suggests that relatives using gaps do not always involve movement. For example, gaps do not always exhibit island effects, as in the complex NP island (4) (contradicting the data in (2) above), in the adjunct island (5) or in the wh-island (6).

- (4) e_i chuan e_j de yifu_j hen piaoli[a]ng de na-ge
 wear REL clothes very pretty REL that-CL
 ren_i
 person
 ‘the person_i that the clothes_j e_i wears e_j are pretty’
 (Aoun & Li 2003: 178)

- (5) na-ge ruguo e_i da-le Zhangsan Lisi
 that-CL if hit-ASP Zhangsan Lisi
 hui hen shengqi de nüsheng_i pao-zou-le.
 will very angry REL girl run-away-SFP
 ‘The girl_i that if e_i hit Zhangsan Lisi would be angry ran away.’

Adjunct relatives are analyzed by Aoun and Li as involving a base-generated head, since they show no reconstruction of the head, and as involving operator movement to the specifier of the relative clause, since they exhibit sensitivity to island conditions, as in (iii).

- (iii) [CP Op_i [IP ... [PP t_i] ...]] [Head NP]
 (Aoun & Li 2003: 175)

- (6) na-ge Zhangsan xiang zhidao e_i weisheme xie-le
 that-CL Zhangsan want know why write-ASP
 zhe-ben shu de zuojia_i likai taiwan le.
 this-CL book REL writer leave Taiwan SFP

‘The writer_i who Zhangsan wonders why e_i wrote this book left Taiwan.’

Most importantly, relatives using gaps are allowed to occur in a coordinate structure with relatives using resumptive pronouns, as in (7a-c).

- (7) a. na-ge [e_i xihuan Zhangsan] de erqie [Lisi
 that-CL like Zhangsan REL and Lisi
 gen ta_i shuo-guo hua] de xuesheng_i
 with she talk-ASP word REL student

‘the student_i that e_i likes Zhangsan and that Lisi talked with her_i before’

- b. na-ge [Zhangsan xihuan e_j] de erqie [Lisi
 that-CL Zhangsan like REL and Lisi
 gen ta_j shuo-guo hua] de xuesheng_j
 with she talk-ASP word REL student

‘the student_j that Zhangsan likes e_j and that Lisi talked with her_j before’

Due to these mismatched properties of gapped and resumptive relative clauses, it will be useful to undertake a thorough re-examination of both types of relative clauses. The organization of this chapter is as follows. In Section 2, I investigate gaps in argument relative clauses in terms of their sensitivity to different island constraints, and to the Weak Crossover Condition. In Section 3, I examine resumptive pronouns in terms of the above conditions, and also provide examples where resumption exhibits reconstruction properties in both island and non-island contexts. In Section 4, I provide more coordinate structure examples of gaps and resumption, and examples where gaps exhibit reconstruction properties within islands. In Section 5, I lay out a revised analysis of both types of argument relative clauses, and argue that while resumption is still derived via base-generation as proposed in Aoun and Li (2003), gaps may be ambiguous between base-generation and movement in general. Section 6 concludes this chapter.

2 Gaps in Relative Clauses

The structural analysis of a relative clause with a gap in some languages is taken to involve movement, as proposed in the head-internal analysis (Chomsky 1977), the head-external analysis (e.g., Schachter 1973; Vergnaud 1974; Kayne 1994), or the hybrid analysis (Sauerland 1998, 2000). The crucial piece of evidence that supports the movement analysis comes from the fact that relative clauses with a gap obey island constraints, which have been taken as a diagnostic of *wh*-movement (Ross 1967).

However, there are other languages where relative clauses with a gap can violate island constraints, such as Palauan (Georgopoulos 1985), Tuki (Bilola 1990),

and Hausa (Tuller 1986). An immediate question that is raised asks about the nature of these gaps in the A'-dependency structure, since they are still bound by A'-binders, namely, having antecedents that are in an A'-position. Furthermore, since this type of language allows both relative clauses with a gap and relative clauses with an overt resumptive pronoun, this raises another question as to whether the gap is simply the null counterpart of the (overt) resumptive pronoun. Interestingly, as these languages are pro-drop, allowing the null pronominal *pro* to occur in certain argument positions of finite clauses, the connection has been drawn between the gap and *pro*. For example, Palauan allows a *pro* not only in subjects, as in (10a), but also in direct objects, as in (10b), and possessors when person and number inflections are present (adapted from Georgopoulos 1985: 61):

- (10) a. ng-remurt (*ngii).
 3S-run he
 ‘[He] is running’
- b. ke-ʔillebed-ii (*ngii).
 2S-PFV.hit-3S him
 ‘You hit him.’

When a gap is associated with a wh-question, a relative clause or a topicalization in these languages, it ignores movement-related island constraints, which is no different from when the overt resumptive pronoun occurs in these structures. As illustrated in

(11), which involves a complex NP island, and in (12), which contains a wh-island, the gap in topicalization does not exhibit island effects at all.

- (11) a buk_i [a ku- dengel -ii [a redil [el
 book IR-1S PFV-know 3S woman COMP
 uldurukl -ii e_i [el mo er a del -ak]]]]
 R-PFV-send 3s COMP go P mother 3S

‘The book_i, I know the woman who sent e_i to my mother.’

(Georgopoulos 1985: 73, ex. 14c)

- (12) a del -ak_i [a diak ku- dengei [el kmo ng-
 mother -1S NEG IR-1S IM-know COMP CL
 ngera_j [a bo lo- ruul e_j e_i el mo belsoil]]
 what IR-FUT IR do L go dinner

‘My mother_i, I don’t know what e_i will cook for dinner.’

(Georgopoulos 1985: 73, ex. 14e)

Thus, it has been argued that the gap position is occupied by a null base-generated *pro* which is A’-bound by a null operator. Accordingly, since these gaps are derived via base-generation, no island effects are observed.

In Chinese, Aoun and Li (2003) provide examples where relative clauses with a gap exhibit island effects, as in (13), and examples where relative clauses with a gap show no island effects, as in (14) (Aoun & Li 2003: 178):

(13) *wo xihuan e_i chuan e_j de yifu_j de na-ge ren_i
 I like wear REL clothes REL that-CL person
 ‘the person_i that I like the clothes_j that e_i wears e_j ’

(14) e_i chuan e_j de yifu_j hen piaol[a]ing de na-ge
 wear REL clothes very pretty REL that-CL
 ren_i
 person
 ‘the person_i that the clothes_j that e_i wears e_j are pretty’

This seems to suggest that some gaps in Chinese relative clauses involve movement, i.e., where island effects are present, and that others do not, i.e., where island effects are absent. More interestingly, since Chinese is also a pro-drop language, a well-reasoned hypothesis of the gaps insensitive to island constraints is that they are *pros* as well. Indeed, adopting Huang’s (1982) analysis of *pro*, Aoun and Li suggest that gaps insensitive to island constraints are associated with a base-generated *pro*. Furthermore, they indicate that cases where island effects are absent are always linked with an island that occurs in the subject position, as in (14), or a topic position.

While Aoun and Li’s first observation that some gaps exhibit island effects and that others do not is correct, their second observation that a violation of island constraints always involves an island in the subject position is, however, not free from error. In the following section, I provide more data indicating that not only do subject and object gaps behave differently, but gaps in further embedded clauses may

vary with respect to their sensitivity to different island conditions. I will then examine gaps in the Weak Crossover Condition in Section 2.2, and show that the weak crossover phenomenon among gaps does not always match the results associated with their sensitivity to islandhood. In Section 2.3, I digress to discuss the asymmetries between empty categories in Chinese subject and object positions in the previous studies; in particular, I focus on their different potential to enter the derivation with a base-generated *pro*.

2.1 *Gaps and Island Conditions*

In Chinese argument relative clauses, gaps are obligatory in both the matrix subject and matrix object positions, but are interchangeable with a resumptive pronoun in the embedded subject and embedded object positions, and are excluded from the oblique object positions (as summarized above in Table 16).

Nevertheless, gaps in different positions are not always consistent with their sensitivity to island conditions. I will discuss gaps in the Wh-island Condition, Adjunct Island Condition and the Complex NP Condition in the following sections.

2.1.1 Gaps in the Wh-island Condition

In the Wh-island Condition, the presence of island effects depends on the relationship between the position of the entire relativized NP and the position of the embedded gap. Subject gaps do not exhibit island effects whether the entire relativized NP is in the subject position, as in (15), or in the object position, as in (16).

Relativized NP in the subject position, gap in subject position

- (15) na-ge Zhangsan xiang zhidao e_i weisheme xie-le
 that-CL Zhangsan want know why write-ASP
 zhe-ben shu de zuojia_i likai taiwan le.
 this-CL book REL writer leave Taiwan SFP

‘The writer_i who Zhangsan wonders why e_i wrote this book left Taiwan.’

Relativized NP in the object position, gap in subject position

- (16) Lisi kan-guo na-ge Zhangsan xiang zhidao e_i
 Lisi see-ASP that-CL Zhangsan want know
 weisheme xie-le zhe-ben shu de zuojia_i
 why write-ASP this-CL book REL writer

‘Lisi saw the writer_i who Zhangsan wonders why e_i wrote this book.’

By contrast, an island effect is observed in object gaps when the entire relativized NP is in the subject position, as in (17), but disappears when the relativized NP is in the object position, as in (18).

Relativized NP in the subject position, gap in object position

- (17) ??/*na-ben Zhangsan xiang zhidao Lisi weisheme xie-le
 that-CL Zhangsan want know Lisi why write-ASP
 e_j de shu_j hen zhongyao.
 REL book very important

‘The book_j which Zhangsan wonders why Lisi wrote e_j is important.’

Relativized NP in the object position, gap in object position

- (18) wo du-guo na-ben Zhangsan xiang zhidao Lisi
I read-ASP that-CL Zhangsan want know Lisi
weisheme xie-le e_j de shu_j
why write-ASP REL book
'I read the book_j which Zhangsan wonders why Lisi wrote e_j .'

Similar to subject gaps, there are no island effects of possessor gaps in the Wh-island Condition, no matter whether the relativized NP is in the subject position, as in (19), or in the object position, as in (20).

Relativized NP in the subject position, gap in a possessed NP

- (19) na-ge Lisi xiang zhidao e_i erzi weisheme xihuan
that-CL Lisi want know son why like
Zhangsan de mama_i likai xuexiao le.
Zhangsan REL mother leave school SFP
'The mother_i who Lisi wonders why e_i 's son likes Zhangsan left school.'

Relativized NP in the object position, gap in a possessed NP

- (20) Wangwu kan-guo na-ge Lisi xiang zhidao e_i erzi
Wangwu see-ASP that-CL Lisi want know son
weisheme xihuan Zhangsan de mama_i
why like Zhangsan REL mother
'Wangwu saw the mother_i who Lisi wonders why e_i 's son likes Zhangsan.'

However, Wh-island effects emerge when gaps are further embedded, regardless of where the entire relativized NP occurs, as illustrated in (21-22) for subject gaps, (23-24) for object gaps and (25-26) for possessor gaps.

Relativized NP in the subject position, gap in subject position

- (21) *na-ge Zhangsan xiang zhidao weisheme Lisi juede e_i
 that-CL Zhangsan want know why Lisi think
 xie-le zhe-ben shu de zuojia_i likai le.
 write-ASP this-CL book REL writer leave SFP

‘The writer_i who Zhangsan wonders why Lisi thinks e_i wrote this book left.’

Relativized NP in the object position, gap in subject position

- (22) *Lisi kan-guo na-ge Zhangsan xiang zhidao weisheme
 Lisi see-ASP that-CL Zhangsan want know why
 Wang juede e_i xie-le zhe-ben shu de ren
 Wang think write-ASP this-CL book REL man_i

‘Lisi saw the man_i who Zhangsan wonders why Wang thinks e_i wrote this book.’

Relativized NP in the subject position, gap in object position

- (23) *na-ben Zhangsan xiang zhidao weisheme Wang
 that-CL Zhangsan want know why Wang
 juede Lisi xie-le e_j de shu_j hen xin.
 think Lisi write-ASP REL book very new

‘The book_j which Zhangsan wonders why Wang thinks Lisi wrote e_j is new.’

Relativized NP in the object position, gap in object position

- (24) *wo du-guo na-ben Zhangsan xiang zhidao weisheme
I read-ASP that-CL Zhangsan want know why
Wangwu juede Lisi xie-le e_j de shu_j
Wangwu think Lisi write-ASP REL book
'I read the book_j which Zhangsan wonders why Wangwu thinks Lisi wrote e_j.'

Relativized NP in the subject position, gap in a possessed NP

- (25) *na-ge Lisi xiang zhidao weisheme Wangwu juede
that-CL Lisi want know why Wangwu think
e_i erzi xihuan Zhangsan de mama_i likai xuexiao le.
son like Zhangsan REL mother leave school SFP
'The mother_i who Lisi wonders why Wangwu thinks e_i's son likes Zhangsan
left school.'

Relativized NP in the object position, gap in a possessed NP

- (26) *Wangwu kan-guo na-ge Lisi xiang zhidao weisheme
Wangwu see-ASP that-CL Lisi want know why
Xiaoying juede e_i erzi xihuan Zhangsan de mama_i
Xiaoying think son like Zhangsan REL mother
'Wangwu saw the mother_i who Lisi wonders why Xiaoying thinks e_i's son
likes Zhangsan.'

2.1.2 Gaps in the Adjunct Island Condition

In the Adjunct Island Condition, I also examine the relationship between the position of the entire relativized NP and the position of the embedded gap.

Let's start with subject gaps. No island effects are observed whether the relativized NP is in the subject position, as in (27), or in the object position, as in (28).

Relativized NP in the subject position, gap in subject position

(27)	na-ge		ruguo	e_i	da-le	Zhangsan	Lisi
	that-CL		if		hit-ASP	Zhangsan	Lisi
	hui	hen	shengqi	de	nüsheng _i	pao-zou-le.	
	will	very	angry	REL	girl	run-away-SFP	

'The girl_i that if e_i hit Zhangsan Lisi would be angry ran away.'

Relativized NP in the object position, gap in subject position

(28)	wo	kan-guo	na-ge	ruguo	e_i	da-le	
	I	see-ASP	that-CL	if		hit-ASP	
	Zhangsan	Lisi	hui	hen	shengqi	de	nüsheng _i
	Zhangsan	Lisi	will	very	angry	REL	girl

'I saw the girl_i that if e_i hit Zhangsan Lisi would be angry.'

Object gaps inside adjunct clauses, in the same manner, do not exhibit island effects whether the entire relativized NP is in the subject position, as in (29a-b), or in the object position, as in (30a-b).

Relativized NP in the subject position, gap in object position

- (29) a. na-ge ruguo Zhangsan qin-le e_j Lisi
 that-CL if Zhangsan kiss-ASP Lisi
 hui hen shangxin de nüsheng;
 will very sad REL girl run-away-SFP
 ‘The girl_j that if Zhangsan kissed e_j Lisi would be very sad ran away.’
- b. na-ge ruguo Zhangsan xihuan e_j Lisi
 that-CL if Zhangsan like Lisi
 hui hen shangxin de nüsheng;
 will very sad REL girl hen piaoliang.
 very pretty
 ‘The girl_j that if Zhangsan likes e_j Lisi would be sad is pretty.’

Relativized NP in the object position, gap in object position

- (30) a. wo kan-guo na-ge ruguo Zhangsan qin-le
 I see-ASP that-CL if Zhangsan kiss-ASP
 e_j Lisi hui hen shangxin de nüsheng;
 Lisi will very sad REL girl
 ‘I saw the girl_j that if Zhangsan kissed e_j Lisi would be very sad.’
- b. ta xihuan na-ge ruguo Zhangsan qin-le
 he like that-CL if Zhangsan kiss-ASP
 e_j Lisi hui hen shangxin de nüsheng;
 Lisi will very sad REL girl
 ‘He likes the girl_j that if Zhangsan kissed e_j Lisi would be very sad.’

Relativized NP in the subject position, gap in subject position

(33) ?/*na-ge ruguo Wangwu juede e_i da-le Zhangsan
 that-CL if Wangwu think hit-ASP Zhangsan
 Lisi hui hen shengqi de nüsheng_i pao-zou-le.
 Lisi will very angry REL girl run-away-SFP
 ‘The girl_i that if Wangwu thinks e_i hit Zhangsan Lisi would be angry ran
 away.’

Relativized NP in the subject position, gap in object position

(34) ?/*na-ge ruguo Wangwu juede Zhangsan qin-le e_j
 that-CL if Wangwu think Zhangsan kiss-ASP
 Lisi hui hen shangxin de nüsheng_j pao-zou-le.
 Lisi will very sad REL girl run-away-SFP
 ‘The girl_j that if Wangwu thinks Zhangsan kissed e_j Lisi would be very sad ran
 away.’

Relativized NP in the subject position, gap in a possessed NP

(35) *na-ge ruguo Wangwu juede e_i erzi da-le
 that-CL if Wangwu think son hit-ASP
 Zhangsan Lisi hui hen shangxin de mama_i likai le.
 Zhangsan Lisi will very sad REL mother leave SFP
 ‘The mother_i that if Wangwu thinks e_i ’s son hit Zhangsan Lisi would be sad
 left.’

2.1.3 Gaps in the Complex NP Condition

Unlike the previous two conditions, in the Complex NP Condition, it is the position of the NP island but not the position of the relativized NP that influences the sensitivity of gaps to the constraint.

First, in subject gaps, no island effects are observed when the NP island is in the subject position, as in (36). But when the NP island is in the object position, island effects are observed, as in (37).

NP island in the subject position, gap in subject position

(36) e_i chuan e_j de yifu_j hen piaoli[a]ng de na-ge
 wear REL clothes very pretty REL that-CL

ren_i

person

‘the person_i that the clothes_j that e_i wears e_j are pretty’

(Aoun & Li 2003: 178)

NP island in the object position, gap in subject position

(37) *wo xihuan e_i chuan e_j de yifu_j de na-ge ren_i
 I like wear REL clothes REL that-CL person

‘the person_i that I like the clothes_j that e_i wears e_j ’

(Aoun & Li 2003: 178)

The same pattern is found in object gaps. Island effects only occur when the NP island is in the object position (39), but not when the island is in the subject position (38).

NP island in the subject position, gap in object position

- (38) na-jian e_i chuan e_j de na-ge ren_i hen shuai
 that-CL wear REL that-CL person very handsome
 de yifu_j
 REL clothes
 ‘the clothes_j that the person_i that e_i wears e_j is handsome’

NP island in the object position, gap in object position

- (39) *na-jian Lisi xihuan e_i chuan e_j de na-ge
 that-CL Lisi like wear REL that-CL
 ren_i de yifu_j
 person REL clothes
 ‘the clothes_i that Lisi likes the person_i that e_i wears e_j ’

The behaviors of possessor gaps in the Complex NP Condition are the same as subject and object gaps. No island effects occur when the NP island is in the subject position (40), but island effects emerge when the island is in the object position (41).

NP island in the subject position, gap in a possessed NP

- (40) na-ge e_i erzi chuan e_j de yifu_j hen gui
 that-CL son wear REL clothes very expensive
 de mama_i
 REL mother
 ‘the mother_i that the clothes_j that e_i ’s son wears e_j are expensive’

NP island in the object position, gap in a possessed NP

- (41) *na-ge Zhangsan xihuan e_i erzi chuan e_j de
 that-CL Zhangsan like son wear REL
 yifu_j de mama_i
 clothes REL mother
 ‘the mother_i that Zhangsan likes the clothes_j that e_i ’s son wears e_j ’

When gaps are further embedded within the NP island, island effects show up for all of them, as illustrated in (42-43) for subject gaps, (44-45) for object gaps, and (46-47) for possessor gaps.

NP island in the subject position, gap in subject position

- (42) *na-ge Zhangsan juede e_i chuan e_j de yifu_j
 that-CL Zhangsan think wear REL clothes
 hen piaoliang de nanren_i
 very pretty REL man
 ‘the man_i that the clothes_j that Zhangsan thinks e_i wears e_j are pretty’

NP island in the object position, gap in subject position

- (43) *na-ge wo xihuan na-jian Zhangsan juede e_i
 that-CL I like that-CL Zhangsan think
 chuan e_j de yifu_j de nanren_i
 wear REL clothes REL man
 ‘the man_i that I like the clothes_j that Zhangsan thinks e_i wears e_j ’

NP island in the subject position, gap in object position

(44) */?na-jian Zhangsan juede e_i chuan e_j de na-ge
that-CL Zhangsan think wear REL that-CL
ren_i hen shuai de yifu_j
person very handsome REL clothes

‘the clothes_j that the person_i Zhangsan thinks that e_i wears e_j is handsome’

NP island in the object position, gap in object position

(45) *na-jian Lisi xihuan Zhangsan juede e_i chuan e_j de
that-CL Lisi like Zhangsan think wear REL
na-ge ren_i de yifu_j
that-CL person REL clothes

‘the clothes_i that Lisi likes the person_i Zhangsan thinks that e_i wears e_j ’

NP island in the subject position, gap in a possessed NP

(46) *na-ge Zhangsan juede e_i erzi chuan e_j
that-CL Zhangsan think son wear
de yifu_j hen gui de mother
REL clothes very expensive REL mama_i

‘the mother_i that the clothes_j that Zhangsan thinks e_i ’s son wears e_j is expensive’

NP island in the object position, gap in a possessed NP

(47) *na-ge Zhangsan xihuan Lisi juede e_i erzi chuan e_j
 that-CL Zhangsan like Lisi think son wear
 de yifu_j de mama_i
 REL clothes REL mother
 ‘the mother_i that Zhangsan likes the clothes_j that Lisi thinks e_i’s son wears e_j’

Below in Table 18 I summarize the sensitivity of subject, object and possessor gaps to different island conditions when these gaps are in the first selected (island) clause.

Table 18: Least Embedded Gaps and Island Configurations

Gaps		Subject	Object	Possessor
Wh-island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	–	+
	Object	+	+	+
Adjunct Island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	+	+
	Object	+	+	+
Complex NP Configuration	<i>grammatical role of NP island</i>			
	Subject	+	+	+
	Object	–	–	–

Key: ‘+’ indicates that gaps are allowed to occur in that type of island configuration; that is, no island effects occur. ‘–’ indicates that gaps are prohibited from occurring in the configuration, as island effects would occur.

Thus, subject and possessor gaps appear to be less sensitive to island conditions than object gaps. If being immune from island constraints suggests that relative clause formation is derived via base-generation, then subject and possessor gaps seem more likely to be associated with base-generation than movement, whereas object gaps

would be ambiguous between base-generation and movement, depending on the condition.

When gaps are further embedded in the selected island clauses, however, island effects occur in all conditions. The sensitivity of these further embedded gaps to different island constraints is summarized in Table 19 below. In these further embedded gaps, the ubiquitous island effects seem to indicate that they are derived via movement.

Table 19: Further Embedded Gaps and Island Configurations

Gaps		Subject	Object	Possessor
Wh-island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	–	–	–
	Object	–	–	–
Adjunct Island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	–	–	–
	Object	–	–	–
Complex NP Configuration	<i>grammatical role of NP island</i>			
	Subject	–	–	–
	Object	–	–	–

In the next section, I examine the properties of gaps with respect to another common test for movement: the Weak Crossover Condition.

2.2 *A'-Gaps and the Weak Crossover Condition*

The Weak Crossover Condition can be taken as a diagnostic of A'-movement as it prohibits a variable left by A'-movement from being bound by an A-binder (e.g., Wasow 1972). We can observe that the trace left behind by A-bar movement is ungrammatical when the chain crosses a coindexed pronoun within a c-commanding

DP. This is the Weak Crossover effect. For example, in (48), a trace t_i is coindexed with a pronoun within a c-commanding DP *his_i sister*, thus exhibiting the Weak Crossover effect (data from Wasow 1972: 136):

(48) ?*Who_i did Mary talk about his_i sister to t_i ?

The Weak Crossover Effect is operative in Chinese relative clauses, as can be seen in (49a-b).

- (49) a. *zhe-ge ta_i laopo bao-zhe e_i de nanren_i
 this-CL he wife hug-ASP REL man
 ‘the man_i that his_i wife is hugging e_i’
- b. *na-ge ta_i mama hen xihuan e_i de nanhai_i
 that-CL he mother very like REL boy
 ‘the boy_i that his_i mother likes e_i a lot’

In these examples, the matrix object gap cannot be coindexed with a coreferential pronoun within a DP that precedes it. The presence of weak crossover effects suggests that movement is involved in the derivation of object relatives.

By contrast, possessor gaps, embedded subject gaps and embedded object gaps do not exhibit weak crossover effects, as illustrated in (50-52), respectively.

Possessor gap (in the matrix clause and in the embedded clause)

- (50) a. na-ge ta_i laopo zheng bao-zhe e_i erzi
that-CL he wife PROG hug-ASP son
de nanren_i
REL man
'the man_i that his_i wife is hugging e_i son'
- b. na-ge ta_i laopo yiwei e_i erzi taoyan shuxue
that-CL he wife think son hate math
de nanren_i
REL man
'the man_i that his_i wife thought e_i son hates math'

Embedded subject gap

- (51) na-ge ta_i mama yizhi juede e_i taoyan shuxue de nanhai_i
that-CL he mother always think hate math REL boy
'the boy_i that his_i mother always thinks that e_i hates math'

Embedded object gap

- (52) na-ge ta_i mama yizhi juede laoshi taoyan e_i de nanhai_i
that-CL he mother always think teacher hate REL boy
'the boy_i that his_i mother always thinks that the teacher hates e_i'

Since only matrix object gaps exhibit weak crossover effects, this suggests that possessor gaps and embedded gaps may be derived via base-generation.

Table 20 summarizes the presence/absence of weak crossover effects of gaps in different positions.

Table 20: Gaps and Weak Crossover Effects

Gaps	Matrix Clause			Embedded Clause		
	Subject	Object	Possessor	Subject	Object	Possessor
WCO Effects	NA	+	-	-	-	-

Key; ‘+’ indicates the presence of WCO effects; ‘-’ indicates the absence of WCO effects; ‘NA’ means that argument position does not apply or is hard to test.

In the following section, I digress to discuss previous studies that suggest an empty category in the subject position is more likely to be a base-generated *pro*, whereas it appears that an empty category in the object position is more likely to involve PF-deletion of a copy involved in A’-movement instead of a base-generated *pro*. By the end, I conclude that this kind of categorization of gaps is insufficient as more data is taken into account.

2.3 *Asymmetries Between Subject and Object Gaps in Chinese*

Based on Chomsky’s NP typology (e.g., 1982, 1995), there are seven types of noun phrases, including covert counterparts of overt noun phrases, as summarized in Table 21. While all the overt NPs can be base-generated (i.e., be present in the lexicon at the beginning of a derivation), among the covert NPs only *pro* and PRO can be base-

generated.³⁰ The NP-trace is a derived object from A-movement, including passives, raising, and so forth, whereas the wh-trace/variable is a derived object from A'-movement, including QR (quantifier raising) and wh-movement.

Table 21: Types of NPs

	Overt	Covert
+anaphor, –pronominal	reciprocals/reflexives	NP-trace
–anaphor, +pronominal	pronoun	<i>pro</i>
+anaphor, +pronominal	--	PRO
–anaphor, –pronominal	name	wh-trace/variable

What is most relevant here is the covert counterpart of the pronoun, which is *pro*. *Pro* is a null pronominal; in other words, it behaves like a normal pronoun, and obeys Binding Principle B. Not all languages have a *pro*, for example, English.

Chinese is a pro-drop language, allowing a pronominal subject to be left unexpressed in a finite clause, hence a *pro*, as in (53) (e.g., Huang 1982, 1984, 1987).

- (53) Zhangsan shuo [[*pro*] xiaowu hui lai].
 Zhangsan say afternoon will come
 ‘Zhangsan said that he will come this afternoon.’

(Huang 1982: 352)

³⁰ A standard assumption of generative grammar (e.g., Chomsky 1995), is that a language is comprised of two components: a lexicon and a computational system. The lexicon specifies what enters into the computational system, and the computational system uses these items to generate derivations and linguistic expressions. Thus, what enters into the computational system is considered a base-generated element, and is distinguished from the empty positions that are generated during derivations (i.e., generated in the computational system).

In addition to the subject position, the object position appears to allow a similar empty category. If we consider the following examples, all of speaker B's answers are perfectly fine. Specifically, the object can be empty in (54c-f).

- (54) Speaker A: Zhangsan kanjian Lisi le ma?
 Zhangsan see Lisi ASP Q
 'Did Zhangsan see Lisi?'
- Speaker B:
- a. ta kanjian ta le.
 he see he ASP
 'He saw him.'
- b. *e* kanjian ta le.
 '[He] saw him.'
- c. ta kanjian *e* le.
 'He saw [him].'
- d. *e* kanjian *e* le.
 '[He] saw [him].'
- e. wo cai [*e* kanjian *e* le].
 I guess see ASP
 'I guess [he] saw [him].'
- f. Zhangsan shuo [*e* kanjian *e* le].
 Zhangsan say see ASP
 'Zhangsan said that [he] saw [him].'

(Huang 1984: 533)

But as argued by Huang (1982, 1984, 1987, 1989), only the subject *e* can be a base-generated *pro*. The *e* in the object position cannot be a (base-generated) *pro*, but must be a variable bound by an A'-binder instead (e.g., a null topic in the sentence). The main argument is based on a subject-object asymmetry among these empty categories. In particular, there appear to be more restrictions on the antecedent of object *e* than subject *e*. First, consider the subject *e* in (55a) with the object *e* in (55b). In (55a), the subject *e* can take the matrix subject *Zhangsan* as its antecedent, whereas in (55b), the object *e* cannot. Were the latter a *pro*, it would be a puzzle why this pronominal, unlike the one in the subject position, cannot be bound by a nonlocal subject, which is not expected by Principle B. But if object *e* is treated as a variable, then the fact that it cannot be bound by an A-binder (e.g., the nonlocal subject) is well accounted for: an R-expression needs to be free from A-binders.

- (55) a. Zhangsan_i xiwang e_i keyi kanjian Lisi.
 Zhangsan hope can see Lisi
 ‘Zhangsan_i hopes that e_i can see Lisi.’
- b. *Zhangsan_i xiwang Lisi keyi kanjian e_i.
 Zhangsan hope Lisi can see
 ‘Zhangsan_i hopes that Lisi can see e_i’

(Huang 1984: 538)

Another piece of evidence that suggests subject *e* can be a *pro*, but object *e* should be an A'-bound variable comes from such examples as (56). The object *e* can

only be bound by the head of a relative clause but not by an NP in the argument position (Huang 1984: 543):

- (56) Li Xiaojie hai zhao-budao yige *e* xinzhong xihuan *e*
 Li Miss still find-cannot one in.heart like
 de nanren.
 REL man
 ‘Miss Li still cannot find a man who she loves in her heart.’
 NOT: ‘Miss Li still cannot find a man who loves her in his heart.’

In this example, there are two empty categories. If each *e* can be bound by the NP *Miss Li* within the main clause, then potentially there should be two interpretations for this sentence. On one interpretation, the subject *e* is bound by *Miss Li*, and the object *e* is bound by the head of the relative clause ‘a man.’ On the other interpretation, the subject *e* is bound by the head of the relative clause and the object *e* bound by *Miss Li*. Nevertheless, only the first interpretation is available. This indicates that object *e* cannot be bound by an NP in the argument position (an A-binder), but can be bound by an operator (an A’-binder). The subject *e* does not have such restrictions.

Thus, subject *e* can be a base-generated null pronominal *pro*, but object *e* cannot and has to be A’-bound. As pointed by Huang (1984), this analysis brings about some desirable consequences, one of which is the subject-object asymmetry in terms of the strong crossover phenomenon. Compare (57a) with (57b):

- (57) a. [e'_i shuo e_i mei kanjian Lisi] de neige
 say no see Lisi REL that
 ren_i
 man
 ‘*the man who_i e_i said e_i didn’t see Lisi’
- b. *[e'_i shuo Lisi mei kanjian e_i] de neige
 say Lisi no see REL that
 ren_i
 man
 ‘*the man who_i e_i said Lisi didn’t see e_i’

(Huang 1984: 559)

As the strong crossover indicates that a pronoun cannot bind a moved NP that it c-commands, the fact that subject *e* in (57a) does not exhibit strong crossover effects but object *e* in (57b) does indicates that the former is not derived via movement but the latter is.

Thus far, Huang’s observation of asymmetries between subject and object empty categories, including available interpretations of their antecedents and the strong crossover phenomenon alike, suggests that in A’-dependencies the subject gap may be associated with a base-generated *pro* when entering the derivation, whereas the object gap cannot enter the derivation associated with a base-generated *pro*, although it can be a result of PF-deletion of a copy in A’-movement. This would seem to be on the right track if only matrix gaps are considered, as the consequence of this

analysis would fit in well with least embedded gaps in terms of their sensitivity to island constraints (as in Table 18) as well as to Weak Crossover Condition (as in Table 20) — object gaps exhibit more movement properties than subject gaps.

Despite the asymmetry between subject and object gaps, Huang (1987) does not exclude the possibility that an object gap can still enter the derivation as a *pro* first and then get bound as a variable by a base-generated operator at LF. This idea is mainly based on Xu’s (1986) examples where an object gap associated with a topicalized head does not exhibit (complex NP) island effects, as in (58a-b).

- (58) a. zhe-ben shu_j, [e_i du guo e_j] de ren_i bu
 this-CL book read ASP REL man not
 duo.
 many
 ‘This book_j, there aren’t that many people who_i read e_j.’
- b. zhe-ben shu_j, wo mei jian guo yige [e_i
 this-CL book I not see ASP one
 neng dudong e_j] de ren_i
 can read.understand REL man
 ‘This book_j, I haven’t met anyone who_i can understand e_j.’

(Xu 1986: 80)

What is worth pointing out here is that the example in (58b) conflicts with what has been observed in gaps in relative clauses. As discussed in Section 2.1.3, the least

embedded object gaps exhibit island effects in the Complex NP Condition when the NP island occurs in the object position, as in (59). But in (58b), the NP island *anyone who can understand [it]* is in the object position, and yet no island effect is observed.

- (59) *na-jian wo xihuan [e_i chuan e_j de na-ge
that-CL I like wear REL that-CL
ren_i] de yifu
person REL clothes
‘the clothes_i that I like [the person_i that e_i wears e_j]

Taken together, even though the least embedded subject gaps in relative clauses appear to exhibit fewer movement properties than object gaps in terms of island sensitivity and weak crossover phenomenon, I contend that it is still possible for gaps in both positions to be derived via movement or base-generation, simply based on mixed properties in both. In the following section, I provide a thorough investigation of resumption in relative clauses, and then return in Section 4 to discuss evidence that indicates gaps sometimes share similar properties with resumptive pronouns in Chinese relative clauses.

3 Overt Resumption

Resumption in A'-dependencies exhibits a variety of properties, but not all are consistent with each other crosslinguistically (see also Rouveret 2011 for a comprehensive examination of resumption). First, relative clauses with a resumptive pronoun in different languages vary in terms of their sensitivity to island constraints.

For example, in Hebrew, Irish, Standard Arabic, Persian, Greek and Tuki, resumption in relative clauses does not exhibit island effects (Hebrew: Borer 1984; Irish: McCloskey 1979; Standard Arabic: Aoun et al. 2010; Persian: Karimi 2001; Greek: Alexopoulou 2006; Tuki: Biloa 1990), whereas resumption in Welsh does exhibit island effects (Tallerman 1983).³¹

Second, resumption does not display the same reconstruction properties across languages. In some languages, reconstruction is sensitive to islands. In Lebanese Arabic, one type of resumptive pronoun allows variable binding in non-island contexts but forbids it when within islands (Aoun et al. 2001). By contrast, French and Jordanian Arabic have one type of resumption that allows variable binding within strong islands (Guilliot & Malkawi 2006, 2007). Furthermore, in nonisland contexts, obligatory resumptive pronouns in Hebrew allow variable binding and embedded idiom interpretation, whereas optional resumptive pronouns do not (Sichel 2014).

Third, some languages do not always exhibit weak crossover effects in resumption, such as Hebrew (Shlonsky 1992) and Welsh (Hendrick 1988), but others do, including Swedish (Engdahl 1985) and Irish (McCloskey 2011).

In Chinese, Aoun and Li (2003) have shown that overt resumption in relative clauses does not exhibit island effects or reconstruction properties. While it is true that resumption does not obey island constraints, in some cases it displays reconstruction properties, thus posing a problem to Aoun and Li's original analysis. I start with a review of Aoun and Li's (2003) observation and analysis of resumption in

³¹ In Greek, Alexopoulou (2006) points out that there are two types of resumption in relative clauses: one is restricted to non-argument positions and sensitive to islands, and the other can occur in both argument and non-argument positions but is not sensitive to islands. My focus is primarily on resumption in argument positions and its related properties.

Chinese relative clauses in Section 3.1. Then, I provide a comprehensive examination of resumption concerning islandhood in Section 3.2, Weak Crossover Condition in Section 3.3, and then the reconstruction properties in Section 3.4.

3.1 Aoun & Li (2003): Overt Resumption in Chinese Relative Clauses

Aoun and Li (2003) argue that overt resumption in Chinese argument relative clauses is derived via base-generation because of two independent pieces of evidence. First, the resumptive pronoun can occur within an island, as in (60a-b).

- (60) a. wo xiang kan na-ge ni yinwei ta_i bu
 I want see that-CL you because he not
 hui lai hen shengqi de xuesheng_i
 will come very angry REL student
 ‘I want to see the student_i that you are angry because he_i would not
 come.’
- b. wo xiang kan na-ge ni yaoqing e_i
 I want see that-CL you invite
 dai ta_j lai de ren_i lai zher de xuesheng_j
 bring him come REL person come here REL student
 ‘I want to see the student_j that you invited the person_i over that brought
 him_j over.’

(Aoun & Li 2003: 170)

In (60a), the resumptive pronoun *ta* ‘him’ is within an adjunct island; in (60b), the pronoun is within a complex NP island. No island effects are obtained, since these sentences are grammatical.

Second, reconstruction of the head noun to the resumptive pronoun is not available, as in (61a-b).

- (61) a. *wo xiang kan ni shuo meigeren_i hui dai
 I want see you say everyone will bring
 ta_j huilai de ziji_i de pengyou_j
 him back REL self REL friend
 ‘I want to see self_i’s friend that you said that everyone_i would bring
 him_j back.’
- b. *wo xiang kan ni shuo meigeren_i hui dai
 I want see you say everyone will bring
 ta_i huilai de wo yijing jieshao gei ta_i
 him back REL I alreadyintroduce to him
 de pengyou_j
 REL friend
 ‘I want to see the friend_j that I have introduced to him_i that you said
 everyone_i would bring him_j back.’

(Aoun & Li 2003: 169)

In the above examples, the quantifier phrase *meigeren* ‘everyone’ cannot bind the anaphor *ziji* ‘self’, which contrasts with the examples where a gap replaces the resumptive pronoun, as in (62a-b).

- (62) a. wo xiang kan ni shuo meigeren_i hui dai t_j
 I want see you say everyone will bring
 huilai de ziji_j de pengyou_j
 back REL self PNM friend
 ‘I want to see self_i’s friend that you said that everyone_i would bring t_i
 back.’
- b. wo xiang kan ni shuo meigeren_i hui dai t_j
 I want see you say everyone will bring
 huilai de wo yijing jieshao gei ta_i de
 back REL I alreadyintroduce to him REL
 pengyou_j
 friend
 ‘I want to see the friend_j that I have introduced to him_i that you said
 everyone_i would bring t_j back.’

(Aoun & Li 2003: 170)

The contrast between overt resumption and gaps in terms of reconstruction indicates that, while the noun head of a relative clause does not originate from the relativized

position in relative clauses using resumptive pronouns, the head of a relative clause using gaps does (or, at least, *can*).

Furthermore, they argue that overt resumption involves a base-generated relative operator in the specifier of CP, where the operator is co-indexed with the head noun NP, as in (63).

(63) [CP Op_i [IP ... [NP pronoun_i] ...]] [Head NP]_i

(Aoun & Li 2003: 179)

The evidence that supports the existence of a null operator in the specifier of CP comes from the fact that Chinese relatives with overt resumptive pronouns cannot contain a *wh*-interrogative, as in (64a), as opposed to relative clauses that contain gaps, as in (64b) (*t* indicates a trace left by A'-movement under Aoun and Li's analysis).

(64) a. *shei xihuan shei dasuan qing ta_i lai yanjiang
 who like who plan ask him come talk
 de zuojia_i
 REL author

‘Who likes the author_i that who planned to ask him_i to come to talk?’

- b. shei xihuan shei dasuan qing t_i lai yanjiang
 who like who plan ask come talk
 de zuojia_i
 REL author

(Aoun & Li 2003: 171)

Specifically, they argue that a relative clause with an overt resumptive pronoun is incompatible with the wh-interrogative because, adopting Huang (1982), the null operator Op_i in the specifier of CP prevents a wh-word in the embedded clause from being licensed by an embedded interrogative operator and thus cannot be interpreted with an interrogative operator in the matrix clause.

In brief, under Aoun and Li's analysis, the resumptive pronoun in Chinese argument relative clauses is always base-generated, and bound by a base-generated A'-operator.

3.2 *Overt Resumption and Island Conditions*

Recall that overt resumptive pronouns in Chinese relative clauses are obligatory in oblique object positions, excluded from the highest subject and direct object positions, and optional in possessor, embedded subject and embedded object positions (as summarized above in Table 17). Similar to the discussion on gaps and island conditions, I discuss resumption in the Wh-island Condition, Adjunct Island Condition, and Complex NP Condition each in turn in the following sections. Specifically, for the first two conditions, I focus on whether the position of the entire relativized NP and the position of the resumptive pronoun make any difference in

terms of sensitivity to island constraints. As for the last condition, I look at the relationship between the NP island and the position of the resumptive pronoun. Furthermore, in order to examine whether there is a contrast between resumptive pronouns in terms of embeddedness, I compare resumptive pronouns in the first selected (island) clause with the ones in further embedded clauses.

3.2.1 Overt Resumption in the Wh-island Condition

In the Wh-island Condition, there is no island effect regardless of which position the relativized NP occurs in, as in (65-66) for oblique object resumptive pronouns, (67-68) for possessor resumptive pronouns, (69-70) for subject resumptive pronouns and (71-72) for object resumptive pronouns.

Relativized NP in the subject position, resumptive pronoun in oblique object position

(65) na-ge Zhangsan xiang zhidao Lisi weisheme dui ta_i
 that-CL Zhangsan want know Lisi why to he
 hui-shou de zuojia_i likai taiwan le.
 wave-hand REL writer leave Taiwan SFP

‘The writer_i who Zhangsan wonders why Lisi waved hands at him_i left Taiwan.’

Relativized NP in the object position, resumptive pronoun in oblique object position

- (66) Lisi kan-guo na-ge Zhangsan xiang zhidao
 Lisi see-ASP that-CL Zhangsan want know
 Wangwu weisheme dui ta_i hui-shou de zuojia_i
 Wangwu why to he wave-hand REL writer
 ‘Lisi saw the writer_i who Zhangsan wonders why Wangwu waved hands at
 him_i.’

Relativized NP in the subject position, resumptive pronoun in possessor position

- (67) na-ge Zhangsan xiang zhidao ta_i erzi weisheme
 that-CL Zhangsan want know she son why
 da-le ren de mama_i likai xuexiao le.
 hit-ASP person REL mother leave school SFP
 ‘The mother_i who Zhangsan wonders why her_i son hit people left school.’

Relativized NP in the object position, resumptive pronoun in possessor position

- (68) Lisi kan-guo na-ge Zhangsan xiang zhidao
 Lisi see-ASP that-CL Zhangsan want know
 ta_i erzi weisheme da-le ren de mama_i
 she son why hit-ASP person REL mother
 ‘Lisi saw the mother_i who Zhangsan wonders why her_i son hit people.’

Relativized NP in the subject position, resumptive pronoun in subject position

(69) na-ge Zhangsan xiang zhidao ta_i weisheme da-le
 that-CL Zhangsan want know he why hit-ASP
 ren de xuesheng_i; likai xuexiao le.
 person REL student leave school SFP

‘The student_i who Zhangsan wonders why he_i hit people left school.’

Relativized NP in the object position, resumptive pronoun in subject position

(70) Lisi kan-guo na-ge Zhangsan xiang zhidao
 Lisi see-ASP that-CL Zhangsan want know
 ta_i weisheme da-le ren de xuesheng_i;
 he why hit-ASP person REL student

‘Lisi saw the student_i who Zhangsan wonders why he_i hit people.’

Relativized NP in the subject position, resumptive pronoun in object position

(71) na-ge Zhangsan xiang zhidao Lisi weisheme da-le
 that-CL Zhangsan want know Lisi why hit-ASP
 ta_i de xuesheng_i; likai xuexiao le.
 he REL student leave school SFP

‘The student_i who Zhangsan wonders why Lisi hit him_i left school.’

Relativized NP in the object position, resumptive pronoun in object position

- (72) Lisi kan-guo na-ge Zhangsan xiang zhidao
Lisi see-ASP that-CL Zhangsan want know
Wangwu weisheme da-le ta_i de xuesheng_i
Wangwu why hit-ASP he REL student
'Lisi saw the student_i who Zhangsan wonders why Wangwu hit him_i.'

When resumptive pronouns are further embedded within the NP island, there are still no island effects, as in (73-74) for oblique object resumptive pronouns, (75-76) for possessor resumptive pronouns, (77-78) for subject resumptive pronouns, and (79-80) for object resumptive pronouns.

Relativized NP in the subject position, resumptive pronoun in oblique object position

- (73) na-ge Zhangsan xiang zhidao weisheme Wangwu
that-CL Zhangsan want know why Wangwu
juede Lisi dui ta_i hui-shou de zuojia_i likai le.
think Lisi to him wave-hand REL writer leave SFP
'The writer_i who Zhangsan wonders why Wangwu wants to know why Lisi waved hands at him_i left.'

Relativized NP in the object position, resumptive pronoun in oblique object position

- (74) Lisi kan-guo na-ge Zhangsan xiang zhidao
Lisi see-ASP that-CL Zhangsan want know
weisheme Xiaoying juede Wangwu dui ta_i hui-shou
why Xiaoying think Wangwu to he wave-hand
de zuojia_i
REL writer
'Lisi saw the writer_i who Zhangsan wonders why Xiaoying thinks Wangwu
waved hands at him_i.'

Relativized NP in the subject position, resumptive pronoun in a possessed NP

- (75) na-ge Lisi xiang zhidao weisheme Wangwu juede
that-CL Lisi want know why Wangwu think
ta_i erzi xihuan Zhangsan de mama_i likai le
she son like Zhangsan REL mother leave SFP
'The mother_i who Lisi wonders why Wangwu thinks her_i son likes Zhangsan
left.'

Relativized NP in the object position, resumptive pronoun in a possessed NP

- (76) Wang kan-guo na-ge Lisi xiang zhidao weisheme Qing
Wang see-ASP that-CL Lisi want know why Qing
juede ta_i erzi xihuan Xin de mama_i
think she son like Xin REL mother
'Wang saw the mother_i who Lisi wonders why Qing thinks her_i son likes Xin.'

Relativized NP in the subject position, resumptive pronoun in subject position

- (77) na-ge Zhangsan xiang zhidao weisheme Lisi juede ta_i
that-CL Zhangsan want know why Lisi think he
xie-le zhe-ben shu de zuojia_i likai le.
write-ASP this-CL book REL writer leave SFP

‘The writer_i who Zhangsan wonders why Lisi thinks e_i wrote this book left.’

Relativized NP in the object position, resumptive pronoun in subject position

- (78) Lisi kan-guo na-ge Zhangsan xiang zhidao weisheme
Lisi see-ASP that-CL Zhangsan want know why
Wangwu juede ta_i xie-le zhe-ben shu de
Wangwu think he write-ASP this-CL book REL
zuoja_i
writer

‘Lisi saw the writer_i who Zhangsan wonders why Wangwu thinks e_i wrote this book.’

Relativized NP in the subject position, resumptive pronoun in object position

- (79) na-zhi Zhangsan xiang zhidao weisheme Wangwu
that-CL Zhangsan want know why Wangwu
juede Lisi mai-le ta_j de maomi_j hen keai.
think Lisi buy-ASP it REL cat very cute

‘The cat_j which Zhangsan wonders why Wangwu thinks Lisi bought it_j is cute.’

Relativized NP in the object position, resumptive pronoun in object position

(80) wo kan-guo na-zhi Zhangsan xiang zhidao weisheme

I see-ASP that-CL Zhangsan want know why

Wangwu juede Lisi mai-le ta_j de maomi_j

Wangwu think Lisi buy-ASP it REL cat

‘I saw the cat_j which Zhangsan wonders why Wangwu thinks Lisi bought it_j.’

3.2.2 Overt Resumption in the Adjunct Island Condition

In the Adjunct Island Condition, none of the resumptive pronouns shows any island effects whether the relativized NP is in the subject or object position, for example, as in (81-82) for oblique object resumptive pronouns, and (83-84) for possessor resumptive pronouns.

Relativized NP in the subject position, resumptive pronoun in oblique object position

(81) na-ge ruguo Zhangsan dui ta_i hui-shou,

that-CL if Zhangsan to she wave-hand

Lisi hui hen shangxin de nüsheng_i likai xuexiao

Lisi will very sad REL girl leave school

le.

SFP

‘The girl_i that if Zhangsan waves hands at her_i Lisi would be sad left school.’

Relativized NP in the object position, resumptive pronoun in oblique object position

- (82) Wangwu kan-guo na-ge ruguo Zhangsan dui ta_i
 Wangwu see-ASP that-CLif Zhangsan to she
 hui-shou, Lisi hui hen shangxin de nüsheng;
 wave-hand Lisi will very sad REL mother
 ‘Wangwu saw the girl_i that if Zhangsan waves hands at her_i Lisi would be
 sad.’

Relativized NP in the subject position, resumptive pronoun in a possessed NP

- (83) na-ge ruguo ta_i erzi da-le Zhangsan
 that-CL if she son hit-ASP Zhangsan
 Lisi hui hen shangxin de mama_i likai xuexiao le.
 Lisi will very sad REL mother leave school SFP
 ‘The mother_i that if her_i son hit Zhangsan Lisi would be sad left school.’

Relativized NP in the object position, resumptive pronoun in a possessed NP

- (84) Wangwu kan-guo na-ge ruguo ta_i erzi da-le
 Wangwu see-ASP that-CLif she son hit-ASP
 Zhangsan Lisi hui hen shangxin de mama_i
 Zhangsan Lisi will very sad REL mother
 ‘Wangwu saw the mother_i that if her_i son hit Zhangsan Lisi would be sad.’

Similar to the Wh-island Condition, resumptive pronouns still do not show any island effects in the Adjunct Island Condition when they are further embedded, as

in (85) for oblique object resumptive pronouns, and (86) for possessor resumptive pronouns.

Relativized NP in the subject position, resumptive pronoun in oblique object position

- (85) na-ge ruguo Wangwu juede Zhangsan dui ta_j
 that-CL if Wangwu think Zhangsan to she
 hui-shou Lisi hui hen shangxin de nüsheng_j
 wave-hand Lisi will very sad REL girl
 pao-zou le.
 run-away SFP
 ‘The girl_j that if Wangwu thinks Zhangsan waves at her_j Lisi would be very sad
 ran away.’

Relativized NP in the subject position, resumptive pronoun in a possessed NP

- (86) na-ge ruguo Wangwu juede ta_i erzi da-le
 that-CL if Wangwu think she son hit-ASP
 Zhangsan Lisi hui hen shangxin de mama_i likai
 Zhangsan Lisi will very sad REL mother leave
 xuexiao le.
 school SFP
 ‘The mother_i that if Wangwu thinks her_i son hit Zhangsan Lisi would be sad
 left school.’

3.2.3 Overt Resumption in the Complex NP Condition

In the Complex NP Condition, no island effect is observed in resumptive pronouns regardless of where the NP island occurs, as in (87-88) for oblique object resumptive pronouns, and (89-90) for possessor resumptive pronouns.

NP island in the subject position, resumptive pronoun in oblique object position

- (87) na-ge e_i du ta_j huishou de nansheng_j
 that-CL to her wave.hand REL boy
 hen shuai de nüsheng_j
 very handsome REL girl
 ‘the girl_j that the boy_i that e_i waved at her_j is very handsome’

NP island in the object position, resumptive pronoun in oblique object position

- (88) na-ge Zhangsan xihuan e_i du ta_j huishou
 that-CL Zhangsan like to her wave.hand
 de na-ge nansheng_i de nüsheng_j
 REL that-CL boy REL girl
 ‘the girl_j that Zhangsan likes the boy_i that that e_i waved at her_j’

NP island in the subject position, resumptive pronoun in a possessed NP

- (89) na-ge ta_i erzi chuan e_j de yifu_j hen
 that-CL she son wear REL clothes very
 gui de mama_i
 expensive REL mother
 ‘the mother_i that the clothes_j that her_i son wears e_j are expensive’

NP island in the object position, resumptive pronoun in a possessed NP

(90) na-ge Zhangsan xihuan ta_i erzi chuan e_j de
that-CL Zhangsan like she son wear REL
yifu_j de mama_i
clothes REL mother
‘the mother_i that Zhangsan likes the clothes_j that her_i son wears e_j’

Moreover, unlike gaps in further embedded clauses, no (or only slight) complex NP effects are observed in resumptive pronouns regardless of which position the NP island occurs in. For example, (91) and (92) illustrate the further embedded subject and object resumptive pronouns in an NP island that occurs in the object position, and both are fine.³²

NP island in the object position, resumptive pronoun in subject position

(91) na-ge Zhangsan xihuan na-kuai Lisi juede ta_i
that-CL Zhangsan like that-CL Lisi think she
mai-le e_j de dangao_j de nüsheng_i
buy-ASP REL cake REL girl
‘the girl_i that Zhangsan likes the cake_j that Lisi thinks she_i bought e_j’

³² Note that there is also another set of Chinese native speakers who consider these further embedded resumptive examples to be ungrammatical in all island conditions. Whether this is due to processing factors or any other grammatical issues, I will leave for future research, since any attempt to answer whether the ungrammaticality of further embedded resumptive pronouns and in fact, further embedded gaps as well, is due to processing or not, is beyond the scope of this study.

NP island in the object position, resumptive pronoun in object position

(92) na-zhi Zhangsan xihuan na-ge Lisi juede e_i
 that-CL Zhangsan like that-CL Lisi think
 mai-le ta_j de nüsheng;_j de maomi;_j
 buy-ASP it REL girl REL cat

‘the cat_j that Zhangsan likes the girl_i that Lisi thinks e_i bought it_j’

The results for overt resumption in different island conditions are summarized in Table 22 and Table 23.

Table 22: Least Embedded Resumptive Pronouns and Island Configurations

Resumptive Pronouns	Oblique Object	Possessor	Subject	Object
Wh-island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	+	+
	Object	+	+	+
Adjunct Island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	+	+
	Object	+	+	+
Complex NP Configuration	<i>grammatical role of NP island</i>			
	Subject	+	+	+
	Object	+	+	+

Key; ‘+’ indicates that resumptive pronoun are allowed to occur in that type of island configuration; that is, no island effects occur.

Table 23: Further Embedded Resumptive Pronouns and Island Configurations

Resumptive Pronouns	Oblique Object	Possessor	Subject	Object
Wh-island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	+	+
	Object	+	+	+
Adjunct Island Configuration	<i>grammatical role of relativized NP</i>			
	Subject	+	+	+
	Object	+	+	+
Complex NP Configuration	<i>grammatical role of NP island</i>			
	Subject	+	+	+
	Object	+	+	+

3.3 Resumption and the Weak Crossover Condition

The resumptive pronoun in relative clauses does not exhibit weak crossover effects, whether it is an obligatory resumptive pronoun in the oblique object position (93), an optional possessor resumption, as in (94a-b), or an optional resumption in the embedded subject position (95), or in the embedded object position (96).

Oblique Object Resumption

- (93) zhe-ge ta_j laopuo (zheng)zai dui ta_j hui-shou
 this-CL he wife PROG to he wave-hand
 de nanren_j
 REL man
 ‘the man_j that his_j wife is waving at him_j’

Possessor Resumption

- (94) a. na-ge ta_i laopo zheng bao-zhe ta_i erzi
that-CL he wife ASP hug-ASP he son
de nanren_i
REL man
'the man_i that his_i wife is hugging his_i son'
- b. na-ge ta_i laopo yiwei ta_i erzi taoyan shuxue
that-CL he wife think he son hate math
de nanren_i
REL man
'the man_i that his_i wife thought his_i son hates math'

Embedded Subject Resumption

- (95) zhe-ge ta_i laopuo juede ta_i zai shuijiao de
this-CL he wife think he PROG sleep REL
nanren_i
man
'the man_i that his_i wife thinks he_i is sleeping'

Embedded Object Resumption

- (96) zhe-ge ta_j laopuo juede Mei xihuan ta_j de nanren_j
 this-CL he wife think Mei like he REL man
 ‘the man_j that his_j wife thinks Mei likes him_j’

The results are summarized in Table 24.

Table 24: Resumptive Pronouns and Weak Crossover Effects

Resumptive Pronouns	Matrix Clause		Embedded Clause			
	Oblique Object	Possessor	Subject	Object	Oblique Object	Possessor
WCO Effects	–	–	–	–	–	–

Key: ‘–’ indicates the absence of WCO effects.

Based on the absence of weak crossover effects and island effects in resumption, it seems that relative clauses using resumptive pronouns do not involve movement at all. In the next section, I will first present a seeming counterexample to the base-generation analysis of overt resumption with respect to reconstruction, but then lay out more data to show that base-generation properties still hold consistently among relatives using resumptive pronouns.

3.4 Overt Resumption and Reconstruction

In this section, I provide new observational data of resumption concerning reconstruction properties.

First, contrary to Aoun and Li (2003), there are cases where overt resumption actually allows variable-binding, as in the matrix oblique object resumption (97). At

first this sentence may sound strange or awkward to some readers, but when a proper context is provided, it sounds fine. Here is a possible context: Imagine that there was a village where all the sons went to war and died. Each of these fathers, unfortunately, was left with only one picture of his son(s) to keep his memory alive. Somehow every father developed the same habit of talking with his son's picture every night before going to bed. One day a journalist from a remote town, whose name is Zhangsan, heard this story, and he wanted to come to visit these fathers and in particular, he wanted to see the picture of each father's son so that he could write an article about this.

Reconstruction of the Matrix Oblique Object Resumptive Pronoun

(97) Zhangsan xiang kankan me-ge fuqin_i (dou) hui dui
 Zhangsan want see every-CL father (all) will to
 ta_j shuohua de ta_i erzi de yizhao_j
 it speak REL his son PNM picture

‘Zhangsan wants to take a look at the picture_j of his_i son that every father_i will talk to’ (Lit. ‘Zhangsan wants to take a look at the picture of his son that every father will talk to it.’)

In addition to the matrix oblique object position, the overt resumption allows variable-binding in the embedded oblique object position (98), and in the embedded object position (99a-b).

Reconstruction of the Embedded Oblique Object Resumptive Pronoun

- (98) mei-ge fuqin_i dou yiwei Lisi hui gen ta_j shuohua
every-CL father all think Lisi will to it talk
de Lisi yao song gei ta_i de wanou_j, (qishi
REL Lisi will give to he REL doll in.fact
Lisi zhishi hui momo ta_j).
Lisi only will touch it

‘the doll_j that Lisi will give to him_i that every father_i thinks that Lisi will talk to’, (in fact Lisi will only touch them.)’ (Lit. ‘The doll that Lisi will give to him that every father thinks that Lisi will talk to.’)

Reconstruction of the Embedded (Direct) Object Resumptive Pronoun

- (99) a. mei-ge-ren_i dou yiwei Lisi yijing mai-xia ta_j
every-CL-person all think Lisi alreadybought-down it
de Lisi yao song gei ta_i de maomi_j,
REL Lisi will give to him REL cat
(qishi Lisi haimei mai).
in.fact Lisi not.yet buy

‘(the) cat_j that Lisi was going to give to him_i that everyone_i thought Lisi already bought, (in fact Lisi hasn’t bought them yet).’ (Lit. ‘The cat_j that Lisi was going to give to him_i that everyone_i thought Lisi already bought it.’)

b. Zhangsan xiang kan ni shuo meige nanren_i
 Zhangsan want see you say each man
 hui dai ta_j huilai de wo yijing jieshao gei
 will bring her back REL I already introduce to
 ta_i de nüpengyou_j
 him REL girl.friend

‘Zhangsan wants to see the girl friend_j that I have introduced to him_i that every man_i would bring back.’ (Lit. ‘Zhangsan wants to see the girlfriend_j I introduced to him_i that every man_i would bring her_j back.’)

If the existence of reconstruction properties suggests a movement analysis, then these examples would indicate that, at least under certain circumstances, relative clauses using resumptive pronouns are associated with movement instead of base-generation. However, more data reveal a different story. Variable-binding in resumption appears to be still available within islands, whether it is the oblique object resumption in Adjunct Island Condition, as in (100), in Wh-island Condition, as in (101), or the object resumption in Adjunct Island Condition, as in (102), and in Wh-island Condition, as in (103).

Reconstruction of oblique object resumption in Adjunct Islands

- (100) na-ge Zhangsan hen nanguo yinwei me-ge
 that-CL Zhangsan very sad because every-CL
 fuqin_i dou hui dui ta_j shuohua de ta_i erzi
 father all will to it talk REL his son
 de yizhao_j
 PNM picture
 ‘that picture_j of his_i son that Zhangsan was sad because every father_i will
 talk to it_j.’

Reconstruction of oblique object resumption in Wh-islands

- (101) na-ge Lisi xiang zhidao Zhangsan weishenme xiang kan
 that-CL Lisi want know Zhangsan why want see
 e_j de me-ge fuqin_i dou hui dui ta_j shuohua
 REL every-CL father all will to it talk
 de ta_i erzi de yizhao_j
 REL he son PNM picture
 ‘that picture_j of his_i son that Lisi wants to know why Zhangsan wants to take
 a look at e_j that every father_i will talk to it_j.’

Reconstruction of object resumption in Adjunct Islands

- (102) na-ge Zhangsan hen nanguo yinwei me-ge
 that-CL Zhangsan very sad because every-CL
 fuqin_i dou xiang bao-zhe ta_j shuijiao de ta_i erzi
 father all want hug-ASP it sleep REL his son
 de yizhao_j
 PNM picture
 ‘that picture_j of his_i son that Zhangsan was sad because every father_i wants to
 hug it_j while sleeping.’

Reconstruction of object resumption in Wh-islands

- (103) na-ge Lisi xiang zhidao Zhangsan weishenme xiang kan
 that-CL Lisi want know Zhangsan why want see
 e_j de me-ge fuqin_i dou yao bao-zhe ta_j
 REL every-CL father all want hug-ASP it
 shuijiao de ta_i erzi de yizhao_j
 sleep REL he son PNM picture
 ‘that picture_j of his_i son that Lisi wants to know why Zhangsan wants to take a
 look at e_j that every father_i needs to hug it_j while sleeping.’

Because variable binding with resumption is not sensitive to island effects, we cannot conclude that reconstruction is necessary to license it, nor that resumption is derived via movement. This pattern is similar to one type of resumption in French and

Jordanian Arabic, which allows variable binding within strong islands without involving movement (e.g., Guillot & Malkawi 2006, 2007). We will conclude that the availability of variable-binding in Chinese resumption in both non-island and island contexts indicates that no movement is involved.³³ Table 25 summarizes the results.

Table 25: Resumptive Pronouns (RPs) and Variable-Binding (VB)

RPs		Matrix Clause		Embedded Clause			
		Oblique Object	Possessor	Subject	Object	Oblique Object	Possessor
VB	non-island contexts	+	NA	NA	+	+	NA
	within islands	+	NA	NA	+	+	NA

Key: ‘+’ indicates the availability of variable binding; ‘-’ indicates the unavailability of variable binding. ‘NA’ means that argument position does not apply or is hard to test.

³³ The reconstruction effects exhibited by resumption are not limited to variable binding cases, but also extend to Condition C cases, as shown below (i). Note that the corresponding sentence where Condition C does not apply is fine, as in (ii).

- (i) *na-ge ta_i hen shengqi yinwei Lisi ganggang dui
 that-CL he very angry because Lisi just.now to
 ta_j paoxiao de Zhangsan_i de meimei_j
 she bawl.at REL Zhangsan POSS sister
 ‘the sister_j of Zhangsan_i that he_i was very angry because Lisi just bawled at her_j’
- (ii) na-ge ta_k hen shengqi yinwei Lisi ganggang dui
 that-CL he very angry because Lisi just.now to
 ta_j paoxiao de Zhangsan_i de meimei_j
 she bawl.at REL Zhangsan POSS sister
 ‘the sister_j of Zhangsan_i that he_k was very angry because Lisi just bawled at her_j’

At first Condition C effects may suggest movement, but since those show up within islands, they do not distinguish islands from nonislands.

4 When Gaps and Resumption in Chinese RCs Surprise Us

In this section, I discuss the shared properties among gaps and resumption in Chinese relative clauses, which suggest that the syntactic rules that govern gaps and resumptive pronouns may not always be distinctive. In Section 4.1, I examine coordinate structures that combines gapped relatives and resumptive relatives. In Section 4.2, I show that gapped relatives not only allow reconstruction in non-island contexts (Aoun & Li 2003), but also within islands, which makes them similar to resumptive relatives, as discussed in the previous section.

4.1 *Coordination of Gapped and Resumptive Relative Clauses*

Since Ross (1967), the Coordination Structure Constraint (CSC) has been applied quite often to test if both conjuncts exhibit the same syntactic properties. Specifically, the CSC excludes the kind of conjunction where extraction only applies to one conjunct but not the other, as in (104a-c).

- (104) a. *The lute which Henry plays and sings madrigals is warped.
b. *The madrigals which Henry plays the lute and sings sound lousy.
c. *The nurse who polished her trombone and the plumber computed my tax was a blonde.

(Ross 1967: 160)

The CSC also applies in Chinese relative clauses, as indicated by the ungrammaticality in (105).

- (105) *na-ben Zhangsan_i du-guo e_j erqie e_i likai
that-CL Zhangsan read-ASP and leave
jia de shu_j hen youqu.
home REL book very interesting
‘The book_j that Zhangsan_i read e_j and e_i left home is very interesting.’

There are also cases where a subject gap and an object gap can be coordinated with each other, as in (106).

- (106) na-ge [e_i xihuan Zhangsan] de erqie [Zhangsan
that-CL like Zhangsan REL and Zhangsan
ye xihuan e_i] de nüsheng
also like REL girl
Lit. ‘the girl_i that e_i likes Zhangsan and that Zhangsan likes e_i ’

The fact that they can be conjoined together suggests that it is possible for both subject gaps and object gaps to be derived via the same rules, whether base-generation or movement.

Importantly, gapped relatives can be conjoined with resumptive relatives, regardless of which argument position the gaps occur in. In (107a), the gap that is coordinated with resumption is in the matrix subject position; in (107b), the gap is in the matrix object position, and in (107c), the gap is in the possessor position.

- (107) a. na-ge [e_i xihuan Zhangsan] de erqie [Lisi
 that-CL like Zhangsan REL and Lisi
 gen ta_i shuo-guo hua] de xuesheng_i
 with she talk-ASP word REL student
 ‘the student_i that e_i likes Zhangsan and that Lisi talked with her_i.’
- b. na-ge [Zhangsan xihuan e_j] de erqie [Lisi
 that-CL Zhangsan like REL and Lisi
 gen ta_j shuo-guo hua] de xuesheng_j
 with she talk-ASP word REL student
 ‘the student_j that Zhangsan likes e_j and that Lisi talked with her_j.’
- c. na-ge [e_i laopo paodiao] de erqie [Lisi
 that-CL wife left REL and Lisi
 gen ta_i shuo-guo hua] de yiyuan_i
 with he talk-ASP word REL councilor
 ‘the councilor_i that e_i’s wife left and that Lisi talked with him_i.’

Coordination is also possible between embedded gaps and resumption, as in (108a-b).

- (108) a. na-ge [Wangwu juede e_i xihuan Zhangsan]
 that-CL Wangwu think like Zhangsan
 de erqie [Lisi zeng gen ta_i shuo-guo hua]
 REL and Lisi once with she talk-ASP word
 de xuesheng;
 REL student
 ‘the student_i that Wangwu thinks e_i likes Zhangsan and that Lisi
 talked with her_i’
- b. na-ge [Wangwu juede [Zhangsan xihuan e_j]
 that-CL Wangwu think Zhangsan like
 de erqie [Lisi zeng gen ta_j shuo-guo hua]
 REL and Lisi once with she talk-ASP word
 de xuesheng;
 REL student
 ‘the student_j that Wangwu thinks Zhangsan likes e_j and that Lisi talked
 with her_j’

Overall, the above examples suggest that gaps and resumption must involve the same types of chain-formation operation.

4.2 *Gaps and Reconstruction Within Islands*

According to Aoun and Li (2003), gaps in relative clauses exhibit reconstruction properties, as in (109a-b). The examples they provide are anaphor-binding in non-island contexts.

(109) a. wo xiang kan ni shuo meigeren_i hui dai t_j
 I want see you say everyone will bring
 huilai de ziji_j de pengyou_j
 back REL self PNM friend
 ‘I want to see self_i’s friend that you said that everyone_i would bring t_j
 back.’

b. wo xiang kan ni shuo meigeren_i hui dai t_j
 I want see you say everyone will bring
 huilai de wo yijing jieshao gei ta_i de
 back REL I alreadyintroduce to him REL
 pengyou_j
 friend
 ‘I want to see the friend_j that I have introduced to him_i that you said
 everyone_i would bring t_j back.’

(Aoun & Li 2003: 170)

Interestingly, the reconstruction properties of gaps, in particular, matrix object gaps, seem to be still available within islands, as in (110-111).

Reconstruction of matrix object gaps in Adjunct Islands

- (110) na-ge Zhangsan hen nanguo yinwei me-ge
 that-CL Zhangsan very sad because every-CL
 fuqin_i dou hui bao-zhe e_j shuijiao de ta_i erzi
 father all will hug-ASP sleep REL his son
 de yizhao_j
 PNM picture
 ‘that picture_j of his_i son that Zhangsan was sad because every father_i will hug
 e_j while sleeping.’

Reconstruction of matrix object gaps in Wh-islands

- (111) na-ge Lisi xiang zhidao Zhangsan weishenme xiang kan
 that-CL Lisi want know Zhangsan why want see
 e_j de me-ge fuqin_i dou hui bao-zhe e_j
 REL every-CL father all will hug-ASP
 shuijiao de ta_i erzi de yizhao_j
 sleep REL he son PNM picture
 ‘that picture_j of his_i son that Lisi wants to know why Zhangsan wants to take a
 look at e_j that every father_i will hug e_j while sleeping.’

This makes them similar to resumptive relative clauses, as shown in examples (100-103) above, as both exhibit reconstruction properties within islands.

In sum, the mixed properties of gaps may suggest that some of them are derived via base-generation and other via movement. In the following section, I lay out an overview of the behavior of gaps and resumption in argument relative clauses, and suggest a new approach.

5 The Revised Proposal

In Section 5.1, I provide the revised analysis of resumptive pronoun and gap structures in Chinese relative clauses. Then, in Section 5.2, I review and adopt Guilliot and Malkawi's (2006, 2007) analysis to account for the reconstruction properties of resumption and some gaps within islands.

5.1 *The Analysis of Resumptive Pronoun and Gap Structures*

Beginning with the analysis of resumption, as shown in Table 22, Table 23, and Table 24, resumption does not exhibit any island-type effects; no island effects and no WCO effects occur. Therefore, I adopt Aoun and Li's (2003) nonmovement analysis of resumption, as schematized in (112).

(112) $[_{CP} Op_i [_{IP} \dots [_{NP} pronoun_i] \dots]] [_{Head} NP]_i$

(Aoun & Li 2003: 179)

This structure involves a base-generated pronoun, which is bound by a base-generated null *Op* in the specifier of the relative clause. All these coreferential elements indicate that they have certain features (e.g., ϕ -features) that match with each other. Note that

we must conclude that the base-generated pronoun chain can be interpreted in such a way as to allow for reconstruction of the head NP for variable binding purposes.

Gaps in argument relative clauses exhibit both movement and nonmovement properties. The least embedded subject and possessor gaps are often associated with a nonmovement derivation, since for one thing, most island effects do not occur, and for another, they can be coordinated with relatives using resumptive pronouns. The least embedded object gaps exhibit two types of island effects and weak crossover effects, which suggests that they can be associated with movement. On the other hand, they do not exhibit island effects in some environments, and are also able to form a coordinate structure with relatives using resumptive pronouns. This suggests that they do not always need to involve movement.

In the further embedded clauses, as illustrated in Table 19 and Table 20, most of the island effects reappear, but none of the gaps exhibits weak crossover effects and all of them are able to be coordinated with relatives using resumptive pronouns. This again suggests that gaps can be derived via movement or base-generation, with constraints on when each must apply.

I do not attempt to explain when a movement pattern is required and when a base-generation pattern is required; however, given the patterns observed above, both must be available analyses for gapped structures in at least some contexts. Hence, I would like to suggest that Chinese relative clauses with gaps are ambiguous. To be more precise, they can either be derived via movement (i.e., Aoun & Li 2003), as in (113) or be derived via base-generation, as in (114), and that yet-to-be-understood constraints govern when one or the other structure applies. Note that the structure

- b. na-ge [Zhangsan xihuan e_j] de erqie [Lisi
that-CL Zhangsan like REL and Lisi
gen ta_j shuo-guo hua] de xuesheng_j
with she talk-ASP word REL student
‘the student_j that Zhangsan likes e_j and that Lisi talked with her_j
before’
- c. na-ge [e_i laopo paodiao] de erqie [Lisi
that-CL wife left REL and Lisi
gen ta_i shuo-guo hua] de yiyuan_i
with he talk-ASP word REL councilor
‘the councilor_i that e_i wife left and that Lisi talked with him_i before’

Such coordinate structures are possible because both conjuncts may have the same structure, one that is associated with base-generation, as schematized in (116).

(116) [[_{CP} Op_i [_{IP} ... [_{NP} pronoun_i] ...]] AND [_{CP} Op_i [_{IP} ... [_{NP} *pro*_i] ...]]] [_{Head} NP]_i

In brief, in this revised proposal I argue that gaps in Chinese argument relative clauses are often ambiguous between movement and nonmovement, and that resumption is always associated with base-generation. In the next section, I will start with a review of Guilliot and Malkawi (2006, 2007), and then explain why resumption and certain gaps in Chinese relative clauses allow reconstruction within islands.

5.2 *Guilliot & Malkawi (2006, 2007): Reconstruction*

Reconstruction has been considered to be a process by which a displaced phrase exhibits properties that indicate it is interpreted in the original place of base-generation (e.g., Chomsky 1977). In particular, since Aoun and Li (2001), cases where the resumptive pronoun and its antecedent are not separated by an island are treated as being associated with movement because reconstruction is available, as in (117). By contrast, when the resumptive pronoun is separated from its antecedent by an island, it is analyzed as being derived via base-generation, as no reconstruction occurs, as in (118) (adapted from Aoun & Li 2001: 381):

Lebanese Arabic

(117) təlmiiz-a_i l-kəsleen ma baddna nχabbir [wala
 student-her the-bad NEG want.1P tell.1P no
 mʃallme]_i ʔanno huwwe zaʃbar b-l-faħs
 teacher that he cheated.3S.MASC in-the-exam
 ‘Her bad student, we don’t want to tell any teacher that he cheated on the exam.’

(118) *təlmiiz-a_i l-kəsleen ma hkiina maʃ [wala
 student-her the-bad NEG talked.1P with no
 mʃallme]_i ʔabl-ma huwwe yuuʃal
 teacher before he arrive.3S.MASC
 ‘Her bad student, we didn’t talk to any teacher before he arrived.’

Thus, according to Aoun and Li (2001), resumption that allows reconstruction is derived via movement and hence cannot cross an island. They call this type of resumption *apparent resumption*. As for cases where the resumption does not allow reconstruction, they call this type the *true resumption*, as it does not involve movement. In other words, such an analysis is based on the fact that reconstruction occurs only in nonisland contexts, and that it is banned from occurring across island boundaries.

However, Guilliot and Malkawi (2006, 2007) show that one type of resumption in Jordanian Arabic and French allows variable-binding reconstruction within islands, as in (119a-b) and (120a-b), respectively. They call this *weak resumption*. Since weak resumption allows reconstruction within islands, these data pose a challenge to Aoun and Li's (2001) original categorization of resumption.

- (119) a. [ʔalib-ha_i l-kassul]_j l-mudiirah ziɕlat laʔannuh
 student-her the-bad the-principal upset.3S.FEM because
 [kul mɕalmih]_i ʃafat-uh_j ǧaʃ
 every teacher saw.3S.FEM-CL cheated.3S.MASC
 bi-li-mtiħan.
 in-the-exam
 ‘Her bad student, the principal got upset because every teacher saw
 him cheating in the exam.’

- b. [ʔalib-ha_i l-kassul]_j ma badku tisʔalu [wala
 student-her the-bad NEG want.1PL ask no
 mʕalmih]_i lajʃ l-mudiira kaḥʃat-uh_j mn
 teacher why the-principal expelled.3S.FEM-CL from
 l-madrase.
 the-shcool

‘Her bad student, you don’t want to ask any teacher why the principal expelled him from school.’

(*Jordanian Arabic*, Guilliot & Malkawi 2007: 118)

- (120) a. [La photo de sa_i classe]_j, tu es fâché parce que chaque prof_i l_j’a
 déchirée.
 ‘The picture of his class, you’re furious because each teacher tore it.’
- b. [Quelle photo de lui_i]_j es-tu fâché parce que chaque homme_i l_j’a
 déchirée?

‘Which picture of him are you furious because each man tore it?’

(*French*, Guilliot & Malkawi 2007: 118)

The other type of resumption in Jordanian Arabic, the type originally documented by Aoun and Li, allows reconstruction only when no island intervenes, which Guilliot and Malakawi called *strong resumption*. As for Aoun and Li’s cases, the examples in (121) indicate that reconstruction is available in nonisland contexts,

whereas the ones in (122) show that no reconstruction holds when intervened by strong islands.

- (121) [ṭalib-ha_i l-kassul]_j ma biddna nxabbir [wala
 student-her the-bad NEG want.1PL tell.1PL no
 mçalmih]_i ?inno hu_j ġaƒ bi-li-mtiħan
 teacher that he cheated.3SM in-the-exam
 ‘Her bad student, we don’t want to tell any teacher that he cheated in the
 exam.’

(Guilliot & Malkawi 2007: 121)

- (122) *[ṭalib-ha_i l-kassul]_j ma ħakjan maç [wala mçalmih]_i
 student-her the-bad NEG talked.1PL with no teacher
 gabl ma hu_j yesal.
 before he arrive.3S.MASC

‘Her bad student, we didn’t talk to any teacher before he arrived.’

(Guilliot & Malkawi 2007: 122)

In order to account for these types of resumption, Guilliot & Malkawi propose that, while strong resumption is derived via movement (i.e., along the lines of Aoun & Li 2001), weak resumption is derived via an NP-deletion analysis of resumptive pronouns. For the processes involved in the derivation of the latter type, they adopt Elbourne’s (2001) unifying theory of third-person pronouns and determiners.

Specifically, pronouns are treated as definite articles, and have the following structures that are associated with two different meanings:

- (123) a. [[the/it] NP]
 b. [the/it *i*]

In (123a), the NP complement is deleted under identity when it has a preceding antecedent. Alternatively, the NP complement is simply a variable, as in (123b).

Thus, the resumption pronoun can be interpreted as having either the structure associated with a full lower NP-complement, which gets deleted at PF, or having the structure where the variable replaces the full NP-complement. In cases where weak resumption obtains variable-binding while intervened by islands, as in (124), the interpretation in (123a) is applied, as illustrated in (125).

- (124) [ʔalib-ha_i l-kassul]_j l-mudiirah ziçlat laʔannuh
 student-her the-bad the-principal upset.3SF because
 [kul mçalmih]_i ʃafat-uh_j ğaʃ bi-li-mtihan.
 every teacher saw.3SF-CL cheated.3SM in-the-exam
 ‘Her bad student, the principal got upset because every teacher saw him
 cheating in the exam.’

- (125) [ṭalib-*ha*_i l-kassul]_j [kul mçalmih]_i ...
 student-her the-bad every teacher
 ...[DP -uh [NP ~~ṭalib-*ha*_i l-kassul~~]]_j
 [DP him [NP ~~bad student of her_i~~]]_j

(Guilliot & Malkawi 2007: 126)

As shown in (125), the lower copy now allows the pronoun *ha*_i ‘her’ to be bound by the quantificational phrase *kul mçalmih*_i ‘every teacher.’

Now let me explain how their analysis can also be applied to the reconstruction properties of Chinese resumption and gaps. In the relatives using resumptive pronouns, when the reconstruction occurs, as in (126), the interpretation associated with the full NP-complement deletion in (123a) is used, as in (127).

- (126) na-ge Zhangsan hen nanguo yinwei me-ge
 that-CL Zhangsan very sad because every-CL
 fuqin_i dou hui dui ta_j shuohua de ta_i erzi
 father all will to it talk REL his son
 de yizhao_j
 PNM picture

‘that picture_j of his_i son that Zhangsan was sad because every father_i will talk to it_j.’

- (127) [me-ge fuqin]_i ... [DP ta_j [~~NP ta_i erzi de yizhao~~]]_j
 every father_i [DP it_j [NP picture of his_i son]]_j

The same analysis can apply to base-generated gaps as well. That is, instead of having an overt pronoun as the head of DP, the structure contains the null pronominal *pro* in the head D position, and ellipsis of its complement NP. For example, (128) corresponds with the interpretation associated with the full NP-complement deletion, as in (129).

- (128) na-ge Zhangsan hen nanguo yinwei me-ge
 that-CL Zhangsan very sad because every-CL
 fuqin_i dou hui bao-zhe e_j shuijiao de ta_i erzi
 father all will hug-ASP sleep REL his son
 de yizhao_j
 PNM picture
 ‘that picture_j of his_i son that Zhangsan was sad because every father_i will hug
 e_j while sleeping.’

- (129) [me-ge fuqin]_i ... [DP *pro*_j [~~NP ta_i erzi de yizhao~~]]_j
 every father_i [DP e_j [NP picture of his_i son]]_j

6 Conclusion

In this chapter, I provide a thorough re-examination of gaps and resumptive pronouns in Chinese argument relative clauses. To begin with, I investigate gaps in relative

clauses in terms of their sensitivity to different island constraints, and to the Weak Crossover Condition, and point out that they exhibit both movement and nonmovement properties. Then, I examine resumptive pronouns in the island conditions, the Weak Crossover Condition and their reconstruction properties in island and non-island contexts. In addition, based on a list of coordinate structure examples that join gapped and resumptive relative clauses, I suggest that the syntactic rules that govern gaps and resumption in Chinese relative clauses may not always be distinctive. More precisely, some gaps may behave like resumptive pronouns, and are derived via base-generation.

Thus, I provide a revised proposal, and argue that while resumption is still derived via base-generation as proposed in Aoun and Li (2003), gaps may be ambiguous between base-generation and movement in general. I do not attempt to explain when a movement pattern is required and when a base-generation pattern is required; however, given the patterns observed above, both must be available analyses for gapped structures in at least some contexts. For example, further development of the idea is still needed to explain why base-generation is not possible in embedded clause contexts, for example, (21-26), but is possible in some less embedded clause contexts, such as (15-16).

The overall implication of the patterns of data examined in this chapter is that while overt resumption is a relatively straightforward pattern, involving base-generation in all contexts, gaps reflect a very complex picture of constraints, with some patterning with movement contexts in other languages and some patterning with

overt resumption. The relative complexity of the syntax of gapped relative clauses, in contrast with the relative straightforwardness of resumptive relative clauses, is clear.

In the next chapter, I suggest how that ambiguity in complexity, particularly in the direct object position, together with some processing properties of resumptive pronouns, can help to provide an explanation for the contrast in acquisition patterns for Chinese gapped and resumptive relative clauses.

CHAPTER 7: TOWARD A UG-BASED PROCESSING ACCOUNT OF THE ACQUISITION PATTERN

1 Introduction

In this chapter, I explore a UG-based processing approach to the acquisition pattern of Chinese relative clauses.³⁴ Before turning to the details, I briefly summarize the experimental results, and discuss the problems faced by the current UG-based proposal, the Structural Distance Hypothesis, and by the non-UG, frequency-based proposal, the Frequency of Exposure Hypothesis. I also discuss the implications for the Bayesian Learning hypotheses, as explained in Chapter 5.

In the experimental study (i.e., Chapter 3), I tested Chinese-speaking children's acquisition, more precisely, the comprehension pattern, of the top three types of relative clauses in the Accessibility Hierarchy (Keenan & Comrie 1977): Subject Gapped relatives, as in (1), Object Gapped relatives, as in (2), and Oblique Object Resumptive relatives, as in (3).

³⁴ Note that UG (Universal Grammar) should not be confused with any kind of learning mechanism that incorporates UG, for example, Yang (2000) or Lidz and Gagliardi (2015). In addition, the current proposal does not serve to propose any kind of learning mechanism (e.g., the mechanism that demonstrates how children learn a language); rather, it is designed to explain the comprehension pattern of Chinese-speaking children in relative clauses obtained in Chapter 3.

(1) Subject Gapped relatives

[___ qin-le maomi de] xiao-gou pao-zou le.
kiss-ASP cat REL little-dog run-away ASP

‘The dog that kissed the cat ran away.’

(2) Object Gapped relatives

[xiao-gou qin-le ___ de] maomi pao-zou le
little-dog kiss-ASP REL cat run-away ASP

‘The cat that the dog kissed ran away.’

(3) Oblique Object Resumptive relatives

[qingwa dui ta tiaowu de] yazi pao-zou le.
frog to it dance REL duck run-away ASP

‘The duck that the frog danced for ran away.’

The results showed that children performed equally well on both Subject Gapped relative clauses and Oblique Object Resumptive relatives, but performed worst on Object Gapped relatives, as summarized in (4) (where “>” means “better than”):

(4) **Children’s Comprehension of RCs**

Subject Gapped RCs = Oblique Object Resumptive RCs > Object Gapped
RCs

These findings suggest two things. First, it is not only hierarchical structure that affects the comprehension of relative clauses, because that would predict that Chinese-speaking children should show differential behaviours across the three types of relatives. Second, the resumption strategy appears to aid the comprehension of relative clauses, since children performed equally well on both Subject Gapped and Oblique Object Resumptive relatives despite the latter being structurally more embedded than the former as well as less frequent in the input.

The Structural Distance Hypothesis, with its prediction of acquisition patterns illustrated in (5), fails to completely account for Chinese-speaking children's comprehension pattern of gapped and resumptive relatives.

(5) **Prediction of Acquisition Patterns by Structural Distance Hypothesis**

Subject Gapped RCs > Object Gapped RCs > Oblique Object Resumptive RCs

Moreover, the results raise at least three questions: (1) If using gaps is the primary strategy to form relative clauses in Chinese, why can the resumption strategy ever outperform the gap strategy in the acquisition of relative clauses? (2) In Keenan and Comrie (1977), the primary RC-forming strategy in the majority of languages is the gap strategy. Therefore, based on the typological data, languages seem to favor a gap strategy. Why then does the resumption strategy in acquisition of Chinese relative clauses neglect this typological preference for gaps? (3) According to Shlonsky (1992), resumptive pronouns serve as a last resort in the grammar when forming A'-

dependencies, so they are not supposed to be used unless wh-movement, i.e., the one typically associated with gaps, fails to form a grammatical construction, for example, in islands. If the resumption strategy is the last resort in grammar, hence only allowed when necessary, then the fact a language allows the resumption strategy is not surprising, but why should a strategy disfavored in the grammar take the lead in acquisition? In other words, the issues at stake concern whether UG principles governing the building of RCs using the gap and the resumption strategies can predict a child's performance, and the answer appears to be negative.

Turning to the Frequency of Exposure Hypothesis, however, the picture is no more promising. The relative frequency of relative clause patterns in the Chinese Treebank, as summarized in (6), suggests that the most frequent Subject Gapped relatives should show a better performance than Object Gapped relatives, followed by Oblique Object Resumptive relatives.

- (6) **Relative Frequency of Relative Clause Patterns in the Chinese Treebank**
Subject Gapped RCs > Object Gapped RCs > Oblique Object Resumptive
RCs

But again, this hypothesis falls short of explaining children's comprehension pattern with Chinese relatives.

Thus, neither the UG-based Structural Distance Hypothesis nor the input-based Frequency of Exposure Hypothesis itself can explain the observed comprehension pattern. In fact, the Structural Distance Hypothesis and the Frequency

of Exposure Hypothesis make the same predictions. For example, since Subject Gapped RCs are more frequent than Object Gapped RCs, both the Structural Distance and Frequency of Exposure hypotheses predict earlier acquisition for the subject relatives — a prediction that is borne out by the observed comprehension pattern in Chinese-speaking children. But as we can see, their predictions both turn out wrong when we bring one more unusual type of relative clause into the discussion, namely, the oblique object relative clause that requires a resumptive pronoun. The main problem lies in the unexpected good performance of Oblique Object Resumptive relatives. In terms of structural distance, the oblique object is farther away from the head noun than subject or direct object from its corresponding head noun. In terms of frequency, the Oblique Object Resumptive relative is the least frequent among the three types of relatives under consideration.

As for the Bayesian approaches, at first they seem to find some support from these experimental results, since Chinese-speaking children can comprehend indirect/oblique relatives well even with very scarce available data in the input. The Bayesian view could be good at explaining why the Subject Gapped RCs and Oblique Object Resumptive RCs are parallel in terms of comprehension, but it still fails to explain why the Object Gapped RCs are not learned as well. If a Bayesian learner can learn from one or two examples of oblique object relatives, it is hard to explain why object relatives cannot be learned well from thousands of object relative examples in the input. In other words, the Bayesian hypothesis does not predict the observed differential learning between Object Gapped RCs and the other two kinds of relatives.

A possible solution, alluded to before, is to recognize the facilitative role of a resumptive pronoun in comprehension — an idea not without precedent. For instance, Rahmany et al. (2014) found that Persian-speaking children’s comprehension of object RCs containing the resumptive elements was better than subject gapped and object gapped RCs. They suggest that resumptive elements aid children’s processing of complex structures because the presence of resumptive elements relieves the processing burden.

But the facilitative property of a resumptive element in comprehension does not seem to be an invariant one. Friedmann and Costa (2011) showed that, whereas hearing-impaired Hebrew children and adolescents’ comprehension of object resumptive relatives was significantly better than object gapped relatives (i.e., 94% vs. 69%), hearing-impaired Palestinian Arabic-speakers’ comprehension of object resumptive relatives was still quite low (i.e., 72%). Friedmann and Costa attribute this between-language difference in resumptive object relatives to language-specific characteristics of resumptive elements along with the assumption that children have a problem with *wh*-movement. In particular, they treat Hebrew resumptive pronouns in objects as strong pronouns and argue that only strong pronouns can serve as a last resort to help save a structure associated with the problematic *wh*-movement. By contrast, Arabic’s resumptive clitics in objects are not considered strong pronouns and hence cannot be used to save a structure associated with *wh*-movement. Furthermore, previous studies on adults’ processing suggest that parsing relatives varies depending on the grammatical status of a resumption strategy in a given language or construction. (e.g., Chacón 2015; Keshev & Meltzer-Asscher 2016).

Taken together, these studies suggest that resumption does not necessarily facilitate processing, and that its processing properties rely on both within-language and cross-language properties. But before discussing the processing proposal, it is worth pointing out an exegetical problem with the original assumptions about UG's role in acquisition. Within the framework of generative grammar, the fact that we can understand sentences, especially ones involving Auxiliary-inversion, as in (7) or the antecedent-anaphor relation, as in (8), suggests that the computational system of language employs minimal structural distance instead of minimal linear distance (e.g., Chomsky 1965; Chomsky 2011; Berwick et al. 2011).³⁵

(7) Can eagles that fly swim?

(Chomsky 2011: 271)

(8) Who do they expect to see each other next week?

(Chomsky 2011: 273)

The propensity of a language learner to rely on structural distance but not on linear order is also called structure-dependence principle. It is the main part of the poverty of stimulus argument, and is what this dissertation sets off to investigate. So far I have considered the influence of structural distance on acquisition, but the potential influence on learning a language (i.e., I-language as in Chomsky 2005) from interactions between UG and other cognitive systems, for example, the third factor

³⁵ But see O'Grady (2012) for a non-UG processing account of the same phenomenon.

(e.g., Chomsky 2005) or a parser linking to the acquisition of syntax (e.g., Omaki & Lidz 2015), has remained unexplored.³⁶ Moreover, Berwick and Chomsky (2011) suggest that several phenomena such as existential sentences (Chomsky 2001), word order, or some principles assumed in the Principles and Parameters approach (P&P, aka Government and Binding, see Chomsky 1981, 1982), including the Theta Criterion, parameterization, and the cause of diversity of languages are not part of UG, but instead reflect *externalization* of internal computation to the sensory-motor system.³⁷ Accordingly, when this view is applied to acquisition patterns, we expect to find data that are not entirely attributed to UG, but to externalization, including the interactions between UG and the processing system (e.g., also see Lidz & Gagliardi 2015; Omaki & Lidz 2015 for discussions on how the development of a parser links to the acquisition of syntax).

³⁶ Regarding the language faculty as an “organ of the body,” Chomsky (2005) proposes three factors of language design that can interact to determine the I-language (i.e., “the notion of structure,” an “internalized language” (e.g., Chomsky 1986)). These three factors are:

1. Genetic endowment, apparently nearly uniform for the species, which interprets part of the environment as linguistic experience, a nontrivial task that the infant carries out reflexively, and which determines the general course of the development of the language faculty. Among the genetic elements, some may impose computational limitations that disappear in a regular way through genetically timed maturation. Kenneth Wexler and his associates have provided compelling evidence of their existence in the growth of language, thus providing empirical evidence for what Wexler (to appear) calls “Lenne-berg’s dream.”
2. Experience, which leads to variation, within a fairly narrow range, as in the case of other subsystems of the human capacity and the organism generally.
3. Principles not specific to the faculty of language.

(Chomsky 2005: 6)

³⁷ According to Berwick and Chomsky (2011), the internal mental objects (e.g., phrases, sentences, etc.) that are constructed by the recursive generative procedure are related to two interfaces with language-external systems: the system of thought and the sensory-motor system that externalizes internal computations and thoughts.

This is the position that I am taking here. I consider a proposal that suggests that the “comprehension pattern” of relatives in Chinese-speaking children is influenced by the interaction between UG (i.e., lexical items and the ways that structured expressions are generated) and the parser. As the parser is shown to be sensitive to language-specific characteristics (e.g., Grillo & Costa 2014; Grillo et al. 2015; Chacón 2015; Keshev & Meltzer-Asscher 2016), I review studies on how the parser behaves in response to language-specific characteristics in Section 2. In Section 3, I review two studies on relative clause processing in Chinese. In Section 4, I formalize the intuition of interactions between UG principles and the parser and account for the Chinese-speaking children’s comprehension pattern. Section 5 concludes this chapter.

2 Parsing Relatives and Language-Specific Characteristics

In this section, I review three studies on adults’ processing of relative clauses using resumption. The results from these studies bring out two crucial factors that can affect processing of relatives. One factor is the grammatical status of resumption in a given language. Both Section 2.2 and Section 2.3 together demonstrate the influence of resumption’s grammatical status on relative clause processing. The other factor is dependency length, and I discuss it in Section 2.3.

Before moving onto the next section, here I briefly describe the processing properties of the so-called “filler-gap dependencies” associated with relatives or wh-questions in order to prepare for the following discussion on processing resumptive dependencies. In a relative clause like (9), “gap” refers to the empty position that is

relativized, and “filler” refers to the phrase, *the dog*, that controls the interpretation of the gap.

(9) the dog that ___ kissed the cat

Several previous studies have found that the parser follows an “active filler strategy” when resolving filler-gap dependencies (e.g., Frazier 1987; Frazier & Flores d’Arcais 1989; Frazier & Clifton 1989, among many others). According to this strategy, when the parser has identified a filler, it will rank the option of assigning it to a gap above other options. In other words, the parser does not look for other options but gaps when it sees a filler. Thus, this strategy will predict that, if a lexical NP argument occurs in the position of a tempting gap for an active filler, the parser will slow down.

For example, to test if the active filler hypothesis makes correct predictions, Frazier and Clifton (1989) used self-paced reading and end-of-sentence acceptability judgment tasks to investigate comprehension of sentences with filler-gap dependencies. In a self-reading task, they used materials such as (10a-d) to identify if the parser immediately searches for gaps or waits when identifying a filler (*where the ^ marks indicate the division into presentation segments*).

- (10) a. Who did the housekeeper[^]from Germany[^]urge[^]the guests[^]to consider?[^]
- b. Who did the housekeeper[^]say she[^]urged[^]the guests[^]to consider?[^]

- c. The housekeeper[^]from Germany[^]urged[^]the guests[^]to consider[^]the new chef.[^]
- d. The housekeeper[^]said she[^]urged[^]the guests[^]to consider[^]the new chef.[^]

(Frazier & Clifton 1989: 97)

The idea of the task is that, if the parser is indeed actively looking for a gap, then the occurrence of a lexical NP *the guests* in the potential gap position in (10a) and (10b) will slow down reading time when compared to the declarative controls (10c-d). This is also called the filled-gap effect, and is caused by immediate reanalysis.

The results showed that the segment following the lexical NP received significantly slower reading times for the question conditions, i.e., the ones associated with filler-gap dependencies, compared to the declarative conditions, hence supporting the active filler hypothesis: The parser actively searches for gaps upon detecting a filler in a filler-gap construction.

When it comes to relatives using resumptive pronouns, questions arise as to whether the parser still uses a similar strategy, for example, considering the resumptive pronoun as a replacement of gap, and actively searches for a resumptive pronoun (as if it knows that gap position is now a resumptive pronoun) when encountering a filler. To answer this question, I review Keshev and Meltzer-Asscher's (2016) study on processing Hebrew resumption In Section 2.1, and review Chacón's (2015) study on processing English resumption In Section 2.2. The results

from both studies suggest that the grammatical status of a resumption strategy in a given language or condition plays a role in processing long-distance dependencies.

2.1 Parsing Grammaticalized Resumption

Keshev and Meltzer-Asscher (2016) investigate if the parser prefers to predict a resumptive pronoun for an early resolution as it enters an island in Hebrew, since Hebrew uses both gaps and resumptive pronouns in relative clause formation. They compared whether there is a difference between grammatical and ungrammatical resumption in terms of active dependency formation, as previous studies investigating languages containing “intrusive” resumption—resumption not considered grammatical—showed that the active filler strategy was blocked within islands (e.g., Stowe 1986; Traxler & Pickering 1996).

First, they indicated that resumption in certain islands in Hebrew is grammatical, such as the Complex NP Condition, but is ungrammatical inside other islands, such as the Coordinate Structure Condition. They used an acceptability judgment task to confirm that resumption does exhibit such a grammaticality contrast between the Complex NP Condition and the Coordinate Structure Condition.

In the second experiment, Keshev and Meltzer-Asscher used this graded island sensitivity of resumption to test if the parser can predict an early resumptive pronoun inside the Complex NP Condition but not inside the Coordinate Structure Condition, so that an active dependency formation will be made and thus exhibit a filled-gap effect in the former but not in the latter condition. They used a self-paced reading task with a filled-gap design to manipulate the potential site of gap/resumptive pronoun in these two island conditions and outside islands. The

results showed that reading times in critical words in the Complex NP Condition were longer than those in the Coordinate Structure Condition, suggesting a filled-gap effect in the former but not in the latter.

Hence, Keshev and Meltzer-Asscher conclude that the processing of filler-gap dependencies in Hebrew varies according to the grammatical status of resumptive pronouns inside different islands. Furthermore, they propose that ‘the parser is “aware” of grammatical constraints in the language’ (p. 18). For example, the parser will avoid postulating gaps in ungrammatical positions such as within certain islands. Similarly, the parser will behave differently depending on the grammatical status of resumptive pronouns when they occur inside different islands.

2.2 *Parsing Ungrammatical Resumption*

Chacón (2015) examines how resumption dependencies in English are constructed, especially when inside islands, as in (11).

(11) These are the things that we don’t know what they are.

(Chacón (2015: 109))

He suggests that there are three possible hypotheses for the comprehension system to construct a resumption dependency in real time. The first one is the “end of sentence hypothesis”, and it states that resumption dependencies are constructed only when the comprehension system does not detect any possible grammatical gap position, which it can generally only be sure of when it reaches the end of the sentence. The second hypothesis is the “filler-driven hypothesis.” It states that, as there is an open filler, the

comprehension system forms a resumption dependency immediately upon seeing a pronoun. The third hypothesis is the “pronoun-driven hypothesis”. The resumption dependency is created only when the pronoun recognizes the head of an open filler-gap dependency as its antecedent. On this view, encountering a pronoun induces the parser to search for an antecedent.

The study manipulated whether the filler was the only choice of antecedent available, or whether multiple potential antecedents were available; the pronoun-driven hypothesis predicts that dependency formation should be improved when the filler is the only available antecedent.

He used a series of offline sentence fragment completion tasks and an eye-tracking study to examine which of these three hypotheses better explains how resumptive dependencies are constructed in English, and the evidence consistently supports the pronoun-driven hypothesis.

The overall findings indicate that, although the resumptive pronouns inside islands appear to be acceptable in English, resumption dependencies are not constructed in an active fashion, which contrasts with Keshev and Meltzer-Asscher’s (2016) findings in Hebrew. Moreover, he attributed this processing difference to the grammaticality contrast of resumption “inside islands” between two languages. That is, Hebrew allows grammatical resumptive pronouns inside islands, but English does not have grammatical resumptive pronouns inside islands (or outside islands).

Taken more broadly, his study shows that the active dependency is “grammatically sensitive”; when we change what is grammatical, we will also see active dependency formation changes. This also predicts that, even in Hebrew, when

a construction does not allow resumptive pronouns, the processing effect should be the same as with ungrammatical resumptive pronouns in English.³⁸

2.3 *Length of Dependency and Parsing Ungrammatical Resumption*

Hofmeister and Norcliffe (2013) conducted two experimental studies to examine whether both acceptability ratings and processing difficulty caused by resumptive pronouns in English vary with dependency length. Experiment 1 uses the acceptability judgment task, where sentences are presented word-by-word to English native speakers. This experiment involves two independent variables: dependency length (long vs. short) and resumption (gap vs. pronoun), as illustrated in (12a-d) (adapted from Hofmeister & Norcliffe 2013: 229):

- (12) a. Mary confirmed that there was a prisoner who the prison officials had acknowledged that the guard helped __ to make a daring escape.
- b. Mary confirmed that there was a prisoner who the prison officials had acknowledged that the guard helped him to make a daring escape.
- c. The prison officials had acknowledged that there was a prisoner that the guard helped __ to make a daring escape.
- d. The prison officials had acknowledged that there was a prisoner that the guard helped him to make a daring escape.

³⁸ I also thank Chacón (p.c., 2016) for clarifying some important points for me when I was writing this section.

In the long-dependency condition, the head noun, *a prisoner*, is separated from the relativized position by both a relative clause and a complement clause. In the short-dependency condition, the head noun is separated from the relativized position only by a single relative clause.

The results showed that adults' acceptability judgments were significantly lower in the long-dependency condition than in the short-dependency condition. Their judgments with sentences containing resumptive pronouns were also lower than the ones with gaps. But there is a significant interaction between dependency length and resumption, which further suggests that the acceptability penalty of sentences with resumptive pronouns is smaller in the long-dependency condition. Hence, although resumptive pronouns never sound better than gaps in English, even in long-dependency sentences, the fact that they induced a uniform effect across sentences with different dependency-lengths suggests that increased embedding depth reduces the disadvantage of resumptive pronouns. This leads to their second experiment, which examines the online processing of relative clauses containing resumptive pronouns.

In Experiment 2, they used the self-paced, moving window display paradigm to extract participants' reading times of the same materials from Experiment 1. The results showed that resumption received significantly faster reading rates than gaps at the tail of the dependency structure. In particular, this faster average reading rate is caused exclusively by the long-dependency condition. In addition, processing at the tail of dependency structures containing gaps slows down in the long-dependency condition.

Thus, Hofmeister and Norcliffe conclude that resumptive pronouns can facilitate processing in high difficulty contexts, but there is no evidence suggesting that resumptive pronouns have a processing advantage over gaps in less difficult contexts.

In sum, based on the findings from the above studies, parsing resumption dependencies is not only sensitive to language-specific properties, such as the grammatical status of resumption strategy in a construction, but can also be sensitive to dependency length, as found in the ungrammatical resumption in English. In the next section, I review two relevant studies on processing relative clauses in Chinese.

3 Parsing Chinese Relatives

Lin (2006) conducted a self-paced moving-window experiment to investigate the processing of subject and object relatives in Chinese. Two variables were manipulated in the experiment: RC type (subject/object of the extraction site) and the grammatical function of the modified noun (subject-modifying/object-modifying) in a sentence. The following examples illustrate all the four conditions, as indicated by the labels (from Lin 2006: 123-124):

(13) Subject-modifying Subject RC

gouyin	yuanzhang	de	shaonü	zhuang-dao	le	yi	yuan
seduce	dean		REL	young.lady	bump-into	ASP	congressman

‘The young lady that seduced the dean bumped into the congressman.’

(14) Subject-modifying Object RC

yuanzhang gouyin de shaonü zhuang-dao le yiyuan
dean seduce REL young.lady bump-into ASP congressman
'The young lady that the dean seduced bumped into the congressman.'

(15) Object-modifying Subject RC

yiyuan zhuang-dao le gouyin yuanzhang de shaonü
congressman bump-into ASP seduce dean REL young.lady
'The congressman bumped into the young lady that seduced the dean.'

(16) Object-modifying Object RC

yiyuan zhuang-dao le yuanzhang gouyin de shaonü
congressman bump-into ASP dean seduce REL young.lady
'The congressman bumped into the young lady that the dean seduced.'

The results showed that subject RCs were easier to comprehend than object RCs and that object-modifying RCs were more difficult to comprehend than subject-modifying RCs. More interestingly, he found that the construction of the filler-gap relation did not start until the arrival of the head noun, which suggests that the parser utilizes the same strategy, namely, a gap-searching process, in Chinese relatives as it does in languages with post-nominal RCs.

Ning (2008) used both the grammaticality judgment task and the self-paced moving-window task to examine Chinese speakers' processing of relatives using gaps

and resumptive pronouns. To begin with, she conducted an on-line grammaticality judgment experiment to examine whether resumptive pronouns have the same grammatical status as gaps in five types of relatives, including: subject relatives (17), long-distance subject relatives (18), object relatives (19), genitive relatives (20), and oblique object relatives (21) (data from Ning 2008: 82-84).³⁹

(17) Subject relatives (gap vs. resumptive pronoun)

- a. souji yuliao de na-wei xuezhe fabiao-guo ji-pian
collect corpus REL that-CL scholar publish-ASP several-CL
wenzhang.
article
'That scholar who collected corpus published several articles.'
- b. ta souji yuliao de na-wei xuezhe fabiao-guo ji-pian
he collect corpus REL that-CL scholar publish-ASP several-CL
wenzhang.
article
Lit. 'That scholar who he collected corpus published several articles.'

³⁹ In Ning's (2008) study, the examples of long-distance subject relatives refer to those that contain two additional (VP-level) adjuncts between the gap/resumptive pronoun and the verb in the RC.

(18) Long-distance subject relatives

- a. xinkudi zai yanjiushi zuanyan jibing wenti
hard in research.office delve.into disease problem
de na-wei yisheng chenzui zai yanjiu
REL that-CL doctor indulge in research
lequ zhong.
pleasure in

‘The doctor who delved hard into the disease problem in the research office indulged in the pleasure of doing research.’

- b. ta xinkudi zai yanjiushi zuanyan jibing
he hard in research.office delve.into disease
wenti de na-wei yisheng chenzui zai
problem REL that-CL doctor indulge in
yanjiu lequ zhong.
research pleasure in

Lit. ‘The doctor who he delved hard into the disease problem in the research office indulged in the pleasure of doing research.’

(19) Object relatives

- a. xuexiao yaoqing de na-wei zuojia ceng huode yiwendajiang.
school invite RELthat-CL authoronce obtain literary.award
‘The author whom the school invited obtained the literary award before.’

- b. xuexiao yaoqing ta de na-wei zuojia ceng huode
 school invite he REL that-CL authoronce obtain
 yiwendajiang.
 literary.award

Lit. 'The author whom the school invited him obtained the literary award before.'

(20) Genitive relatives

- a. nuer fabiao xinshu de na-wei fuqin
 daughter publish new.book REL that-CL father
 zao chengwei zhongshi de duzhe.
 alreadybecome faithful PNM reader

'That father whose daughter published a new book has already become a faithful reader.'

- a. ta nuer fabiao xinshu de na-wei
 he daughter publish new.book REL that-CL
 fuqin zao chengwei zhongshi de duzhe.
 father alreadybecome faithful PNM reader

Lit. 'That father who his daughter published a new book has already become a faithful reader.'

(21) Oblique object relatives

- a. yisheng ti kaiyao de na-wei huanzhe
 doctor for prescribe REL that-CL patient

 bu guanxin ziji de shenti.
 NEG care.about self POSS body

 ‘That patient whom the doctor prescribed for did not care his body.’
- b. yisheng ti ta kaiyao de na-wei huanzhe
 doctor for he prescribe REL that-CL patient

 bu guanxin ziji de shenti.
 NEG care.about self POSS body

 Lit. ‘That patient whom the doctor prescribed for him did not care
 about his own body.’

The results showed that Chinese speakers demonstrated a significant preference for gaps over resumptive pronouns in subject relatives, long-distance subject relatives, and object relatives, while in oblique object relatives, they showed a significant preference for resumptive pronouns and in genitive relatives they showed no significant preference for either strategy.⁴⁰

Since oblique object relatives containing gaps were judged ungrammatical in the judgment task, Ning excluded this type from the self-paced reading comprehension task. Overall, the results from the reading task showed that the parser

⁴⁰ At first Ning (2008) did not find a significant difference between resumptive pronouns and gaps in the oblique object relatives. But after she realized that there were some ambiguous examples, which led to misinterpretations, she excluded those examples and then found a significant difference between the two strategies.

spent more time on reading relatives with resumptive pronouns than relatives with gaps, even though this difference did not reach significance in subject and genitive relatives. Specifically, for long-distance subject relatives, the reading time was significantly longer for relatives with resumptive pronouns than relatives with gaps *after* the head nouns (i.e., *na-wei yisheng* ‘that doctor’ in (18a-b)) were parsed. In object relatives, resumptive relatives took longer time to read than gapped relatives in the head noun positions.

Based on the results from the grammaticality judgment and self-paced reading experiments, Ning concluded that for the parser, 1) obligatory resumptive pronouns, e.g., in oblique object relatives, must always be the preferred strategy, and 2) optional gaps, e.g., in subject relatives, long-distance subject relatives, object relatives, will be the preferred strategy since resumptive pronouns will likely lead to a garden path effect, namely, reanalysis phenomenon, and 3) there may be no preference for either strategy in relatives when structural complexity such as embeddedness increases, e.g., genitive relatives.

In brief, these studies indicate two findings that are relevant to the current processing proposal. First, subject relatives are easier to parse than object relatives. Second, when a structure is grammatical with a resumptive pronoun and ungrammatical with a gap, the resumptive pronoun is preferred over the gap. This is a confirmation of the phenomenon that the parser’s behaviour in relative clauses is sensitive to language-specific properties, especially with regard to whether a position is grammatical with a resumptive pronoun or not.

4 A UG-Based Processing Proposal of Comprehension Patterns

In this section, I lay out a UG-based processing proposal to account for the Chinese-speaking children’s comprehension pattern by drawing evidence from the above studies.

First, if the grammatical resumption can be processed in an active dependency formation as gaps are (e.g., Keshev & Meltzer-Asscher 2016; Chacón 2015), then the fact that Chinese allows grammatical resumption in relative clause formation suggests that that the comprehension system may not necessarily distinguish the resumption strategy from gap strategy in any intrinsic way, regardless of the former being used as a secondary strategy within and across languages.⁴¹

Second, the distributional contrast between gaps and resumptive pronouns in Chinese relative clauses, as in Table 26, suggests that there is a tendency for relativized positions that are less embedded, i.e., higher in the syntactic structure, to use gaps, such as subjects and objects in matrix clauses, and for relativized positions that are more embedded, i.e., lower in the syntactic structure, to use (grammaticized) resumptive pronouns, such as oblique objects in the matrix clause and arguments in further embedded clauses.

Table 26: Distribution of Gaps and Resumptive Pronouns (RPs) in Chinese

Argument RCs

Matrix Clause				Embedded Clause	
Subject	Object	Oblique Object	Possessor	Subject	Object
Gap	Gap	RP	Gap/RP	Gap/RP	Gap/RP

⁴¹ This, however, is a hypothesis that still needs to be supported by empirical evidence.

In other words, even though both gaps and resumptive pronouns are considered grammatical strategies in forming Chinese relative clauses, the higher positions in the syntactic structure prefer gaps, whereas the lower positions either prefer resumptive pronouns, as in oblique objects, or show no specific preference, as in possessors, embedded subjects and objects.⁴² This tendency can also be schematized as in Figure 3 below:

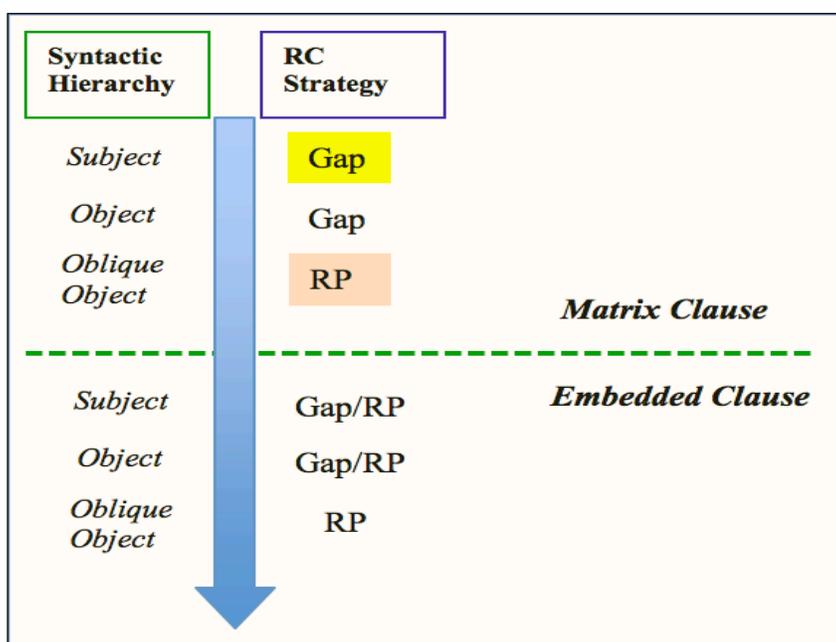


Figure 3: Distribution/Tendency of Chinese Gap/RP Strategy in the Hierarchical Structure

Thus, Chinese-speaking children may leverage this language-specific tendency in their acquisition of relative clauses. This makes it easier to learn subject relatives

⁴² That being said, this empirical question can be explored by examining Chinese speakers' preference (or more specifically, acceptance) rate of gaps/resumptive pronouns in different positions (see e.g., Ning 2008). In addition, since there is a possibility that the input frequency of gaps and resumptive pronouns in different argument positions might match this tendency, it is not clear whether this distributional tendency is guided by the constraints in UG or by input frequency. I will leave it for future research.

using gaps (i.e., taking advantage of the highest position) AND oblique object relatives using resumptive pronouns (i.e., meriting from the point when resumptive pronouns “start” being used).⁴³ As for the relativized object position, however, it does not appear to benefit directly from this tendency as much as a subject with a gap or an oblique object with a resumptive pronoun.

Third, as discussed in Chapter 6, resumptive pronouns are unanimously grammatical inside all island-type positions, and also in more embedded positions that are not islands. In contrast to resumptive pronouns, gaps are sometimes grammatical and sometimes ungrammatical when they occur in the island-type configuration involving only a single layer of embedding. When occurring in an embedded clause which is itself further embedded in an island-type configuration, almost all gaps are ungrammatical, except the crossover configurations.

Thus, this entire pattern appears to be associated with a processing advantage for resumption, since it is grammatical in more embedded positions and in all island-type positions. That is, comprehending dependencies of greater length or dependencies into complex island configurations appears to be aided by resumption (e.g., Hofmeister & Norcliffe 2013; Beltrama & Xiang 2016). Specifically, adopting Hofmeister and Norcliffe’s suggestion for English resumption and also Ning’s (2008) observation in Chinese, Chinese resumption may also live a double life — one aiding dependency processing when the structure gets more complicated and the other

⁴³ The tendency that the highest subject needs to occur with gaps but not resumptive pronouns can also be subsumed under a known constraint in many languages that allow resumptive pronouns: The Highest Subject Restriction (e.g., McCloskey 1990 on Irish; Borer 1984 on Hebrew; Shlonsky 1992 on Hebrew and Palestinian Arabic and his similar comments on Welsh, also see Rouveret 2011, where he considers this restriction to be a general property of resumptive pronouns).

serving as a requirement of grammar. Perhaps the fact that Chinese-speaking children's better comprehension of oblique object relatives using resumptive pronouns than gapped subject and object relatives suggests that resumption in the former facilitates processing. Nevertheless, to determine when resumption can serve as a processing aid and when it cannot in a language still requires a thorough investigation.

Fourth, in Chapter 6, I demonstrate that gaps in Chinese relative clauses reflect a very complex picture of constraints, with some patterning with movement in some contexts and some patterning with nonmovement in other contexts, whereas overt resumption is a relatively straightforward pattern, involving base-generation in all contexts. Therefore, I suggest that the syntactic ambiguity that exists in gapped relatives makes it more difficult to learn gapped relatives than resumptive relatives. When Chinese-speaking children are learning gapped relative clauses, they may need to discover which structure is assigned to which interpretation, and a similar decision-making process occurs when they are trying to comprehend gapped relative clauses. Whereas gapped subject relatives can still benefit from being in the highest argument position, gapped object relatives not only suffer the effects induced by their longer dependency, but also suffer from being associated with the ambiguity of gaps.⁴⁴

⁴⁴ Again, this is a hypothesis that requires testing and further research. But if this idea is on the right track and predicts that it is the structural ambiguity of gapped relatives in Chinese that makes them more difficult to comprehend than resumptive relatives, then perhaps the same story can be applied to previous findings in Hebrew-speaking children's comprehension of gapped and resumptive relatives. According to Sichel (2014), while gapped relatives in Hebrew are associated with movement, relatives using optional resumptive pronouns are ambiguous between movement and nonmovement structures. The structural ambiguity of resumptive relatives in Hebrew may be part of the reasons why studies like Friedmann et al. (2009) or Arnon (2005) did not find a significantly better performance on object resumptive relatives than object gapped relatives in Hebrew-speaking children's comprehension.

In brief, resumption itself does not necessarily facilitate comprehension or learning, but must be taken together with language-specific properties such as the hierarchical structure, the grammaticality of particular structures, and also the distributional cues in order to yield the best result. This kind of cooperative interpretation of data works well in accounting for the acquisition of Chinese relatives. Children performed equally well on oblique object resumptive relatives and subject gapped relatives, since both merit from the cooperative work of grammar and processor. By contrast, object gapped relatives do not receive a boost from either the grammar or processor, and may actually suffer relative penalties from both, hence showing the worst performance in children's acquisition.

5 Conclusions

In this chapter, I started with a discussion of the Structural Distance Hypothesis and the Frequency of Exposure Hypothesis in terms of their predictions for Chinese-speaking children's comprehension pattern of relative clauses. I pointed out that neither the UG-based Structural Distance Hypothesis nor the input-based Frequency of Exposure Hypothesis itself can explain the observed comprehension pattern, and that the main problem lies in the unexpected good performance of Oblique Object Resumptive relatives. Thus, I explored a UG-based processing approach by drawing evidence from previous studies on the interactions between processing and language-specific properties of resumption. On the one hand, children can rely on the grammatical status of resumption to actively construct resumption dependencies as in Hebrew (Keshev & Meltzer-Asscher 2016; Chacón 2015). On the other hand, they use the processing tendencies of gaps and resumption to aid in determining the

correct parse of the associated structure (e.g., Ning 2008). In other words, resumption itself does not necessarily facilitate comprehension or learning, but it needs to work together with language-specific properties such as the hierarchical structure, the grammaticality of structures, and the distributional cues in order to yield the best result in learning.

APPENDICES

APPENDIX A: SUMMARY OF THE PILOT STUDY

Because it was difficult to find native Mandarin-speaking children in Tucson, USA, the pilot study was conducted using the English-version materials with English-speaking children in Tucson during 2012 and 2013. Their ages ranged from 3 to 5-year-old.

The vocabulary used in the pilot study includes the animal names and action verbs that are generally recognizable to children younger than 2 year-old according to the MacArthur Communicative Development Inventory norms (Fenson et al. 1993). The current experimental items that are modified based on the pilot study include:

- (1) the layout of all sentences in each story;
- (2) the number of sentences in each story (e.g., originally there were 6 test sentences per story and the total number of sentences of an experiment was 48);
- (3) some action verbs that are easier to be recognized by the preschoolers in Taiwan (double-checked with the school teachers before running the experiment);
- (4) other improvements (including how the experimenters delivered stories, how the experimenters take notes, and when to ask follow-up questions, etc.).

Overall, the pilot study with the English-speaking children ran smoothly. The average time that each participant spent in the experiment was 20 through 25 minutes.

APPENDIX B: MATERIALS USED IN THE COMPREHENSION EXPERIMENT

Pre-test

- (1) laohu gen maer zai yiqi chi hanbao.
tiger and horse ASP together eat hamburger
'The tiger and the horse were eating the hamburger together.'
- (2) zhe-zhi yu zai chi hanbao.
this-CL fish ASP eat hamburger
'The fish was eating a hamburger.'
- (3) houzi pao-zou yihou, zhe-zhi yu pao-lai chi
monkey run-away after, this-CL fish run-come eat
pingguo.
apple
'After the monkey ran away, the fish ate the apple.'
- (4) maer pao-zou le.
horse run-away ASP
'The horse ran away.'

Story 1:

- (1) you yi-zhi xiaogou zai wazi-shang tiaowu.
 have one-CL dog in sock-above dance
 ‘A dog was dancing on the sock.’
- (2) you yi-zhi maomi zai wazi-shang tiaowu.
 have one-CL cat in sock-on dance
 ‘A cat was dancing on the sock.’
- (3) qin-le maomi de xiaogou pao-zou le.
 kiss-ASP cat REL dog run-away ASP
 ‘The dog that kissed the cat ran away.’
- (4) zhe-zhi maomi zai chi yu.
 this-CL cat ASP eat fish
 ‘This cat was eating a fish.’
- (5) xiaogou qin-le de maomi pao-zou le.
 dog kiss-ASP REL cat run-away ASP
 ‘The cat that the dog kissed ran away.’

Story 2:

- (1) you yi-zhi yazi zai yaolan-li shuijiao.
 have one-CL duck in cradle-in sleep
 ‘A duck was sleeping in the playpen/cradle.’
- (2) you yi-zhi qingwa zai yaolan-li shuijiao.
 have one-CL frog in cradle-in sleep
 ‘A frog was sleeping in the playpen/cradle.’
- (3) dui yazi tiaowu de qingwa pao-zou le.
 to duck dance REL frog run-away ASP
 ‘The frog that danced for the duck ran away.’
- (4) zhe-zhi yazi zai zhu shucaitang.
 this-CL duck ASP cook vegetable-soup
 ‘This duck was cooking vegetable soup.’
- (5) qingwa dui ta tiaowu de yazi pao-zou le.
 frog to it dance REL duck run-away ASP
 ‘The duck that the frog danced for ran away.’

Story 3:

- (1) you yi-zhi tuzi zai chi xiangjiao.
 have one-CL rabbit ASP eat banana
 ‘A rabbit was eating a banana.’
- (2) you yi-zhi shizi zai chi xiangjiao.
 have one-CL lion ASP eat banana
 ‘A lion was eating a banana.’
- (3) shizi pai-le de tuzi pao-zou le.
 lion pat-ASP REL rabbit run-away ASP
 ‘The rabbit that the lion patted ran away.’
- (4) zhe-zhi shizi zai kai-che.
 this-CL lion ASP drive
 ‘This lion was riding on a car.’
- (5) pai-le tuzi de shizi pao-zou le.
 pat-ASP rabbit REL lion run-away ASP

‘The lion that patted the rabbit ran away.’

Story 4:

- (1) you yi-zhi daxiang zai chui dizi.
have one-CL elephant ASP blow flute
‘An elephant was playing the flute.’
- (2) you yi-zhi xongmao zai chui dizi.
have one-CL panda ASP blow flute
‘A panda was playing the flute.’
- (3) xongmao dui ta hui-shou de daxiang pao-zou
panda to it wave-hand REL elephant run-away
le.
ASP
‘The elephant that the panda waved at ran away.’
- (4) zhe-zhi daxiang zai kan shu.
this-CL elephant ASP read book
‘This elephant was reading a book.’
- (5) dui daxiang hui-shou de xongmao pao-zou
to elephant wave-hand REL panda run-away
le.
ASP
‘The panda that waved at the elephant ran away.’

Story 5:

- (1) you yi-zhi maomi zai qi jiaotache.
have one-CL cat ASP ride bike
‘A cat was riding a bike.’
- (2) you yi-zhi xiaogou zai qi jiaotache.
have one-CL dog ASP ride bike
‘A dog was riding a bike.’
- (3) qin-le maomi de xiaogou pao-zou le.
kiss-ASP cat REL dog run-away ASP
‘The dog that kissed the cat ran away.’
- (4) zhe-zhi xiaogou zai chi dangao.
this-CL dog ASP eat cake
‘This cat was eating a cake.’
- (5) xiaogou qin-le de maomi pao-zou le.
dog kiss-ASP REL cat run-away ASP
‘The cat that the dog kissed ran away.’

Story 6:

- (1) you yi-zhi yazi zai maozi-shang tiao-lai-tiao-qu.
have one-CL duck in hat-on jump-come-jump-go
‘A duck was jumping on a hat.’
- (2) you yi-zhi qingwa zai maozi-shang tiao-lai-tiao-qu.
have one-CL frog in hat-on jump-come-jump-go
‘A frog was jumping on a hat.’
- (3) dui yazi tiaowu de qingwa pao-zou le.

- to duck dance REL frog run-away ASP
 ‘The frog that danced for the duck ran away.’
- (4) zhe-zhi yazi zai shua toufa.
 this-CL duck ASP brush hair
 ‘This duck was brushing his hair.’
- (5) qingwa dui ta tiaowu de yazi pao-zou le.
 frog to it dance REL duck run-away ASP
 ‘The duck that the frog danced for ran away.’

Story 7:

- (1) you yi-zhi tuzi zai tui pisa.
 have one-CL rabbit ASP push pizza
 ‘A rabbit was pushing a pizza.’
- (2) you yi-zhi shizi zai tui pisa.
 have one-CL lion ASP push pizza
 ‘A lion was pushing a pizza.’
- (3) shizi pai-le de tuzi pao-zou le.
 lion pat-ASP REL rabbit run-away ASP
 ‘The rabbit that the lion patted ran away.’
- (4) zhe-zhi shizi zuo zai shu-shang.
 this-CL lion sit in tree-on
 ‘This lion was sitting in the tree.’
- (5) pai-le tuzi de shizi pao-zou le.
 pat-ASP rabbit REL lion run-away ASP
 ‘The lion that patted the rabbit ran away.’

Story 8:

- (1) you yi-zhi xongmao zai xi panzi.
 have one-CL panda ASP wash plate
 ‘A panda was washing a plate.’
- (2) you yi-zhi daxiang zai xi panzi.
 have one-CL elephant ASP wash plate
 ‘An elephant was washing a plate.’
- (3) xongmao dui ta hui-shou de daxiang pao-zou
 panda to it wave-hand REL elephant run-away
 le.
 ASP
 ‘The elephant that the panda waved at ran away.’
- (4) zhe-zhi xongmao zai wan xiezi.
 this-CL panda ASP play shoe
 ‘This panda was playing with a shoe.’
- (5) dui daxiang hui-shou de xongmao pao-zou
 to elephant wave-hand REL panda run-away
 le.
 ASP
 ‘The panda that waved at the elephant ran away.’

APPENDIX C: CORRECT YES(Y)/NO(N) ANSWERS TO EACH SENTENCE IN EACH STORY

<i>Sentence /Story</i>	<i>Story 1</i>	<i>Story 2</i>	<i>Story 3</i>	<i>Story 4</i>	<i>Story 5</i>	<i>Story 6</i>	<i>Story 7</i>	<i>Story 8</i>
1 (Filler)	Y	N	Y	N	Y	N	Y	N
2 (Filler)	N	Y	N	Y	N	Y	N	Y
3 (Target)	Y	Y	N	N	N	N	Y	Y
4 (Filler)	Y	N	Y	N	Y	N	Y	N
5 (Target)	N	N	Y	Y	Y	Y	N	N

**Total Y/N number: 20/20*

**Total Y/N number of target sentences: 8/8*

**Total Y/N number of filler sentences: 12/12*

APPENDIX D: PERCENTAGE ACCURACY OF BOTH TYPES OF RESPONSES IN FILLERS

<i>response/age group</i>	<i>younger</i>	<i>middle</i>	<i>older</i>
Yes	98%	99%	99%
No	92%	97%	94%

APPENDIX E: TREGEX SEARCH PATTERNS FOR ADJUNCT RELATIVE CLAUSES

Adjuncts	Ext (extent)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-EXT < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	TPC (topic)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (/^NP.*-TPC/ < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	LOC (locative)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-LOC < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	TMP (temporal)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-TMP < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	NP-DIR (direction)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-DIR < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	NP-MNR (manner)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-MNR < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	NP-PRP (purpose or reason)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-PRP < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP
	ADV (adverbial)	@NP < (CP < (/^WH.*-[0-9]+)/#1%i < (-NONE- < *OP*)) < (CP << (NP-ADV < (-NONE- < /^*T*-([0-9]+)/#1%i)) < DEC)) < NP

APPENDIX F: RAW COUNTS AND FREQUENCY OF ADJUNCT RELATIVE CLAUSES

Type of Relative Clause (Treebank Tag and its indication)	Raw counts	Frequency (out of 24,837 relatives)
Ext (extent)	33	0.13
TPC (topic)	1350	5.44
LOC (locative)	49	0.2
TMP (temporal)	99	0.4
NP-DIR (direction)	7	0.03
NP-MNR (manner)	107	0.43
NP-PRP (purpose or reason)	38	0.15
ADV (adverbial)	31	0.12
Total	1714	6.9

REFERENCES

- Alexopoulou, T. (2006). Resumption in relative clauses. *Natural Language & Linguistic Theory*, 24(1), 57-111.
- Arnon, I. (2005). Relative clause acquisition in Hebrew: towards a processing-oriented account. In A. Brugos, M. R. Clark-Cotton & S Ha (Eds.), *Proceedings of the 29th Annual Boston Conference on Language Development* (pp. 37-48). Somerville, MA: Cascadilla Press.
- Aoun, J., Choueiri, L., & Benmamoun, E. (2010). *The Syntax of Arabic*. New York: Cambridge University Press.
- Aoun, J., Choueiri, L., & Hornstein, N. (2001). Resumption, movement, and derivational economy. *Linguistic Inquiry*, 32(3), 371-403.
- Aoun, J., & Li, Y. A. (2003). *Essays on the Representational and Derivational Nature of Grammar: The Diversity of Wh-constructions*. Cambridge, Mass: MIT Press.
- Beltrama, A. & M. Xiang. (2016) Unacceptable but comprehensible: the facilitation effects of resumptive pronouns. Unpublished Manuscript. (to appear in *Glossa*).
- Berwick, R. C. & Chomsky N. (2011). The biolinguistic program: the current state of its development. In Di Sciullo, A-M. & C., Boeckx (Eds.), *The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of Human Language Faculty* (pp. 19–41). Oxford: Oxford University Press.
- Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, 35(7), 1207-1242.

- Biloa, E. (1990). Resumptive pronouns in Tuki. *Studies in African Linguistics*, 21(2), 211.
- Borer, H. (1984). Restrictive relatives in modern Hebrew. *Natural Language & Linguistic Theory*, 2(2), 219-260.
- Botwinik, I., Bshara, R., & Armon-Lotem, S. (2015). Children's production of relative clauses in Palestinian Arabic: Unique errors and their movement account. *Lingua: International Review of General Linguistics*, 156, 40-56.
- Brame, M. K. (1968). A new analysis of the relative clause: Evidence for an interpretive theory. Unpublished Manuscript, MIT.
- Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, Mass: Harvard University Press.
- Carlson, G. N. (1977). Amount relatives. *Language*, 53(3), 520-542.
- Carnie, A. (2010). *Constituent structure* (2nd ed.). Oxford;New York;: Oxford University Press.
- Chacón, D. A. (2015). *Comparative Psychosyntax*. Doctoral dissertation, University of Maryland, College Park.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, 10(7), 335-344.
- Chen, J., & Shirai, Y. (2015). The acquisition of relative clauses in spontaneous child speech in Mandarin Chinese. *Journal of Child Language*, 42(2), 394.
- Chen, Z., Grove, K. & Hale, J. (2012). Structural expectations in Chinese relative clause comprehension. In J. Choi et al. (Eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics* (pp. 29-37). Somerville, MA:

Cascadilla Press.

- Cheng, Lisa L.-S. & Sybesma, R. (2005). A Chinese relative. In Broekhuis, H. et al. (Eds.), *Organizing Grammar: Studies in Honor of Henk van Riemsdijk* (pp. 69-76). New York;Berlin;: Mouton de Gruyter.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1977). On wh-movement. In P. W. Culicover, T. Wasow & A. Akmanjian (Eds.), *Formal Syntax* (pp. 71-132). New York: Academic Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht, the Netherlands: Foris Publications.
- Chomsky, N. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, Mass: MIT Press.
- Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origins, and Use*. New York: Praeger.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (2001). Derivation by Phase. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 1-52). Cambridge, Mass: MIT Press.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1-22.
- Chomsky, N. (2011). Language and other cognitive systems. What is special about language? *Language Learning and Development*, 7(4), 263-278.
- Chung, S., & McCloskey, J. (1987). Government, barriers, and small clauses in modern Irish. *Linguistic Inquiry*, 18(2), 173-237.
- Corrêa, L. M. S. (1995). An alternative assessment of children's comprehension of

- relative clauses. *Journal of Psycholinguistic Research*, 24(3), 183-203.
- Costa, J., Friedmann, N., Silva, C., & Yachini, M. (2014). The boy that the chef cooked: Acquisition of PP relatives in European Portuguese and Hebrew. *Lingua: International Review of General Linguistics*, 150, 386.
- Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, Mass: MIT Press.
- Crain, S., Thornton, R., Boster, C., Conway, L., Lillo-Martin, D., & Woodams, E. (1996). Quantification without qualification. *Language Acquisition*, 5(2), 83-153.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25(2), 108-127.
- Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, 81(4), 882-906.
- Elbourne, P. (2001). E-type anaphora as NP-deletion. *Natural Language Semantics*, 9(3), 241-288.
- Engdahl, E. (1985). Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, 23(1), 3-44.
- Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1993). *The MacArthur Communicative Development Inventories: user's guide and technical manual*. San Diego: Singular Publishing Group.
- Frazier, L. (1987). Syntactic Processing: Evidence from Dutch. *Natural Language &*

- Linguistic Theory*, 5(4), 519-559.
- Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4(2), 93-126.
- Frazier, L., & Flores d'Arcais, G.,B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, 28(3), 331.
- Friedl, J. E. F., (2006). *Mastering Regular Expressions* (3rd ed.). Sebastopol, CA: O'Reilly.
- Friedmann, N., & Costa, J. (2011). Last resort and no resort: Resumptive pronouns in Hebrew and Palestinian Arabic hearing impairment. In A. Rouveret (Ed.), *Resumptive Pronouns at the Interfaces* (pp. 223-240). Amsterdam, NLD: John Benjamins Publishing Company.
- Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119(1), 67-88.
- Georgopoulos, C. (1985). Variables in Palauan syntax. *Natural Language & Linguistic Theory*, 3(1), 59-94.
- Gerken, L., Dawson, C., Chatila, R., & Tenenbaum, J. (2015). Surprise! Infants consider possible bases of generalization for a single input example. *Developmental Science*, 18(1), 80-89.
- Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15(3), 289-316.
- Gordon, P. (1996). The truth-value judgment task. In D. McDaniel, C. McKee and H. S. Cairns (Eds.), *Methods for Assessing Children's Syntax* (pp. 211-231). Cambridge, Mass: MIT Press.

- Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of Cognition. (2008). In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology*. Cambridge, United Kingdom: Cambridge University Press.
- Grillo, N., & Costa, J. (2014). A novel argument for the universality of parsing principles. *Cognition*, *133*(1), 156-187.
- Grillo, N., Costa, J., Fernandes, B., & Santi, A. (2015). Highs and lows in English attachment. *Cognition*, *144*, 116-122.
- Gu, G. 2001. (2001). *A Study of Resumptive Pronouns*. Doctoral dissertation, The Chinese University of Hong Kong, Hong Kong.
- Guilliot, N. & N. Malkawi. (2006). When resumption determines reconstruction. In D. Baumer, D. Montero & M. Scanlon (Eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics* (pp.168-176). Somerville, MA: Cascadilla Press.
- Guilliot, N. & N. Malkawi. (2007). Reconstruction without movement. In L. Eguren & O. Fernández-Soriano (Eds.), *Coreference, Modality, and Focus: Studies on the Syntax-Semantics Interface* (pp. 111-131). Amsterdam: John Benjamins Publishing Company.
- Haegeman, L. M. V. (1991). *Introduction to Government and Binding Theory* (2nd ed.). Cambridge, Mass.: Blackwell.
- Hamburger, H. & S. Crain. (1982). Relative acquisition. In S. A. Kuczaj (Ed.), *Language Development. Vol. 1, Syntax and Semantics* (pp. 245-274). Hillsdale NJ: Lawrence Erlbaum Associates.

- Han, C., Lidz, J., & Musolino, J. (2007). V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry*, 38(1), 1-47.
- Hendrick, R. (1988). *Anaphora in Celtic and Universal Grammar*. Dordrecht: Kluwer.
- Hendriks, P., & Koster, C. (2010). Production/comprehension asymmetries in language acquisition. *Lingua*, 120(8), 1887-1897.
- Hofmeister, P. & Norcliffe, E. (2013). Does resumption facilitate sentence comprehension? In P. Hofmeister & E. Norcliffe (Eds.), *The Core and the Periphery: Data-Driven Perspectives on Syntax Inspired by Ivan A. Sag* (pp. 225-246. Stanford, CA: CSLI Publications.
- Hsiao, F., & Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, 90(1), 3-27.
- Hsiao, Y., & MacDonald, M. C. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology*, 4, 767.
- Hsu, C. N., Hermon, G., & Zukowski, A. (2009). Young children's production of head-final relative clauses: Elicited production data from Chinese children. *Journal of East Asian Linguistics*, 18(4), 323-360.
- Hu, S., Gavarro, A., & Guasti, M. (2016). Children's production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics*, 37(2), 323-346.
- Hu, S., Gavarró, A., Vernice, M., & Guasti, M. T. (2016). The acquisition of

- Chinese relative clauses: Contrasting two theoretical approaches. *Journal of Child Language*, 43(1), 1-21.
- Huang, C.-T. James. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Huang, C. - J. (1984). On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531-574.
- Huang, C. - J. (1987). Remarks on empty categories in Chinese. *Linguistic Inquiry*, 18(2), 321-337.
- Huang, C.-T. James. (1989). Pro-drop in Chinese: a generalized control theory. In O. Jaeggli & K. J. Safir (Eds.), *The Null Subject Parameter*, ed., 185-214. Dordrecht: Kluwer.
- Huang, C. J., Li, Y. A., & Li, Y. (2009). *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, 45(3), 337-374.
- Jäger, L., Chen, Z., Li, Q., Lin, C. C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, 79-80, 97-120.
- Karimi, S. (2001). Persian complex DPs: How mysterious are they? *Canadian Journal of Linguistics* 46: 63-96.
- Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass: MIT Press.
- Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1), 63-99.

- Kemp, C., Perfors, A., & Tenenbaum, J. B. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, 10(3), 307-321.
- Keshev, M. & Meltzer-Asscher, A. (2016). Active dependency formation in islands: How grammatical resumption affects sentence processing. Manuscript submitted for publication.
- Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, 22(6), 860-897.
- Kidd, E., Lieven, E. V. M., & Tomasello, M. (2006). Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development*, 21(2), 93-107.
- Kidd, E., Lieven, E. V. M., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences*, 32(1), 132-142.
- Lau, E. (2016). The role of resumptive pronouns in Cantonese relative clause acquisition. *First Language*, 36(4), 355-382.
- Lee, H.-T. T. (1992). The inadequacy of processing heuristics – evidence from relative clause acquisition in Mandarin Chinese. In T. Lee (Ed.), *Research on Chinese Linguistics in Hong Kong* (pp. 47-85). Hong Kong: The Linguistic Society of Hong Kong.
- Li, Y. A. (1990). *Order and Constituency in Mandarin Chinese*. Dordrecht: Kluwer.
- Lidz, J. & Gagliardi, A. (2015) How Nature Meets Nurture: Universal Grammar and

- Statistical Learning. *Annual Review of Linguistics* 1, 333-353.
- Lin, C.-J. C. (2006). *Grammar and Parsing: a Typological Investigation of Relative- Clause Processing*. Doctoral dissertation, University of Arizona, Tucson.
- Lu, P.-T. (2012). *Structures of the Mandarin Jiao Constructions*. MA thesis. National Tsing Hua University, Hsinchu.
- MacWhinney, B. (2000). *The CHILDES project: Tools for Analyzing Talk* (3rd ed.). Mahwah, N.J.; Lawrence Erlbaum.
- Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, 20(1), 121-136.
- McCloskey, J. (1979). *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Dordrecht: Springer Netherlands.
- McCloskey, J. (1985). The modern Irish double relative and syntactic binding. *Ériu*, 36, 45-84.
- McCloskey, J. (1990). Resumptive pronouns, A-bar binding, and levels of representation in Irish. In R. Hendrick (Ed.), *Syntax of the Modern Celtic Languages* (pp. 199-248). San Diego: Academic Press.
- McCloskey, J. (2002). Resumption, successive cyclicity, and the locality of operations. In S. D. Epstein & T. D. Seeley (Eds.), *Derivation and Explanation in the Minimalist Program* (pp. 184-226). London: Blackwell.
- McCloskey, J. (2005). Resumption. In M. Everaert & Henk van Riemsdijk (Eds.), *The Blackwell Companion to Syntax* (pp. 94-117). Mass;Oxford;: Blackwell

Publishing.

- McCloskey, J. (2011). Resumptive pronouns, A'-binding, and levels of representations in Irish. In A. Rouveret (Ed.), *Resumptive Pronouns at the Interfaces* (pp. 65-119). Amsterdam, NLD: John Benjamins Publishing Company.
- McDaniel, D., & Cowart, W. (1999). Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition*, 70(2), B15-B24.
- McKee, C., McDaniel, D., & Snedeker, J. (1998). Relatives children say. *Journal of Psycholinguistic Research*, 27(5), 573-596.
- Min, R.-F. (1994). The Acquisition of Referring Expressions by Young Chinese Children: A Longitudinal Study of the Forms and Functions of Early Noun Phrases. Unpublished Doctoral dissertation, Catholic University of Nijmegen, Nijmegen, the Netherlands.
- Munn, A. (1994). A minimalist account of reconstruction asymmetries. In M. González (Ed.), *Proceedings of NELS 24* (pp. 397-410), Amherst. University of Massachusetts, GLSA.
- Ning, C. (1993). *The Overt Syntax of Relativization and Topicalization in Chinese*. Doctoral dissertation, University of California, Irvine.
- Ning, L.-H. (2008). The Grammar and Processing of Resumptive Pronouns in Chinese Relative Clauses. MA thesis. National Chengchi University, Taipei.
- O'Grady, W. (2012). Three factors in the design and acquisition of language. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(5), 493-499.
- Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic

- knowledge. *Language Acquisition*, 22(2), 158-192.
- Omaki, A., White, I. D., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: Children's incremental interpretation in English and Japanese wh-questions. *Language Learning and Development*, 10(3), 206.
- Perfors, A., Tenenbaum, J. B., & Regier, T. (2006). Poverty of the stimulus? A rational approach. *28th Annual Conference of the Cognitive Science Society*.
- Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 118(3), 306-338.
- Perfors, A., Tenenbaum, J. B., & Wonnacott, E. (2010). Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language*, 37(3), 607-642.
- Rahmany, R., Marefat, H., & Kidd, E. (2014). Resumptive elements aid comprehension of object relative clauses: Evidence from Persian. *Journal of Child Language*, 41(4), 937-948.
- Regier, T., & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition*, 93(2), 147-155.
- Ross, J. R. 1938. (1967). *Constraints On Variables in Syntax*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Rouveret, A. (2011). Some issues in the theory of resumption: A perspective on early and recent research. In Rouveret (Ed.), *Resumptive Pronouns at the Interfaces* (pp. 1-62). Amsterdam, NLD: John Benjamins Publishing Company.
- Safir, K. (1998). Reconstruction and Bound anaphora: Copy Theory without Deletion

- at LF. Unpublished Manuscript, Rutgers University, New Brunswick, New Jersey.
- Sauerland, U. (1998). *On the Making and Meaning of Chains*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Sauerland, U. 2000. Two structures for English restrictive relative clauses. In M. Saito et al. (Eds.), *Proceedings of the Nanzan GLOW* (pp. 351-366). Nanzan University, Nagoya, Japan.
- Schachter, P. (1973). Focus and relativization. *Language*, 49(1), 19-46.
- Sells, P. 1984. (1984). *Syntax and Semantics of Resumptive Pronouns*. Doctoral dissertation, University of Massachusetts, Amherst.
- Shlonsky, U. (1992). Resumptive pronouns as a last resort. *Linguistic Inquiry*, 23(3), 443-468.
- Sichel, I. (2014). Resumptive pronouns and competition. *Linguistic Inquiry*, 45(4), 655-693.
- Stowe, L. (1986). Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, 1(3), 227-245.
- Su, Y-C. (2004). Relatives of Mandarin children. Paper presented in the 2004 Generative Approaches to Language Acquisition in North America, the University of Hawaii at Manoa.
- Tallerman, M. (1983). Island constraints in Welsh. *York Papers in Linguistics 10*: 197-204.
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24(4), 629-640.

- Theakston, A.L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, 28(1), 127-152.
- Theakston, A.L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language*, 31(1), 61-99.
- Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, 35(3), 454-475.
- Tsai, H.-C. J. (2008). On gapless relative clauses in Chinese. *Nanzan Linguistics: Special Issue 5*: 109-124.
- Tsai W.-T. D. (1997). On the absence of island effects. *Tsing Hua Journal of Chinese Studies*, New Series 27:125–149.
- Tuller, L. A. (1986). *Bijjective Relations in Universal Grammar and the Syntax of Hausa (Africa)*. Doctoral dissertation, UCLA.
- Vergnaud, J. R. (1974). *French Relative Clauses*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Wasow, T. (1972). *Anaphoric Relations in English*. Doctoral dissertation, MIT, Cambridge, Massachusetts.
- Wei, J.-H., T.-H. J. Lin & H.-Y. C. Huang. (2014). Wh-in-situ and QR in Mandarin generic sentences. In R. E. Santana-LaBarge (Ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics* (pp. 484-492). Somerville, MA: Cascadilla Press.

- Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, 56(3), 165-209.
- Xu, L. (1986). Free empty category. *Linguistic Inquiry*, 17(1), 75-93.
- Xu, Yu. (2009). *The Syntax, Processing and Second Language Acquisition of Chinese Relative Clauses*. Doctoral dissertation, University of Arizona, Tucson.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245-272.
- Xue, N., Xia, F., Chiou, F., & Palmer, M. (2005). The Penn Chinese TreeBank: Phrase structure annotation of a large corpus. *Natural Language Engineering*, 11(2), 207-238.
- Zaenen, A., Engdahl, E., & Maling, J. M. (1981). Resumptive pronouns can be syntactically bound. *Linguistic Inquiry*, 12(4), 679-682.
- Zhang, N. (2008). Gapless relative clauses as clausal licensors of relational nouns. *Language and Linguistics* 9, 1005-1028.
- Wu, F. (2009). *Factors affecting relative clause processing in Mandarin*. Doctoral dissertation, University of Southern California, LA.
- Yang, C. D. (2000). *Knowledge and Learning in Natural Language*. Doctoral dissertation, MIT, Cambridge, Massachusetts.